Policy with stochastic hysteresis

Georgii Riabov and Aleh Tsyvinski[*]

April 16, 2021

Abstract

The paper develops a general methodology for analyzing policies with path-dependency (hysteresis) in stochastic models with forward looking optimizing agents. Our main application is a macro-climate model with a path-dependent climate externality. We derive in closed form the dynamics of the optimal Pigouvian tax, that is, its drift and diffusion coefficients. The dynamics of the present marginal damages is given by the recently developed functional Itô formula. The dynamics of the conditional expectation process of the future marginal damages is given by a new total derivative formula that we prove. The total derivative formula represents the evolution of the conditional expectation process as a sum of the expected dynamics of hysteresis with respect to time, a form of a time derivative, and the expected dynamics of hysteresis with the shocks to the trajectory of the stochastic process, a form of a stochastic derivative. We then generalize the results. First, we propose a general class of hysteresis functionals that permits significant tractability. Second, we characterize in closed form the dynamics of the stochastic hysteresis elasticity that represents the change in the whole optimal policy process with an introduction of small hysteresis effects. Third, we determine the optimal policy process.

1 Introduction

Modern macroeconomics is stochastic and is built on recursive methods. The actions in these models are history dependent but the past is represented by a small number of finite-dimensional state variables. For example, Ljungqvist and Sargent (2018) write that "finding a recursive way to handle history dependence is a major achievement of the past 35 years and an important methodological theme". They call these developments an "imperialistic response of dynamic programming" to Prescott's (1977) critique of the impossibility of using dynamic programming in government policy design. In this paper, we aim to study environments in which it may in general be difficult or impossible to write a recursive representation. We use the term hysteresis to describe settings in which the trajectory of policies or actions may generally affect the structure of the environment. That is, the effects of policies are path-dependent.

Our main economic application is a generalization of the macroeconomy-climate model of Golosov, Hassler, Krusell, and Tsyvinski (2014) to the path-dependent climate externalities. There are two main reasons to consider hysteresis in such models. First, there is important recent evidence from climate sciences that a number of climate variables show significant hysteresis behavior. Intergovernmental Panel on Climate Change (IPCC) lists, for example, hysteresis that is present in (1) vegetation change; (2) changes in the ice sheets; (3) ocean acidification, deep ocean warming and associated sea level rise; (4) models of the feedback between the ocean and the ice sheets; and (5) models of the Atlantic meridional overturning circulation (Collins, et al. 2013 in IPCC Fifth Assessment Report – AR5). While cumulative carbon emission is an important benchmark used in macro-climate models, recent research in climate sciences (further discussed in Section 3.2) points to importance of hysteresis in a number of climate variables that constitute a natural science foundation of these models. Moreover, the path-dependency in these models is often rather sophisticated as the exact sequence of the events may matter. The IPCC Special Report further concludes that path dependence of carbon budgets "remains an important knowledge gap" (Rogelj et al. 2018). Second, one of the central themes in the recent developments in the economics literature on climate change is incorporating path-dependency. A recent survey by Aghion, et al. (2019) summarizes these developments and argues that path-dependency in which both history and expectations matter is one of the core insights in this literature (see, for example, Acemoglu, Aghion, Bursztyn, and Hemous (2012), Acemoglu, Akcigit, Hanley, and Kerr (2016), Acemoglu, Aghion, Barrage and Hemous (2019), and Aghion, Dechezlepretre, Hemous, Martin, and Van Reenen (2016)). Finally,

careful consideration of uncertainty has been another important theme of recent research on economics of climate change. Our application thus puts at the forefront the analysis of an economy where a forward looking social planner makes emission decisions under uncertainty and faces a path-dependent emissions externality. In contrast to the literature and, specifically to Golosov, Hassler, Krusell, and Tsyvinski (2014), we significantly relax the assumption of how the previous emission choices enter the climate damages. In our setting, the whole path of the emission (hysteresis) matters as opposed to a depreciated stock of the previous emissions. Specifically, we analyze a particular climate hysteresis functional that captures the most important elements of the setting with general hysteresis.

Our main result in this macro-climate application is a closed form characterization of the dynamics of the marginal externality damage which is also the Pigouvian tax that implements the optimal allocation. We start by deriving the first-order conditions for the optimum and showing that the marginal externality damage from emissions is comprised of the marginal contemporaneous damages and the conditional expectation of the cumulative future damages. Both of these terms are path-dependent and their dynamics cannot be derived using the usual Itô formula which applies only to functions of the current state and not the functionals of the trajectories. We use two different sets of tools to provide the dynamics of the present and the expected cumulative damages.

The present marginal damages are already represented as an explicit path-dependent functional and we use recently developed functional Itô calculus for the non-anticipative functionals (Dupire 2009, 2019, Cont and Fournie 2013) to represent this term. The functional Itô formula uses two new concepts of derivatives introduced by Dupire (2009, 2019). The vertical derivative is an analogue of the space derivative and the horizontal derivative is an analogue of the time derivative for the functionals. These derivatives evaluate, respectively, the effects of a discontinuous bump in the underlying trajectory and the effects of a time extension of the trajectory while fixing the terminal position. The Dupire's functional Itô formula then gives the dynamics of the present damages with the drift determined by the horizontal and the second vertical derivative, and the diffusion coefficient given by the vertical derivative of the marginal present damages. Using the notion of the horizontal and vertical derivatives the functional Itô formula allows us to convey the similar intuition as in the case of no hysteresis for which the usual Itô formula is applicable. However, in contrast with the case of no hysteresis path-dependency may lead to significant effects on the dynamics even in the case of absent contemporaneous damages.

The most challenging part of the paper is to characterize the dynamics of the expected future

marginal damages represented by the conditional expectation process. The difficulty comes from the fact that with passage of time both the damages themselves and the information set (filtration) are changing. The main new theoretical tool that we develop – the total derivative formula for conditional expectation processes – allows us to characterize the dynamics of such processes. Moreover, this formula is broadly applicable to a variety of macroeconomic settings where the conditional expectations processes are central. There are two terms in the semimartingale decomposition of the conditional expectations process that the total derivative formula delivers. The first term can be intuitively thought of as a time derivative and represents how the conditional expectation of the future marginal damages evolves with respect to time. The second term, given as a conditional expectation of the Malliavin derivative of the cumulative future marginal damages, can be thought of as a stochastic derivative with respect to the underlying process and represents how the conditional expectation changes with the changes in the stochastic process. It is useful to compare the results to the case with no hysteresis. The main difference is that in the case of no hysteresis, the policy does not have future effects. With hysteresis, the contemporaneous effects of the policy may be very small but the expected future effects of the policy may be very large – small actions today my have significant future consequences. Moreover, these effects change with both time and the stochastic shocks.

Combining the dynamics of the present and the conditional expectation of the future marginal damages, we obtain in closed form the drift and diffusion of the optimal Pigouvian tax correcting the path-dependent externality. The dynamics are given by the semimartingale decomposition of the marginal externality damages which can be though of in terms of the planner changing the tax with the passage of time and with the realization of uncertainty. Alternatively, this can be thought of as determination of what features of the model matter to the first order for the drift and diffusion of the optimal tax. Specifically, those are given by the functional Itô formula and by our total derivative formula for the conditional expectation processes. Even if there are no direct contemporaneous effects, the dynamics of the optimal policy may be significantly impacted by either the past choices or by the future effects. In other words, both the past and the future non-trivially matter.

We then show that one can significantly generalize the application of the climate hysteresis. First, we propose a very general class of hysteresis functionals that allow significant tractability of the analysis of optimal policies. The Fréchet derivatives of these functionals have two sources of variation: an instantaneous influence in the given (current) period and an integral influence of

perturbations in previous periods. The assumption is mild and only requires that the derivative of a functional at a given time is absolutely continuous up to that period and may have an atom at the present time. We show that every Fréchet differentiable functional is a pointwise limit of functionals from this class. Importantly, these functionals have an attractive property of conveniently separating the effects of the past and the present. As we are considering a very general environment with path-dependence, it may be difficult or impossible to use dynamic programming to write a recursive representation of the problem. Instead, we directly find the first order conditions for the optimal problem by considering all local variations of the policies. The assumption on the derivative of the hysteresis functional allows us to conveniently separate the first order condition in two parts: the marginal effects of the policy on current period and the conditional expectation of the cumulative future marginal effects of policy paths in the future periods.

Our next goal is to explicitly find the evolution equation for the stochastic hysteresis elasticity. This concept of stochastic hysteresis elasticity is similar to the usual concept of elasticity as it captures the change in the variables following a small change in policy. The difference is that the stochastic elasticity is the change in the whole optimal process. Alternatively, one can think of the stochastic elasticity as the asymptotics of the optimal policy when there is an infinitesimal additional hysteresis. Specifically, we show that the stochastic elasticity is an Itô process and write its explicit semimartingale decomposition. A fundamental property of the solution is that both of the terms – the present and the expected future marginal effects of policy paths – in the first order condition for optimality and in the formula for the stochastic elasticity are path-dependent. Similarly to our climate hysteresis application we use the functional Itô formula and the total derivative formula to represent the dynamics of the stochastic elasticity. The semimartingale decomposition of the present and the cumulative marginal effects of policies show that the stochastic elasticity, that is, the change in the optimal policy process following introduction of hysteresis, is an Itô process. The difference with the no-hysteresis case is that now the drift and diffusion coefficients are themselves functionals of the trajectory of the policy. That is, hysteresis leads to the path-dependent stochastic process of the elasticity for which we calculate the drift and diffusion coefficients in the explicit form. Finally, we return to the analysis of the optimal policies with general hysteresis. The evolution of the optimal policy can be analyzed using exactly the same tools as the stochastic elasticity. The explicit form we obtained for the stochastic elasticity thus can be thought of as a characterization of the optimal policy process for small hysteresis, or small hysteresis asymptotics. We use the functional Itô formula to represent the current marginal effects and the total derivative formula to

represent the conditional expectation process of the future marginal effects of policies. The drift and diffusion coefficients of the optimal policy are then determined implicitly by the equations having a similar form as the explicit solution for the stochastic elasticity but also depending on the process of the optimal policy itself.

We finally provide several examples of how to calculate the stochastic hysteresis elasticity. All of these examples are straightforward applications of the main semimartingale decomposition formula that we developed.

Literature

The closest to our work is a sequence of papers by Borovička, Hansen, and Scheinkman (2014), Borovička, Hansen, Hendricks, and Scheinkman (2011), and Borovička and Hansen (2016). They define the concept of the shock elasticity, relate it to impulse responses familiar to macroeconomists, and, importantly, formalize it using Malliavin derivatives. The shock elasticity is an impulse response of the pricing kernel to the marginal perturbation of the underlying stochastic cash flow process. Our work goes beyond their results and delivers the explicit semimartingale decomposition of the evolution of the stochastic elasticity for the path-dependent functionals. In Section 7.3 we also provide a connection to the Clark-Ocone formula that features prominently in Borovička, Hansen, and Scheinkman (2014).

The only paper that we are aware of that uses functional Itô formula in economic settings is an insightful work of Cvitanić, Possamaï, and Touzi (2017). In the dynamic moral hazard setting where the agent controls the volatility of the process, they show how that without loss of generality the set of admissible contracts can be represented as the path-dependent processes arising from the functional Itô formula. In our environment, both the functional Itô formula and the total derivative formula are applied. We discuss this in more details in Section 7.1.

The concept of stochastic elasticity is similar to the analysis of the models of investment under uncertainty for the case of small sunk costs or for models with small menu costs. In those models, there is a form of hysteresis where due to the region of inaction, a temporary change may have a permanent effect. The analysis of such models have important implications for the large amounts of inertia that arise in the models of small menu costs such as in Akerlof and Yellen (1985) and Mankiw (1985). Dixit (1991) develops a method of analytical approximation for such hysteresis models. Reis (2006) and Alvarez, Lippi, and Paciello (2011, 2016) derive a similar analytical characterization in the models with inattentive producers. These are essentially the same as our

calculation of stochastic elasticity for the environments that can be characterized by Itô formula. A series of papers by Alvarez and Lippi (2019), Alvarez, Le Bihan, and Lippi (2016) and Alvarez, Lippi, and Oskolkov (2020) focus on the cumulative output effects of shock in environments with the menu and adjustment costs frictions. These cumulative output effects are the output impulse response functions[1] that are similar to the future expected marginal effects of policy in our paper. Our analysis can be thought of as generalizing the results to the environments in which there is path-dependence.

An important early paper by Detemple and Zapatero (1991) studies an asset pricing model with habit formation using Malliavin calculus.[2] Huang and Sockin (2018) study optimal dynamic taxation in continuous time and calculate impulse response to misreporting using Malliavin derivatives. As Sannikov (2014) for the model of moral hazard they connect their analysis to the finance literature on the sensitivity of the "Greeks" of the options (see, e.g., Fournié, Lasry, Lebuchoux, Lions, and Touzi 1999). Their analysis relates to our calculation of Malliavin derivatives of the policy process in the last part of Section 6.4.

Hysteresis or path-dependence is important in a variety of macroeconomic and other settings. The natural rate of unemployment may directly depend on the previous path of unemployment and, hence, on the path of monetary policy (Blanchard and Summers 1986). Examples of recent explicit focus on path dependency and monetary policy is Berger, Milbradt, Tourre, and Vavra (2018), Gali (2020), and Jorda, Singh, and Taylor (2020). Hysteresis is also central in the models of secular stagnation as in Summers (2014) and Eggertsson, Mehrotra, and Robbins (2019). Adjustment costs in investment generate hysteresis (for example, see a survey in Dixit 1992). Acemoglu, Egorov, and Sonin (2020) exposit a wide variety of political economy models with path-dependence of policies or institutions. Egorov and Sonin (2020) additionally discuss models in which there is non-Markovian dependence on the sequence of policy events. Page (2006) classifies types of history dependence and distinguishes between the path-dependent policies (where the exact sequence of events matters) and the "phat" dependent strategies (where the events but not their sequence matters). The models of increasing returns such as Arthur (1989) and David (1985), and a review in Arrow (2000) feature potentially very complicated dependence on the paths.

---

[1]Similarly, the impulse responses play an important role in Pavan, Segal, and Toikka (2014), Garrett and Pavan (2012), and Bergemann and Strack (2015) in dynamic mechanism design, Makris and Pavan (2017) in the dynamic taxation, and Makris and Pavan (2018) who connect these two literatures. See also a review in Pavan (2017). In general, the recursive formulation of a number of these problems may be challenging and the tools we develop may prove to be useful in those contexts.

[2]See also Serrat (2001), Bhamra and Uppal (2009) and Cvitanic and Malamud (2009).

We now briefly discuss the relationship of the economic environment we consider with mathematical literature. The analysis of deterministic dynamic systems with hysteresis is well established (e.g., Krasnosel'skii and Pokrovskii 2013). In our model, path dependency arises from the optimization of the forward looking agents under uncertainty rather than being directly posed as a property of a dynamic system. The problem that we study also does not fit easily into the standard framework of the infinite-dimensional stochastic optimal control such as Fabbri, Gozzi, and Swiech (2017), see discussion in Section 7.4. The reason is that the hysteresis functional depends on the whole path of the policy. If we treat the policy as a control process we are obliged to take the space of paths (of variable length) as a control space which leads to the analysis of path-dependent HJB equations the understanding of which is in its nascency.

2   Environment

We describe an environment of a stochastic problem with forward looking agents and path-dependency.

2.1   Baseline setting

Let $w = (w_t)_{0 \le t \le T}$ be a Brownian motion.[3] Let the objective of the planner be described by the quadratic loss function:

$$\max_c \mathbb{E} \int_0^T -\frac{1}{2} (c_t - w_t)^2 \, dt. \tag{1}$$

Here maximization is taken over all (progressively measurable) policy processes $c : [0,T] \times C([0,T]) \to \mathbb{R}$. That is, policy c_t depends on information about w up to moment t.

The basic problem is easily solved by interchanging integral and expectation and yields the optimal process c_t^* :

$$c_t^* = w_t. \tag{2}$$

The optimal policy simply tracks the Brownian motion w_t.

2.2   General environment

Consider a problem where there is an additional path-dependent effect of policy h_t:

$$\max_c \mathbb{E} \int_0^T (-\frac{1}{2} (c_t - w_t)^2 - \epsilon h_t(c)) dt, \tag{3}$$

---

[3] All of the results can be straightforwardly extended to more general diffusion processes.

where $h : [0, T] \times C[0, T] \rightarrow \mathbb{R}$ is an adapted functional on the space of trajectories (i.e. $h_t(c)$ is defined only by the restriction of c on $[0, t]$), and ϵ is a parameter. One can think of the functional h as an additional effect (a cost or a benefit) that depends on the history of policy up to time t. The solution of this problem is denoted c^ϵ. We define hysteresis as dependence of the effects of policy $h_t(c)$ on the path of the previous policies $c_{[0,t]} = c^t$.

3 Climate externality

This section generalizes the macro-climate model of Golosov, et al. (2014) to the case where climate externality is path-dependent. Section 3.1 shows how to map some of the key features of that paper to the framework described in the previous section. Section 3.2 discusses evidence on hysteresis in the climate externality. Section 3.3 proposes a specific path-dependent climate hysteresis functional that captures most of the insights of the general framework that we develop later in Section 4. Sections 3.4-3.7 solve in closed form the dynamics of the marginal climate externality or the Pigouvian tax implementing the optimum, that is, determine its drift and diffusion coefficients.

3.1 Planner's problem with a climate externality

We first note that the structure of the problem (3) captures some of the main features of the macroeconomic model with a climate externality of Golosov, et al. (2014). In its essence, a macro-climate model consists of two primary blocks. The first is the specification of the economy which is a standard dynamic stochastic general equilibrium model. The planner maximizes the expected utility of consumption (we assume no discounting)

$$\mathbb{E} \int_0^T U\left(c_t^{\text{cons}}\right) dt,$$

subject to the standard feasibility constraint with capital accumulation, where c_t^{cons} is a consumption good. The production function is given by

$$Y_t = F\left(K_t, E_t, S_t\right),$$

where K_t is capital, E_t is energy consumption (measured in carbon emission units), and S_t is a climate variable at time t. The climate variable is in general a functional $S_t = \tilde{S}_t\left(E^t\right)$ of the

path of emissions $E^t = E_{[0,t]}$ (equation (4) in Golosov, et al. (2014) is, in fact, exactly this general formulation). The climate variable S_t affects the economy via a damage function $D_t(S_t)$ so that the output is reduced multiplicatively $(1 - D_t(S_t)) \times \tilde{F}(K_t, E_t)$. Moreover, the damages are assumed to be exponential $1 - D_t(S_t) = \exp(-\gamma_t S_t)$ for some parameter γ_t and the utility $U(x) = \log(x)$. The essence of the model is how the path of emissions E^t translates to damages to the economy. Let us now reduce the model further to focus on this key dimension. Assume that production does not require capital and the feasibility constraint for the economy is then static:

$$c_t^{cons} = Y_t = e^{-\gamma_t S_t} F(E_t).$$

The planner then maximizes

$$\max_{E_t} \mathbb{E} \int_0^T \left(\tilde{U}(E_t) - \gamma_t S_t \right) dt,$$

which is essentially equivalent to the general problem (3). We further assume that $\tilde{U}(E_t) = -\frac{1}{2}(E_t - w_t)^2$ but this is again done purely for leanness of the model and can be immediately extended.

The key simplification, however, that the literature makes is in simple dependence of the climate variable S_t on the path of the emissions. For example, Golosov, et al. (2014)) assume that

$$S_t = \int_0^t d_s E_s ds,$$

where $d_s \in [0,1]$ is a carbon depreciation rate and then also further structure is placed on d_s. In other words, the climate variable is equal to the stock of the depreciated emissions.

The climate-economy model in its essence reduces to how the previous energy consumption choices and the associated carbon emissions impact today's and future economy. In contrast to the literature, we significantly relax the assumption of how the previous emission choices enter the damages. In our setting, the whole path of the emissions (hysteresis) matters through the functional $h_t(c^t)$ as opposed to a depreciated stock of the previous emissions. We, on purpose, stripped down the climate model to focus only on the general path-dependent effects of the emissions and their interaction with the uncertainty.

3.2 Evidence on climate hysteresis

There are two primary sets of evidence which imply that hysteresis is important in the economic models of climate change.

The first is a large set of recent evidence on importance of hysteresis in climate sciences. The most comprehensive and authoritative source for such research is the Intergovernmental Panel on Climate Change (IPCC). The working group for the physical science basis of the long-term climate change in the Fifth Assessment Report – AR5 (Collins, et al. 2013), while arguing for the attractiveness of the use of the cumulative carbon emission, notes that a number of climate variables and models show significant hysteresis behavior. These are the models of (1) vegetation change; (2) changes in the ice sheets; (3) ocean acidification, deep ocean warming and associated sea level rise; (4) models of the feedback between the ocean and the ice sheets; and (5) models of the Atlantic meridional overturning circulation. The report also argues that "the concepts of climate stabilization and targets is that stabilization of global temperature does not imply stabilization for all aspects of the climate system. For example, some models show significant hysteresis behaviour in the global water cycle, because global precipitation depends on both atmospheric CO_2 and temperature". This is also consistent with the study of Zickfield et al. (2012) who argue that, while total cumulative emissions may be an important approximation for many models, there are a number of exceptions where path-dependency is important: among the variables with timescales of several centuries, such as deep ocean temperature and sea level rise, and for the peak responses of atmospheric CO_2 or for the surface ocean acidity. We now briefly discuss some other recent evidence of the hysteresis behavior in individual components of the climate system. Eliseev et al. (2014) is a comprehensive study of the permafrost behavior that finds significant evidence of hysteresis especially for the higher concentration of greenhouse gases in the atmosphere. Another important aspect of ice thawing is the release of permafrost carbon which is found to be both very significant and highly path-dependent (Gasser et al. 2018). Garbe, et al. (2020) analysis "reveals a strong, multi-step hysteresis behaviour of the Antarctic Ice Sheet" which is potentially reinforced by a number of additional feedback mechanisms. Nordhaus (2019) augments the DICE model with the effects of Greenland ice sheet disintegration and finds that the baseline or the no-policy effects are significantly different when even a simple formulation of the hysteresis is incorporated while the optimal policy results are similar. Boucher et al. (2012) analyzes the response of a number of models to a significant increase in the CO_2 concentration and finds that hysteresis is particularly

pronounced for the terrestrial (such as long lived soil carbon sinks and vegetation) and marine variables (such as the sea-level rise) as well as in global mean precipitation. Nohara et al. (2013) finds that the hysteresis effects are important for the regional climate change which is notable given, for example, the recent focus on the significant differences in optimal climate policy at the regional level by Hassler, et al. (2020). Wu et al. (2012) describe significant hysteresis in the hydrological cycle that leads to one of the most direct impacts of global warming affecting droughts, floods and water supplies. This is important for the recent economic literature that studies the impact of the floods and the sea level (for example, Bakkensen and Barrage 2017, Barrage and Furst 2019; Hong, Wang, and Young, 2020 where hysteresis may be also compounded with the path-dependency in beliefs). Summarizing, while the cumulative emission and its variants are certainly a central benchmark for the development of the climate-economy models, recent research in climate sciences points to importance of hysteresis with sophisticated trajectory dependence in a number of important climate variables.

The second reason for why hysteresis is important lies in the economics part of the climate-economy models. One of the central themes in the recent advances in this literature has been incorporating path-dependency. A recent survey by Aghion, et al. (2019) summarize these developments: "The core insight is that technological innovation is a path-dependent process in which history and expectations matter greatly in determining eventual outcome" leading to "important implications for climate policy design" and "research and knowledge production are path-dependent, deployment of innovations is path-dependent and the incentives for technology adoption create path dependence". Several papers in the literature develop various parts of these insights. In Acemoglu, Aghion, Bursztyn, and Hemous (2012), dirty technologies have an advantage in the market size and in the initial productivity and, hence, there is path-dependency in the direction of innovation and production. One of the most important findings of Acemoglu, Akcigit, Hanley, and Kerr (2016) is that the nature of innovations in clean or dirty technologies is path dependent. Similarly, there is path-dependence in innovation in the model of the consequences of the shale gas revolution by Acemoglu, Aghion, Barrage and Hemous (2019). Grubb, et al. (2020) and Baldwin, Cai, and Kuralbayeva (2020) develop models of path-dependency due to the costs of switching from the dirty technologies. Aghion, Dechezleprêtre, Hemous, Martin, and Van Reenen (2016) provide extensive empirical evidence of path dependence in innovation in auto industry from aggregate spillovers and from the firm's own innovation history. Meng (2016) finds significant evidence for strong path dependence in energy transition for the U.S. electricity sector over the 20th century, focusing on coal.

Fouquet (2016) is a summary of evidence that energy systems are subject to strong and long-lived path dependence due to technological, infrastructural, institutional and behavioral lock-ins. A recent strand of the climate economics literature (Lemoine and Traeger 2014; Lontzek, Cai, Judd, and Lenton 2015; van der Ploeg and de Zeeuw 2018, and Cai and Lontzek 2019) has attempted to incorporate hysteretic effects through the tipping point modeling. Dietz, Rising, Stoerk, and Wagner (2020) synthesize a number of different approaches in this literature into a meta-model and argue for the need to develop sophisticated models of hysteretic behavior.[4]

Finally, careful consideration of uncertainty such as in Temzelides (2016), Li, Narajabad, and Temzelides (2016), Traeger (2017), Cai and Lontzek (2019), Brock and Hansen (2018), Van den Bremer and van der Ploeg (2018), Barnett, Brock, and Hansen (2020), Giglio, Kelly, and Stroebel (2020), Kotlikoff, Kubler, Polbin, and Scheidegger (2020) and Lemoine (2021) has been another important theme of recent research on economics of climate change.

3.3 Climate hysteresis functional

In this section, we analyze a particular hysteresis functional that captures the most important elements of the abstract setting that we develop in Section 4 and allows to transparently show how the solution in the general environment works.

Specifically, let

$$h_t\left(c^t\right) = g_t\left(w^t\right)c_t + \int_0^t k_{s,t}\left(w^t\right)c_s ds, \tag{4}$$

where c_t denotes the amount of emissions at time t.[5] Here, there are two sources of the effects of emissions. The first source is contemporaneous. The amount of emission c_t yields climate externality equal to $g_t\left(w^t\right)$, where g_t is a functional of the path of uncertainty w^t. The second source is the effect of the past emissions c_s. Each of the emissions in the previous period c_s contributes $k_{s,t}\left(w^t\right)$ to the climate externality h_t at time t, where $k_{s,t}$ is a functional of the path of uncertainty w^t. It is important to note that both of these terms are path-dependent.[6]

We observe that with the functional (4) the optimization problem (3) becomes intractable for the stochastic control approach. Indeed, two possibilities for such approach would be either to introduce a control c^t or to introduce an additional control $x_t = h_t(c_t)$. Both approaches lead to

---

[4]Section 6.6 further elaborates on an example of the use of our methodology with tipping points.

[5]Strictly speaking, the functional in this section depends on both the path c^t and w^t while we consider in Section 4 the functional $h_t\left(c^t\right)$. The extension to allow for the dependence on the path w^t is immediate. Further, for the analysis of the stochastic elasticity in Section 5.1, the underlying optimal process $c_t^* = w_t$ and, hence, this is without loss of generality.

[6]Remark 2 in Section 4.1 further discusses motivation for this example as a representation of stochastic integrals.

complicated infinite-dimensional HJB equations. The first is due to the complicated structure of the control space. The second is due to the path dependent first order condition for the process x.[7]

3.4 First-order conditions

We derive the first order conditions of the problem (3) with the climate hysteresis functional (4) by perturbing the process c by a an adapted process z and computing the derivative in ν at $\nu = 0$:

$$
\partial_\nu E \int_0^T \left(-\frac{1}{2}(c_t + \nu z_t - w_t)^2 - \epsilon \left(g_t \left(w^t \right) (c_t + \nu z_t) + \int_0^t k_{s,t} \left(w^t \right) (c_s + \nu z_s) \, ds \right) \right) dt \Bigg|_{\nu=0} =
$$
$$
= E \int_0^T \left(z_t w_t - c_t z_t - \epsilon \left(g_t \left(w^t \right) z_t + \int_0^t k_{s,t} \left(w^t \right) z_s ds \right) \right) dt =
$$
$$
= E \int_0^T z_t \left(w_t - c_t - \epsilon \left(g_t \left(w^t \right) + \int_t^T k_{t,s} \left(w^s \right) ds \right) \right) dt = 0,
$$

where the last line is obtained by changing the order of integration. Since z is an arbitrary adapted process we get the first order conditions

$$
c_t = w_t - \epsilon \left(g_t \left(w^t \right) + E \left[\int_t^T k_{t,s}(w^s) ds \Big| \mathcal{F}_t \right] \right). \tag{5}
$$

The term

$$
\Lambda_t = g_t \left(w^t \right) + E \left[\int_t^T k_{t,s}(w^s) ds \Big| \mathcal{F}_t \right] \tag{6}
$$

plays the central role in Golosov, et al. (2014) and is the marginal externality damage from emissions, it also is equal to the optimal Pigouvian tax that corrects this externality. Specifically, in our setting there are two effects of the marginal change in the emission c_t. First, there is the immediate damage $g_t \left(w^t \right)$. Second, there are marginal damages in each future period s, given by $k_{t,s} \left(w^s \right)$. These damages are then integrated and evaluated as the expectation conditional on the information up to time t.

It is important to note the difference with the case with no hysteresis, where the externality is a function $g_t \left(w_t \right)$ and not a functional of the trajectory w^t. First, the future expected marginal damage $E \left[\int_t^T k_{t,s}(w^s) ds \Big| \mathcal{F}_t \right]$ is absent in that case as the change in emissions today does not have effects on the future. Second, the contemporaneous effects $g_t \left(w^t \right)$ also may be significantly different

---

[7]Of course, for some special cases there is no need to use the theory we develop. In Section 6 we show, whenever possible, how to derive the results using known tools. One interesting class of examples can be solved by extracting a martingale and then using the Clark-Ocone formula. In Section 6.3 and Remark 3 we show how to do this for the case of cumulative hysteresis $h_t \left(c_t \right) = c_t \int_0^t c_s ds$; in Section 6.4 and Remark 4 we show how to do this for when the kernel $k_{s,t}$ is mutliplicative. We further expand on this class of examples in Boulatov, Riabov, and Tsyvinski (2020).

14

from $g_t(w_t)$ as it depends now on the path w^t. Even if emissions have no immediate effects, the past effects and the expected cumulative effects may be very significant. In other words, both the past and the future non-trivially matter.

Our primary interest is in providing the evolution of the marginal externality damage, that is, in the semimartingale decomposition $d\Lambda_t = \alpha_t(w^t) dt + \beta_t(w^t) dw_t$. This decomposition can be thought of as the evolution of the marginal externality damage (or the optimal Pigouvian tax) with respect to time dt, the drift, and with respect to the realizations of uncertainty dw_t, the diffusion coefficient. Since both the terms $g_t(w^t)$ and $E\left[\int_t^T k_{t,s}(w^s)ds\middle|\mathcal{F}_t\right]$ are path-dependent, we cannot use the standard Itô formula as it applies only to the functions and not to the functionals of the path. We proceed by analyzing the present and the future effects separately as the analysis requires the use of different methods.

3.5 Present effects of the emissions path

The present damage $g_t(w^t)$ of the current emissions is already a well-defined functional of the path w^t. We use a recently developed Dupire's functional Itô formula for the non-anticipative functional (Dupire 2009, 2019; Cont and Fournie 2013) to represent this term.

We first introduce two derivatives due to Dupire (2009, 2019): the vertical and the horizontal derivative. These are, respectively, the functional analogues of the space and of the time derivatives.

Definition 1. (Vertical Derivative) Let $D[0, t]$ be the space of càdlàg paths. The vertical derivative of a functional $h : D[0, t] \to \mathbb{R}$ is the limit

$$\partial_{c_t} h(c) = \lim_{\epsilon \to 0} \frac{h(c + \epsilon e_t) - h(c)}{\epsilon},$$

where $e_t(s) = 1_{s=t}$.

Intuitively, this derivative measures an influence on the functional of a (discontinuous) bump to the path at time t (Figure 1). If the functional is just the function of current realization of the path, this yields the usual derivative $h'(c_t)$.

Definition 2. (Horizontal derivative) Let $z : [0, T] \times D[0, T] \to \mathbb{R}$ be an adapted functional on the space of trajectories. The horizontal derivative Δ_t is defined as a limit

$$\Delta_t z_t(c) = \lim_{\epsilon \to 0} \frac{z_{t+\epsilon}(c_{\cdot,\epsilon}) - z_t(c)}{\epsilon},$$
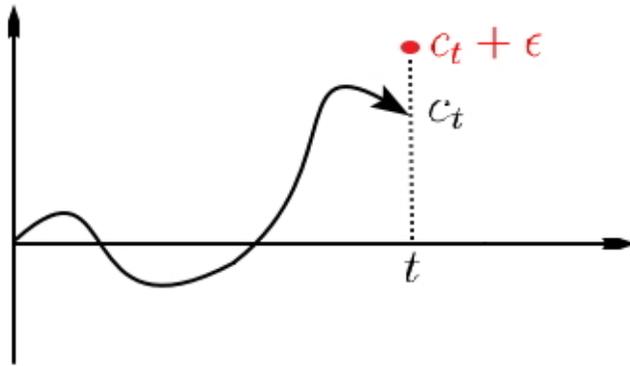
15

Figure 1: Dupire's vertical derivative: path perturbation

where $c_{\cdot,\epsilon}$ is an extension of the path c from $[0, t]$ to $[0, t + \epsilon]$ by $c_{s,\epsilon} = c_t$ for $t \le s \le t + \epsilon$.

Intuitively, this derivative freezes the path of the underlying process at time t, extends it for a small time ϵ and evaluates the functional $z_{t+\epsilon}$ on this extended path (Figure 2). If the functional z_t is just the function of the value of the path at time t, this yields the usual time derivative of $(\partial_t z_t)(c_t)$.
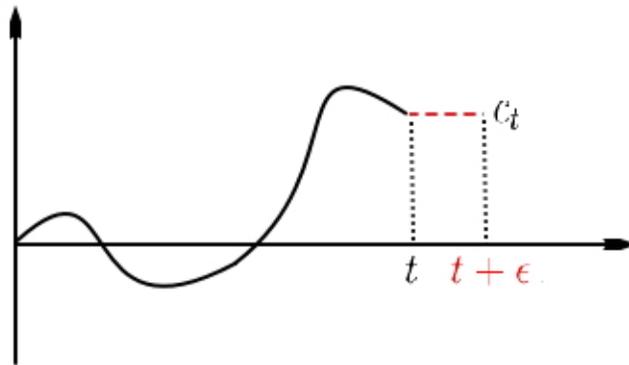

Figure 2: Dupire's horizontal derivative: path perturbation

If the functional $g_t\left(w^t\right)$ does not depend on the path of policies and is a function $f\left(w_t, t\right)$ of only the current realization of uncertainty and of time, we can use the usual Itô formula. For its application, we need to ensure that the first and the second order space derivative $f_1'\left(w_t, t\right)$, $f_1''\left(w_t, t\right)$ and the time derivative $f_2'\left(w_t, t\right)$ exist. When, as is the case in this section, there is hysteresis to apply the functional Itô formula we need to make a similar assumption on the functional $g_t\left(w^t\right)$ as in the case of it being a function but with different notions of derivatives.

16

Lemma 1. *(Functional Itô Formula: Dupire, 2009, 2019; Cont and Fournie, 2013) Let x_t be an Itô process, $dx_t = b_t dt + \sigma_t dw_t$, where the drift coefficient b_t and the diffusion coefficient σ_t are adapted functionals of w. Let g_t be horizontally differentiable and twice vertically differentiable functional on the space of cádlág paths. Then the process $g_t(x^t)$ is an Itô process with the stochastic differential*

$$dg_t\left(x^t\right) = \Delta_t g_t(x^t) dt + \partial_{c_t} g_t(x^t) dx_t + \frac{1}{2} \partial_{c_t}^2 g_t(x^t)(\sigma_t)^2 dt. \tag{7}$$

The lemma gives a semimartingale decomposition of the functional $g_t\left(x^t\right)$. It is similar in form to the usual Itô formula that applies to the functions of the state. However, the time derivative of a function is replaced by the horizontal derivative of the functional Δ_t; the first and second space derivatives of the function are replaced by the first and second vertical derivatives ∂_{c_t} and $\partial_{c_t}^2$. Of course, if there is no path-dependency and g_t is a function rather than a functional, one recovers the usual Itô formula.

Applying functional Itô formula to $g_t\left(w^t\right)$ we obtain

$$dg_t\left(w^t\right) = \left(\Delta_t g_t(w^t) + \frac{1}{2} \partial_{c_t}^2 g_t(w^t)\right) dt + \partial_{c_t} g_t(w^t) dw_t.$$

Summarizing the results in this section: the drift of the dynamics of the present effects of the emissions is determined by the horizontal and the second vertical derivative and the diffusion coefficient is given by the vertical derivative of the present marginal effects of the policy functional. This result both parallels and differs from the case of no hysteresis where, similarly, the first and second order derivatives matter for the dynamics. However, the notion of the derivative is very different as the horizontal and the vertical derivatives measure how hysteresis functionals change with the whole past trajectory. Even if there are no direct contemporaneous effects of the policy, its dependence on the past may lead to significant effects on the dynamics of optimal policy. At the same time, using these different notions of the derivatives the functional Itô formula allows us to convey the similar intuition as in the case of no hysteresis. In particular, uncertainty represents itself in the second vertical derivative of the functional $g_t\left(w^t\right)$ by affecting the drift and in the first vertical derivative of the functional affecting the diffusion coefficient.

3.6 Expected future effects of the emissions paths

The most challenging part of the paper is to characterize the dynamics of the expected future marginal effects of current emission represented by the conditional expectation process $E\left[\int_t^T k_{t,s}(w^s)ds\Big|\mathcal{F}_t\right]$. Let $\xi_t = \int_t^T k_{t,s}(w^s)ds$. Note that the conditional expectation process $\mathbb{E}\left[\xi_t|\mathcal{F}_t\right]$ is in general not a martingale. For example, if $k_{s,t} = 1$, then $\xi_t = T - t$ and we have $\mathbb{E}\left[\xi_t|\mathcal{F}_t\right] = T - t$. Also note that ξ_t is not adapted in general. The difficulty more broadly comes from the fact that in the conditional expectation both the process ξ_t and the filtration $\mathcal{F}_t$ are changing with time.

3.6.1 Total derivative formula for conditional expectations

In this section, we derive a new total derivative formula for conditional expectation processes – this is the main new theoretical tool that we develop in this paper. A more general statement of the result and a complete proof is given in the Appendix. Here, we state a simpler version of the result and an outline of a heuristic proof.

Definition 3. The Malliavin derivative of the cylindrical functional $f(w) = g(w_{t_1}, \ldots, w_{t_k})$, where $t_1 < \ldots < t_k$ and g is a smooth function, is defined by $D_t f(w) = \sum_{i=1}^k \partial_i g(w_{t_1}, \ldots, w_{t_k}) 1_{t \leq t_i}$. This defines a closable linear operator $D : L^2(\Omega) \to L^2(\Omega \times [0, T])$. Its closure is also denoted by D and is called the Malliavin derivative operator.

Intuitively, a Malliavin derivative is a change in a functional due to a perturbation of the whole path of the process (Figure 3).
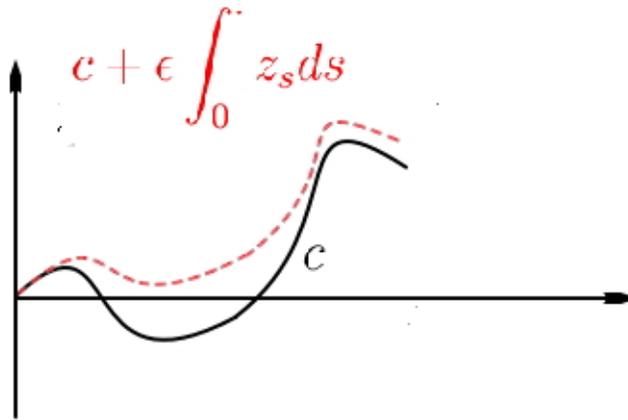

Figure 3: Malliavin derivative: path perturbation

The next proposition states the total derivative formula, the main technical tool that we develop in this paper.

Proposition 1. *(Total derivative formula) Let an adapted square integrable process $(X_t)_{0 \le t \le T}$ be represented as*

$$X_t = \mathbb{E}[\xi_t | \mathcal{F}_t].$$

Let $(\xi_t)_{0 \le t \le T}$ be a Malliavin differentiable square integrable absolutely continuous process, that is, a process of the form $\xi_t = \xi_0 + \int_0^t \eta_s ds$, where ξ and η may be anticipative. Then,

$$dX_t = \mathbb{E}[\partial_t \xi_t | \mathcal{F}_t] dt + \mathbb{E}[D_t \xi_t | \mathcal{F}_t] dw_t, \tag{8}$$

where D is a Malliavin derivative.

Proof. The Clark-Ocone formula gives

$$\xi_t = \mathbb{E}[\xi_t] + \int_0^T \mathbb{E}[D_r \xi_t | \mathcal{F}_r] dw_r.$$

Taking the conditional expectation leads to

$$X_t = \mathbb{E}[\xi_t] + \int_0^t \mathbb{E}[D_r \xi_t | \mathcal{F}_r] dw_r.$$

The total derivative is then

$$dX_t = \mathbb{E}[\partial_t \xi_t] dt + \mathbb{E}[D_t \xi_t | \mathcal{F}_t] dw_t + \left(\int_0^t \mathbb{E}[D_r \partial_t \xi_t | \mathcal{F}_r] dw_r \right) dt =$$
$$= \mathbb{E}[D_t \xi_t | \mathcal{F}_t] dw_t + \left(\mathbb{E}[\partial_t \xi_t] + \int_0^t \mathbb{E}[D_r \partial_t \xi_t | \mathcal{F}_r] dw_r \right) dt.$$

Note now that by the Clark-Ocone formula applied to the derivative $\partial_t \xi_t$:

$$\partial_t \xi_t = \mathbb{E}[\partial_t \xi_t] + \int_0^T \mathbb{E}[D_r \partial_t \xi_t | \mathcal{F}_r] dw_r,$$

and taking the conditional expectations:

$$\mathbb{E}[\partial_t \xi_t | \mathcal{F}_t] = \mathbb{E}[\partial_t \xi_t] + \int_0^t \mathbb{E}[D_r \partial_t \xi_t | \mathcal{F}_r] dw_r.$$

The total derivative is then given by (8). □

The lemma provides a time-dependent extension of the Clark-Ocone formula for processes that can be represented as a conditional expectation of an absolutely continuous process.[8] The first term, $\mathbb{E}[\partial_t \xi_t | \mathcal{F}_t]$, can be thought of as a time derivative of this variable and represents how the conditional expectation evolves with respect to time. The second term, $\mathbb{E}[D_t \xi_t | \mathcal{F}_t]$, can be thought of as a stochastic derivative with respect to the underlying process w_t and represents how the conditional expectation changes with the changes in w_t.[9]

3.6.2 Semimartingale decomposition of the expected future effects of policy paths

We now make mild assumptions on the kernel $k_{t,s}$ – absolute continuity of the kernel in variable t and Malliavin differentiability as a functional from w^s – and apply the total derivative formula to the expected future marginal damages, $E\left[\int_t^T k_{t,s}(w^s)ds \Big| \mathcal{F}_t\right]$:

$$dE\left[\int_t^T k_{t,s}(w^s)ds \Big| \mathcal{F}_t\right] = \underbrace{\mathbb{E}\left[\partial_t \int_t^T k_{t,s}(w^s)ds \Big| \mathcal{F}_t\right]dt}_{\text{time derivative}} + \underbrace{\mathbb{E}\left[D_t \int_t^T k_{t,s}(w^s)ds \Big| \mathcal{F}_t\right]dw_t}_{\text{stochastic derivative}} =$$

$$= \underbrace{\mathbb{E}\left[\int_t^T \partial_t k_{t,s}(w^s)ds \Big| \mathcal{F}_t\right] - k_{t,t}\left(w^t\right)dt}_{\text{time derivative}} + \underbrace{\mathbb{E}\left[\int_t^T D_t k_{t,s}(w^s)ds \Big| \mathcal{F}_t\right]dw_t}_{\text{stochastic derivative}}.$$

The meaning of this equation is straightforward. The conditional expectation $dE\left[\int_t^T k_{t,s}(w^s)ds \Big| \mathcal{F}_t\right]$ measures the effects of the change in the emissions in period t on the expected future marginal externality $k_{t,s}(w^s)$ in all periods s ($s \in [t, T]$). The time evolution of the conditional expectation is given by the analogue of the time derivative. The stochastic evolution of the conditional expectation is given by a stochastic derivative of the cumulative change in all marginal effects of emission in time t on all future periods s with respect to variation in the underlying process w.

---

[8]See Section 7.3 for the more detailed discussion of the relationship of our total derivative formula to the Clark-Ocone formula.

[9]The process (ξ_t) in Proposition 1 need not be adapted, and we can understand the process $X_t = \mathbb{E}[\xi_t | \mathcal{F}_t]$ as the adapted projection of the process ξ. In Proposition 1 we thus prove that an adapted projection of an absolutely continuous process is necessarily an Itô process. In the Appendix 9.6 we show the converse to this clam – any Itô process can be represented as such projection.

3.7 Dynamics of the marginal externality damage

We can now collect the results of the previous two sections. We obtain the dynamics of the marginal externality damage

$$d\Lambda_t = -\epsilon \left(\Delta_t g_t(w^t) + \frac{1}{2}\partial_{c_t}^2 g_t(w^t) + \mathbb{E}\left[\int_t^T \partial_t k_{t,s}(w^s)ds \Big| \mathcal{F}_t \right] - k_{t,t}\left(w^t\right) \right) dt+$$
$$+ \left(\partial_{c_t} g_t(w^t) + \mathbb{E}\left[\int_t^T D_t k_{t,s}(w^s)ds \Big| \mathcal{F}_t \right] \right) dw_t.$$

This, we believe, is an important result as it solves in the closed form the dynamics of the evolution of the marginal externality damage or the optimal Pigouvian tax. This formula shows how the policymaker should change the tax with the passage of time and with the realizations of uncertainty – that is, we determined the drift and the diffusion coefficients of the process of the optimal Pigouvian tax.

A related way to think about this result is that it determines what features of the model matter up to the first order. For the time updating, the functional Itô formula shows that the horizontal and the second order vertical derivative are important and the total derivative formula shows that the "time derivative" of the expected cumulative damages are important. For the updating with the movement of uncertainty, the functional Itô formula shows that the vertical derivative and the total derivative formula shows that the "stochastic derivative" of the expected cumulative marginal damages are important.

We also immediately obtain the dynamics of the optimal emission policies from equation (5) in closed form:

$$dc_t = -\epsilon \left(\Delta_t g_t(w^t) + \frac{1}{2}\partial_{c_t}^2 g_t(w^t) + \mathbb{E}\left[\int_t^T \partial_t k_{t,s}(w^s)ds \Big| \mathcal{F}_t \right] - k_{t,t}\left(w^t\right) \right) dt+$$
$$+ \left(1 - \epsilon\partial_{c_t} g_t(w^t) - \epsilon\mathbb{E}\left[\int_t^T D_t k_{t,s}(w^s)ds \Big| \mathcal{F}_t \right] \right) dw_t.$$

We thus extended the result of the macro-climate model of Golosov, et al. (2014) to the case with the path-dependent hysteresis given by (4). Note that all of the results here are given in closed form. The next sections show that the main insights derived with this specifications apply to general hysteresis functionals.

4 General hysteresis functionals

We now return to the analysis of the problem (3) for a general hysteresis functional.

4.1 The class $\mathcal{A}_t$ of path-dependent functionals

In this section we describe a class of path-dependent functionals $\mathcal{A}_t$ which, in a sense, is parallel to continuously differentiable functions of a real-valued argument. It will turn out that this class, while very general, allows significant tractability in the analysis of path-dependent problems. The introduction of this class of functionals is one contribution of our paper.

Remark 1. Adding a path-dependent functional h_t can considerably change the structure of the optimal policy. For example, consider the functional $h_t(c) = c_{\frac{t}{2}}$ that describes the effects of the policies at the middle of the time period $[0, t]$. We show in the appendix that the optimal policy is no longer continuous and has a jump at $t = \frac{T}{2}$.

Let us recall that a Fréchet derivative of a functional $h_t : C[0, t] \to \mathbb{R}$ at a point $c \in C[0, t]$ is a signed measure μ on $[0, t]$ such that

$$h_t(c + z) = h_t(c) + \int_0^t z_s d\mu(t, s) + o(\|z\|), \|z\| \to 0.$$

The primary difficulty with this formulation is that it features a very general dependence of the measure μ on time t. The next assumption structures this dependence.

Assumption 1. *(Class $\mathcal{A}_t$) Suppose the (Fréchet) derivative of the functional h_t has an absolutely continuous part and an atom at point t:*

$$h_t(c + z) = h_t(c) + \partial_{c_t} h_t(c) z_t + \int_0^t \delta_s h_t(c) z_s ds + o(\|z\|), \|z\| \to 0.$$

The family of such functionals is denoted by $\mathcal{A}_t$.

This assumption on the derivative of functionals means that there are two sources of variation for h_t: an instantaneous influence of a perturbation at the moment t given by an atom and an integral influence of perturbations at previous moments $s \leq t$ which is an absolutely continuous process. Assumption 1 is mild – it only requires that the derivative of a functional (which is a measure) is absolutely continuous and has an atom at the present time.

The class $\mathcal{A}_t$ contains a variety of functionals:

$$h_t(c) = f(c_t),$$

which is state-dependent but not path-dependent;

$$h_t(c) = f_t(c_t, \int_0^t g_t(c_s)ds),$$

which jointly depends on state c_t and the integral influence of the path $\int_0^t g_t(c_s)ds$ and, moreover, both the joint dependence f_t and the effects of the past policies g_t depend on time t;

$$h_t(c) = f_t(\int_0^t \int_0^t g_t(c_s, c_r)dsdr, ...)$$

where now there is a joint dependence on the past values c_s and c_r via a repeated integral.

The proposition that follows shows that this assumption covers a very general class of functionals.

Proposition 2. *Every Fréchet differentiable functional $g : C[0,T] \to \mathbb{R}$ is a pointwise limit of functionals from the class $\mathcal{A}_T$.*

Proof. In the appendix. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

We next explore the nature of the present marginal influence of policy $\partial_{c_t} h_t$ which is itself a path-dependent functional. The lemma that follows (proven for a more general case in the appendix) connects Dupire's vertical derivative to the derivatives of the functionals in the class $\mathcal{A}_t$.

Lemma 2. *Let the functional h_t be in the class $\mathcal{A}_t$, then $\partial_{c_t} h_t(c)$ is the vertical derivative in the Dupire sense.*

Remark 2. We now discuss some additional motivation behind Assumption 1. Consider, for example, hysteresis given by an Itô process:

$$h_t(c) = \int_0^t b_s dc_s, \qquad\qquad\qquad\qquad\qquad\qquad (9)$$

where b is an absolutely continuous function. The functional $h_t(c)$ in equation (9) is well-defined if and only if the process $(c_s)_{0 \leq s \leq t}$ is of bounded variation. This is not true for an arbitrary progressively measurable process $(c_s)_{0 \leq s \leq t}$, and we cannot expect that the optimal policy process

will be of bounded variation. For example, in our baseline case of Section 2.1 $c_t^* = w_t$ and then almost all realizations of c^* are of unbounded variation. However, if the coefficient b is absolutely continuous, then we can integrate (9) by parts and rewrite $h_t(c)$ in the form:

$$h_t(c) = -\int_0^t b_s' c_s ds + b_t c_t.$$

Further, h_t is just a linear functional from c :

$$h_t(c+z) - h_t(c) = b_t z_t - \int_0^t b_s' z_s ds.$$

So, $h_t \in \mathcal{A}_t$ with $\partial_{c_t} h(c) = b_t$, $\delta_s h_t(c) = -b_s'$.

4.2 Optimal policy: the first order conditions

In this section, we derive the first order conditions for the optimal process c^ϵ. While in some circumstances one can write a recursive formulation even for a path-dependent problem and then find a Hamilton-Jacobi-Bellman equation, in general it is difficult or impossible to do it. Here, we instead find the first order conditions for optimal policy using a variational method.

Proposition 3. *(The first order conditions for the optimum)* *Let the optimal policy process* c^ϵ *solve (3), then*

$$c_t^\epsilon = w_t - \epsilon \left(\partial_{c_t} h_t((c^\epsilon)^t) + E\left[\int_t^T \delta_t h_s((c^\epsilon)^s) ds \Big| \mathcal{F}_t \right] \right). \tag{10}$$

Proof. Perturb the process c by νz, where z is an adapted process, and compute the derivative in ν at $\nu = 0$:

$$\partial_\nu E \int_0^T \left(-\frac{1}{2}(c_t + \nu z_t - w_t)^2 - \epsilon h_t(c^t + \nu z^t) \right) dt \Big|_{\nu=0} =$$

$$= E \int_0^T \left(z_t w_t - c_t z_t - \epsilon \left(\partial_{c_t} h_t(c^t) z_t + \int_0^t \delta_s h_t(c^t) z_s ds \right) \right) dt =$$

$$= E \int_0^T z_t \left(w_t - c_t - \epsilon \left(\partial_{c_t} h_t(c^t) + \int_t^T \delta_t h_s(c^s) ds \right) \right) dt = 0.$$

Since z is an arbitrary adapted process, we get the first order conditions (10). The importance of Assumption 1 is evident in particular in the third line of the proof. If we did not impose this assumption, then instead of the integral $\int_t^T \delta_t h_s(c^s) ds$ we would have a general measure $\mu_t(c^s, [t, T])$ which can even be discontinuous in t. Assumption 1 imposes a smooth structure for this measure

by requiring that the measure has an absolutely continuous part. □

This equation (10) has a natural economic meaning. When hysteresis ϵh_t is present, the solution to the optimal problem $(c_t^* = w_t)$ is modified by the two terms. The first term,

$$I_t = \epsilon \partial_{c_t} h_t(c^\epsilon),$$

is the instantaneous marginal effect of the change in the policy c_t^ϵ on functional h_t. The second term gives the future marginal effects of policy c_t^ϵ

$$F_t = \epsilon E\left[\int_t^T \delta_t h_s((c^\epsilon)^s) ds \middle| \mathcal{F}_t\right].$$

on the values of all future hysteresis functionals $h_s\left((c^\epsilon)^s\right)$. For each time s (where $t \leq s \leq T$), the marginal effect of changing policy in period t is represented by the derivative $\delta_t h_s((c^\epsilon)^s)$ of the functional h_s with respect to change in policy at time t. These marginal effects are evaluated as a conditional expectation at time t and hence represent the expected future marginal effects. It is useful to think about the term F_t as the cumulative impulse response of the change of the policy today on all future hysteresis functionals. This is similar to Alvarez and Lippi (2019), Alvarez, Le Bihan, and Lippi (2016), Alvarez, Lippi, and Oskolkov (2020) and Borovička, Hansen, Scheinkman (2014), Borovička, Hansen, Hendricks, and Scheinkman (2011), and Borovička and Hansen (2016). Note that F_t is a conditional expectation process that may non-trivially change with time as both the marginal effects $\delta_t h_s$ and the filtration $\mathcal{F}_t$ changes.

Both terms depend on the path of policy $c_{[0,s]}^\epsilon = (c^\epsilon)^s$. As seen here, the assumption that the functional h_t is in the class $\mathcal{A}_t$ allows us to conveniently separate the first order condition in two parts: the marginal effects of the policy on current period I_t and the conditional expectation of the cumulative future marginal effects of policy paths in the future, F_t.

5 Characterizing the general problem

This section is divided into two main parts. Section 5.1 derives a closed-form characterization of the change in the optimal policy when hysteresis is small. Section 5.2 provides a characterization of the optimal policy.[10]

---

[10]It may be useful for a reader to also consider an example in the appendix that provides a parallel characterization for the case with no path-dependency.

5.1 Stochastic hysteresis elasticity and its dynamics

In this section we derive a closed-form characterization of the change in the optimal policy when hysteresis is small. We call this first order process a stochastic hysteresis elasticity. We characterize the dynamics of the stochastic elasticity in closed form by providing its semimartingale decomposition which is the main contribution of this section.

5.1.1 Stochastic hysteresis elasticity

We first formally define the stochastic hysteresis elasticity. This is the first order process that represents the change in optimal policy process c^* in response to introduction of an infinitesimal path-dependent hysteresis functional h_t .

Definition 4. (Stochastic hysteresis elasticity) Let c_t^* be a solution to the baseline problem (1) and c_t^ϵ be a solution to the problem (3), where hysteresis is given by the functional ϵh_t, $\epsilon \to 0$. The first-order process or the stochastic hysteresis elasticity C_t^h is such that

$$c_t^\epsilon = c_t^* + \epsilon C_t^h + o(\epsilon).$$

The stochastic elasticity C_t^h is the change to the first order in the optimal policy process in response to a small change in the hysteresis h. In this sense, it is similar to the usual concept of elasticity but now determines how the whole process c_t changes. From now on, to ease notation we drop the dependence of C_t^h on h and denote it simply by C_t.

Differentiating (10) with respect to ϵ at $\epsilon = 0$, and recalling that $c_t^0 = c_t^* = w_t$ we find the process C_t explicitly

$$C_t = -\partial_{c_t} h_t(w^t) - E\left[\int_t^T \delta_t h_s(w^s) ds \middle| \mathcal{F}_t\right]. \tag{11}$$

This equation is already interesting by itself as it presents the change in the optimal policy plan due to the introduction of the path-dependent effect. Moreover, it is given in closed form.

We now turn to presenting the main result of this section – showing that C_t is an Itô process and, most importantly, writing its explicit semimartingale decomposition. We are seeking a representation

$$dC_t = \beta\left(w^t\right) dt + \gamma\left(w^t\right) dw_t,$$

where β and γ are potentially path-dependent. In other words, we want to determine the dynamics of the stochastic elasticity C_t.

5.1.2 Present effect of the policy path

This section provides a semimartingale decomposition of the term $\partial_{c_t} h_t(c^*)$ – the present effect of the path of policies $(c^*)^t = w^t$. The principal tool that we use is Dupire's functional Itô formula in Lemma 1 that applies to the non-anticipative functionals of the path. We now make an assumption that allows the use of this formula.

Assumption 2. *There exists a non-anticipative functional $q : [0,T] \times D[0,T] \to \mathbb{R}$ which is horizontally differentiable, twice vertically differentiable and is such that $q_t(c) = \partial_{c_t} h_t(c)$. In other words, the derivative $\partial_{c_t} h_t(c)$ can be extended to a $C_b^{1,2}$-functional on the space of cádlág paths. We denote this extension also by $\partial_{c_t} h_t(c)$.*

Applying the functional Itô formula, and noting that $c_t^* = w_t$ we obtain a semimartingale decomposition of the functional $\partial_{c_t} h_t(c^*)$ that depends on the whole history of policy $(c^*)^t = w^t$:

$$d(\partial_{c_t} h_t(c^*)) = \Delta_t \left(\partial_{c_t} h_t(w^t) \right) dt + \partial_{c_t} \left(\partial_{c_t} h_t(w^t) \right) dw_t + \frac{1}{2} \partial_{c_t} \left(\partial_{c_t}^2 h_t(w^t) \right) dt,$$

and gathering the terms we obtain the following lemma.

Lemma 3. *(Dynamics of the present effect, I_t) The semimartingale decomposition of the present marginal effects, I_t, of the policy path is given by*

$$d(\partial_{c_t} h_t(c^*)) = \left(\Delta_t(\partial_{c_t} h_t(w^t)) + \frac{1}{2} \partial_{c_t}^3 h_t(w^t) \right) dt + \partial_{c_t}^2 h_t(w^t) dw_t. \tag{12}$$

This part of the derivation is already interesting as a stand-alone result. The reason why one can apply the Dupire and Cont-Fournie analysis is that $\partial_{c_t} h_t(c^*)$ is already represented as a functional of the path. The decomposition (12) then has the same intuitive meaning as the standard Itô's formula but now applies to the functional of the past, not the function of the present realization. Note that we are already applying the functional Itô's formula to the marginal effects, that is, to the vertical derivative of the functional $\partial_{c_t} h_t(w^t)$. Hence, there are the second, $\partial_{c_t}^2$, and the third, $\partial_{c_t}^3$, vertical derivatives as well as the mixed derivative $\Delta_t \partial_{c_t}$.

5.1.3 Expected future effects of policy paths

We now make an assumption on the derivative of the functional, $\delta_t h_s(c^*)$, that is needed to apply the total derivative formula of Proposition 1.[11]

---

[11]In Section 7.2, we further discuss the smoothness assumptions for the functionals h_t.

Assumption 3. *The derivative $\delta_t h_s(c^*)$ is an absolutely continuous in t and is a Malliavin differentiable functional of the path w^s.*

It follows that the future marginal effect of a given policy path is given by

$$\frac{d}{dt}\int_t^T \delta_t h_s(c^*)ds = -\delta_t h_t(c^*) + \int_t^T \frac{\partial}{\partial t}(\delta_t h_s)(c^*)ds.$$

and its conditional expectation is given by the total derivative formula (8):

$$d\mathbb{E}\left[\int_t^T \delta_t h_s(c^*)ds\Big|\mathcal{F}_t\right] = \underbrace{\mathbb{E}\left[\partial_t \int_t^T \delta_t h_s(c^*)ds\Big|\mathcal{F}_t\right]dt}_{\text{time derivative}} + \underbrace{\mathbb{E}\left[D_t \int_t^T \delta_t h_s(c^*)ds\Big|\mathcal{F}_t\right]dw_t}_{\text{stochastic derivative}}.$$

The meaning of this equation is straightforward. The conditional expectation

$$F_t = \mathbb{E}\left[\int_t^T \delta_t h_s(c^*)ds\Big|\mathcal{F}_t\right]$$

describes how a change in policy in period t determines the expected future marginal effects of that policy $\delta_t h_s(c^*)$ in all periods s ($s \in [t, T]$). The time evolution of the conditional expectation is given by the analogue of the time derivative. The stochastic evolution of the conditional expectation is given by a stochastic derivative of the cumulative change in all marginal effects of policy t on all future periods s with respect to variation in the underlying process w. Collecting the terms gives us the dynamics of the future effects of policy and shows that it is an Itô process.

Lemma 4. (Dynamics of the future effects, F_t) *The semimartingale representation of the future marginal effects, F_t, of policy paths is given by:*

$$d\mathbb{E}\left[\int_t^T \delta_t h_s(c^*)ds\Big|\mathcal{F}_t\right] =$$
$$= \left(-\delta_t h_t(w^t) + \int_t^T \mathbb{E}\left[\frac{\partial}{\partial t}(\delta_t h_s)(w^s)\Big|\mathcal{F}_t\right]ds\right)dt + \mathbb{E}\left[D_t \int_t^T \delta_t h_s(w^s)ds\Big|\mathcal{F}_t\right]dw_t. \qquad (13)$$

5.1.4 Dynamics of the stochastic elasticity

We now combine the results of the decompositions of the present effects (12) and the future effects (13) of policy and recalling that $c_t^* = w_t$:

$$dC_t = - \underbrace{\left(\Delta_t \partial_{c_t} h_t(w^t) dt + \frac{1}{2} \partial_{c_t}^3 h_t(w^t) dt + \partial_{c_t}^2 h_t(w^t) dw_t \right)}_{\text{Present effects: functional Itô}} -$$

$$- \underbrace{\left(-\delta_t h_t(w^t) dt + E\left[\int_t^T \frac{\partial}{\partial t}(\delta_t h_s)(w^s) ds \middle| \mathcal{F}_t \right] dt + E\left[\int_t^T D_t \left(\delta_t h_s(w^s) \right) ds \middle| \mathcal{F}_t \right] dw_t \right)}_{\text{Future effects: total derivative}},$$

where in the last term we interchanged the Malliavin derivative and the integral.

Collecting the terms gives the semimartingale decomposition of the first order process C_t in closed form.

Theorem 1. *(**Dynamics of stochastic elasticity**) The semimartingale decomposition of the stochastic elasticity C_t is given by*

$$dC_t = - \left(\Delta_t \partial_{c_t} h_t(w^t) + \frac{1}{2} \partial_{c_t}^3 h_t(w^t) - \delta_t h_t(w^t) + E\left[\int_t^T \frac{\partial}{\partial t}(\delta_t h_s)(w^s) ds \middle| \mathcal{F}_t \right] \right) dt -$$

$$- \left(\partial_{c_t}^2 h_t(w^t) + E\left[\int_t^T D_t \left(\delta_t h_s(w^s) \right) ds \middle| \mathcal{F}_t \right] \right) dw_t. \tag{14}$$

This theorem shows that the optimal process c_t^* changes with the introduction of hysteresis, that is, the additional infinitesimal path-dependent effect of policies. The equation (14) gives the first order process in closed form as

$$dC_t^h = \beta_t \left(w^t \right) dt + \gamma_t \left(w^t \right) dw_t,$$

where β and γ are path-dependent coefficients.

5.2 Dynamics of optimal policy

We now characterize the dynamics of the optimal policy.

In this section, we assume that $\epsilon = 1$, without loss of generality. The first order conditions (10) become

$$w_t = c_t + \partial_{c_t} h_t(c^t) + E\left[\int_t^T \delta_t h_s(c^s) ds \middle| \mathcal{F}_t \right].$$

Let us take the differential of this equation assuming that c is an Itô process, i.e.

$$dc_t = \alpha_t dt + \beta_t dw_t.$$

The present marginal effects of the policy are differentiated using the functional Itô formula (7):

$$d\left(\partial_{c_t} h_t(c^t)\right) = \Delta_t \partial_{c_t} h_t(c^t) dt + \partial_{c_t}^2 h_t(c^t) dc_t + \frac{1}{2} \partial_{c_t}^3 h_t(c^t) \beta_t^2 dt.$$

The future expected marginal effects of the policy are differentiated using the total derivative formula (8):

$$d\left(E\left[\int_t^T \delta_t h_s(c^s) ds \Big| \mathcal{F}_t\right]\right) = E\left[\partial_t \int_t^T \delta_t h_s(c^s) ds \Big| \mathcal{F}_t\right] dt + E\left[\int_t^T D_t\left(\delta_t h_s(c^s)\right) ds \Big| \mathcal{F}_t\right] dw_t =$$

$$= E\left[\int_t^T \partial_t\left(\delta_t h_s(c^s)\right) ds \Big| \mathcal{F}_t\right] dt - \delta_t h_t(c^t) dt + E\left[\int_t^T D_t\left(\delta_t h_s(c^s)\right) ds \Big| \mathcal{F}_t\right] dw_t.$$

The differential of the first order conditions is then given by

$$dw_t = dc_t + \Delta_t \partial_{c_t} h_t(c^t) dt + \partial_{c_t}^2 h_t(c^t) dc_t + \frac{1}{2} \partial_{c_t}^3 h_t(c^t) \beta_t^2 dt +$$

$$+ E\left[\int_t^T \partial_t \delta_t h_s(c^s) ds \Big| \mathcal{F}_t\right] dt - \delta_t h_t(c^t) dt + E\left[\int_t^T D_t\left(\delta_t h_s(c^s)\right) ds \Big| \mathcal{F}_t\right] dw_t.$$

Collecting the terms near dt and dw_t we derive dynamics of the optimal policy:

$$\begin{cases} 0 = (1 + \partial_{c_t}^2 h_t(c^t))\alpha_t + \frac{1}{2}\partial_{c_t}^3 h_t(c^t)\beta_t^2 + \Delta_t \partial_{c_t} h_t(c^t) + E\left[\int_t^T \partial_t \delta_t h_s(c^s) ds \Big| \mathcal{F}_t\right] - \delta_t h_t(c^t), \\ 1 = (1 + \partial_{c_t}^2 h_t(c^t))\beta_t + E\left[\int_t^T D_t\left(\delta_t h_s(c^s)\right) ds \Big| \mathcal{F}_t\right], \end{cases}$$

and the next theorem characterizes the drift and diffusion coefficients of the optimal policy.

Theorem 2. *(**Optimal policy**) The dynamics of optimal policy is given by:*

$$\alpha_t = -\frac{1}{1 + \partial_{c_t}^2 h_t(c^t)}\left(\frac{1}{2}\partial_{c_t}^3 h_t(c^t)\beta_t^2 + \Delta_t \partial_{c_t} h_t(c^t) + E\left[\int_t^T \partial_t \delta_t h_s(c^s) ds \Big| \mathcal{F}_t\right] - \delta_t h_t(c^t)\right),$$

$$\beta_t = \frac{1 - E\left[\int_t^T D_t \delta_t h_s(c^s) ds \Big| \mathcal{F}_t\right]}{(1 + \partial_{c_t}^2 h_t(c^t))}.$$

We can compare this theorem to the results without hysteresis, where there are only contempo-

raneous effects of the policy $f(c_t)$, the details of which are in the appendix. The diffusion coefficient β_t is determined by two terms. The first term, in the denominator, contains the vertical derivative of the present marginal effects of the policy $\partial_{c_t}\left(\partial_{c_t}h_t(c^t)\right)$ and measures the change in the present marginal costs due to a "bump" in the trajectory of the costs. This is an analogue of the term $f''(c_t)$ for the case without hysteresis. The second term, in the numerator, is the stochastic derivative $E\left[\int_t^T D_t\delta_t h_s(c^s)ds\Big|\mathcal{F}_t\right]$ which measures impact of the change in the trajectory of the stochastic process on the expected future marginal effects in all periods. This term is not present in the case without hysteresis.

The drift coefficient α_t is determined by three terms. The first term, in the denominator, is the same relative scaling as in the case of β_t. The second set of terms is the time evolution of the present, $\Delta_t\partial_{c_t}h_t(c^t)$, and future, $E\left[\partial_t\int_t^T \delta_t h_s(c^s)ds\Big|\mathcal{F}_t\right]$, marginal costs. The analogue of the horizontal derivative $\Delta_t\partial_{c_t}h_t(c^t)$ for the case without hysteresis would be the time derivative $\partial_t f(c_t, t)$. The time evolution of the future marginal costs are is in the case without hysteresis. The third term is the quadratic variance of the present marginal costs, $\frac{1}{2}\partial_{c_t}^3 h_t(c^t)\beta_t^2$. It is a path-dependent analogue of $\frac{1}{2}f'''(c_t)(\beta_t)^2$ in the case without hysteresis.

The equations Theorem 2 characterize the drift and diffusion coefficients implicitly in contrast to the explicit form we obtained in Theorem 1. The explicit form we obtained for the stochastic elasticity thus can be thought of as a characterization of the optimal process for small hysteresis, or small hysteresis asymptotics.

Remark. We now show that assuming that the marginal effect of policy $\delta_t h_s$ are in the class $\mathcal{A}_s$, we can compute Malliavin derivatives $D_t\delta_t h_s(c^s)ds$ in Theorem 2 in a more detailed way.

Corollary. *Let $\delta_t h_s$ be in the class $\mathcal{A}_s$, then*

$$D_t\delta_t g_s(c^s) = \partial_{c_t}\delta_t h_s(c^s)D_t c_s + \int_t^s \delta_r\delta_t h_s(c^s)D_t c_r dr,$$

where the tangent process $D_t c_r$ is given by

$$\begin{cases} d(D_t c_s) = (D_t\alpha_s)ds + (D_t\beta_s)dw_s, \;\; s > t \\ D_t c_t = \beta_t. \end{cases} \tag{15}$$

Proof. Let us perturb the underlying Brownian motion: $S(w, \epsilon z)_t = w_t + \epsilon\int_0^t z_r dr$, $0 \leq t \leq T$. Then

by the definition of the Malliavin derivative

$$c_t(S(w, \epsilon z)) = c_t(w) + \epsilon \int_0^t D_r c_t(w) z_r dr + o(\epsilon),$$

where $D_r c_t$ is a tangent process. Hence,

$$\delta_t h_s(c^s(S(w, \epsilon z))) = \delta_t h_s(c^s(w)) + \epsilon \partial_{c_s} \delta_t h_s(c^s) \int_0^s D_r c_s(w) z_r dr +$$

$$+ \epsilon \int_0^s \delta_r \delta_t h_s(c^s) \int_0^r D_u c_r(w) z_u du dr + o(\epsilon) =$$

$$= \delta_t h_s(c^s(w)) + \epsilon \int_0^s z_r \left(\partial_{c_s} \delta_t h_s(c^s) D_r c_s(w) + \int_r^s \delta_u \delta_t h_s(c^s) D_r c_u(w) du \right) dr + o(\epsilon).$$

The result follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

This characterization is useful as it uses the tangent process which has a particularly simple form. When c is an Itô process, its tangent process is also an Itô process with the drift and diffusion coefficients being the Malliavin derivatives of the drift and diffusion coefficients of the process c.

6   Examples

In this section, we present a number of examples of the hysteresis functionals h_t to illustrate how to use our theoretical results. We also provide, whenever possible, an alternative derivation using other tools. Section 6.1 is the user's guide that describes the steps needed to apply the theory we developed. The first example, Section 6.2, revisits the case of no hysteresis. In this case, only the present effects of the policy are present and the functional Itô formula reduces to the usual Itô formula. There is no need to use the total derivative formula. The second example, Section 6.3, considers cumulative hysteresis. The effects of the polices are a function of the current policy and the integral of the past policies. For the case of the multiplicative dependence of the effects of the current policies and the cumulative hysteresis, the stochastic elasticity takes a very simple form. The present marginal effects of the policy path only have the horizontal derivative which measures how the cumulant of the past policies changes with time and all of the vertical derivatives are equal to zero. The total derivative formula for the conditional expectation of the future marginal effects only has the stochastic derivative component which itself has a very simple form that measures how lengthy the effects of the stochastic shock are. We expand on this class of examples in Boulatov, Riabov, and Tsyvinski (2020) where we analyze such cases extracting a martingale and not using

the main tools of this paper – the functional Itô formula and the total derivative formula. Our third example, Section 6.4, studies hysteresis that depends both on time and the past policies. This example captures three important parts of path-dependent policies. First, there is an effect of the present policy. Second, there is an effect of the hysteresis. The hysteresis is represented by an integral with a kernel that has joint dependence on current time and past policies. That is, path dependency sophisticatedly changes with both the time and evolution of the past polices. Third, there is a joint dependence between the present and the past. For this example, we need to utilize all of the tools developed in the paper. The functional Itô formula gives the evolution of the present marginal effects of the policies in terms of horizontal and vertical derivatives. The total derivative formula straightforwardly gives the evolution of the conditional expectation process in terms of the time and stochastic derivative. We also consider a deterministic setting with hysteresis in Section 6.5. As there is no stochasticity, there is only evolution with respect to time or, rather, the history of policies. We decompose its effects into the effects of the present and the past history. The present marginal effect of policy then evolves as the first-order horizontal and the vertical derivatives. The future marginal effects of policies evolve as their time derivative. In Section 6.6 we consider an example with a tipping point in which the time when the stochastic process reaches the maximum becomes a reference point.

6.1 The user's guide

In order to apply results from the previous section one needs to calculate a number of derivatives of the functional $(h_t)_{0 \le t \le T}$. In calculation of the first order conditions (10), we need to find a Fréchet derivative of h_t:

$$h_t(c + z) - h_t(c) = \partial_{c_t} h_t(c) z_t + \int_0^t \delta_s h_t(c) z_s ds + o(||z||).$$

For the application of Dupire's functional Itô's formula and finding the SDE for the present effects of the path in Section 5.1.2 and equation (12) we need to calculate the horizontal derivative $\Delta_t \partial_{c_t} h_t(c)$ and two vertical derivatives $\partial_{c_t}^2 h_t(c)$, $\partial_{c_t}^3 h_t(c)$. For the application of the total derivative formula (13) and finding the evolution equation for the future effects of the path we need to calculate two types of derivatives of $\delta_t h_s(c)$: (a) time derivative $\frac{\partial}{\partial t} \delta_t h_s(c)$, and (b) the Malliavin derivatives $D_t \delta_t h_s(c)$,

   In other words, we first break the first order condition for the effects of policies of the present and the effects of the past. The dependence of variables in the past implies that any change in a

variable affects all the future states. We then use the functional Itô formula to derive the evolution equation for the effect on the present and the total derivative formula for the effects on the future. The total derivative formula also requires calculations of the Malliavin derivatives.

6.2 No hysteresis: state-dependent perturbations

We start with the simplest case – the state-dependent perturbation where h is a function of the current state rather than a functional. Consider the perturbation $h_t(c) = f(c_t)$, where $f : \mathbb{R} \to \mathbb{R}$ is a smooth function. The list of derivatives becomes the following.

The Fréchet derivative of the functional is just the derivative of f :

$$h_t(c + z) - h_t(c) = f(c_t + z_t) - f(c_t) = f'(c_t)z_t + o(||z||).$$

Correspondingly, the present marginal effects of policy are given by the regular derivative

$$\partial_{c_t} h_t(c) = f'(c_t),$$

and the effects of the past are

$$\delta_s h_t = 0 \text{ for } s < t.$$

In particular, we see that $h_t \in \mathcal{A}_t$ and $\delta_s h_t \in \mathcal{A}_t$.

We now turn to the analysis of the present effect of the policy path. The horizontal derivative $\Delta_t \left(\partial_{c_t} h_t(c) \right)$ is equal to zero. Indeed, it is defined as a limit

$$\Delta_t \partial_{c_t} h_t(c) = \lim_{\epsilon \to 0} \frac{\partial_{c_{t+\epsilon}} h_{t+\epsilon}(c_{\cdot,\epsilon}) - \partial_{c_t} h_t(c)}{\epsilon},$$

where $c_{\cdot,\epsilon}$ is an extension of the path c from $[0, t]$ to $[0, t + \epsilon]$ by $c_{s,\epsilon} = c_t$ for $t \le s \le t + \epsilon$. It follows that[12]

$$\Delta_t \partial_{c_t} h_t(c) = \lim_{\epsilon \to 0} \frac{f'(c_{t+\epsilon,\epsilon}) - f'(c_t)}{\epsilon} = \lim_{\epsilon \to 0} \frac{f'(c_t) - f'(c_t)}{\epsilon} = 0.$$

Vertical differentiation of the present marginal effect of policy is the regular differentiation of f :

$$\partial_{c_t}^2 h_t(c) = f''(c_t), \ \partial_{c_t}^3 h_t(c) = f'''(c_t).$$

---

[12]If we had $h_t(c) = f(c_t, t)$ the only difference would be that $\Delta_t \partial_{c_t} h_t(c) = \partial_{c_t,t}^2 f(c_t, t)$

We now turn to the analysis of the expected future marginal effects. The derivative in t of the future marginal effects in period $s : \delta_t h_s(c)$, i.e. $\frac{\partial}{\partial t}\delta_t h_s(c)$ is equal to zero, since $\delta_t h_s(c) = 0$:

$$\frac{\partial}{\partial t}\delta_t h_s(c) = 0.$$

The Malliavin derivative of $\delta_t h_s(c)$ is also zero, since $\delta_t h_s(c) = 0$:

$$D_t \delta_t h_s(c) = 0.$$

The formula for the dynamics of the stochastic elasticity (14) becomes:

$$dC_t = -\left(\underbrace{\Delta_t \partial_{c_t} h_t(w^t)}_{=0}\,dt + \frac{1}{2}\underbrace{\partial_{c_t}^3 h_t(w^t)}_{=f'''(w_t)}dt + \underbrace{\partial_{c_t}^2 h_t(w^t)}_{=f''(w_t)}dw_t\right) -$$
$$- \underbrace{\left(-\delta_t h_t(w^t)dt + E\left[\int_t^T \partial_t \delta_t h_s(w^s)ds\Big|\mathcal{F}_t\right]dt + E\left[\int_t^T D_t\delta_t h_s(w^s)ds\Big|\mathcal{F}_t\right]dw_t\right)}_{=0},$$

and

$$dC_t = -\frac{1}{2}f'''(w_t)dt - f''(w_t)dw_t.$$

We summarize the result. First, the present effect of the path from the Section 5.1.2 are given by the standard space derivatives. The functional Itô formula reduces to the usual Itô's formula. Second, all of the future effects of the path from are zero, hence, the conditional expectation is zero.

One can, of course, immediately get the result of this section by applying the Itô formula to $f'(c_t^*)$, see the Appendix.

6.3 Cumulative hysteresis

Consider the hysteresis functional $h_t(c) = c_t \int_0^t c_s ds$. Here, the path-dependence enters cumulatively as the integral of the past realization. The cumulative hysteresis is then multiplicative with the current policy c_t.

We start with the Fréchet derivative of h_t that appears in the first order conditions of the problem (10). It is given by varying the whole path $c_{[0,t]}$ by a variation $z_{[0,t]}$:

$$h_t(c+z) - h_t(c) = (c_t + z_t)\int_0^t (c_s + z_s)\,ds - c_t\int_0^t c_s ds = \int_0^t c_t z_s ds + z_t\int_0^t c_s ds + o(||z||).$$

Correspondingly, the present marginal effects of policy are given by the integral of the path of previous policies

$$\partial_{c_t} h_t(c) = \int_0^t c_s ds.$$

As all previous c_s enter identically in the integral, the effects of the past marginally contributes c_t:

$$\delta_s h_t = c_t \text{ for } s < t.$$

In particular, we see that $h_t \in \mathcal{A}_t$ and $\delta_s h_t \in \mathcal{A}_t$.

The first order process is given by (11), recalling that $c_t^* = w_t$:

$$C_t = -\partial_{c_t} h_t(c_{[0,t]}^*) - E\left[\int_t^T \delta_t h_s(c_{[0,s]}^*) ds \Big| \mathcal{F}_t\right] = -\underbrace{\int_0^t w_s ds}_{\text{Present effect}} - \underbrace{E\left[\int_t^T w_s ds \Big| \mathcal{F}_t\right]}_{\text{Future effect}}.$$

Note that in the future effects it is $\delta_t h_s(c^t) = c_s^* = w_s$ that enters as it shows how policy in period t ($t \leq s \leq T$) affects future period s.

We now turn to calculation of the derivatives needed for the application of the Dupire's functional Itô formula (12). Recall that the horizontal and two vertical derivatives are needed. The horizontal derivative of the present effects of the path $\partial_{c_t} h_t(c)$ is simply the time derivative of the integral:

$$\Delta_t \left(\partial_{c_t} h_t(c)\right) = \Delta_t \left(\int_0^t c_s^* ds\right) = c_t^* = w_t.$$

The vertical differentiation of the present effects of the path $\partial_{c_t} h_t(c)$ is given by bumping the path and is equal to zero:

$$\partial_{c_t} \left(\partial_{c_t} h_t(c)\right) = \partial_{c_t} \left(\int_0^t c_s^* ds\right) = 0,$$

$$\partial_{c_t}^3 h_t(c) = 0.$$

We now turn to characterizing the expected future effects of policy using the total derivative formula (8). Since the marginal effect of the policy at time t on time s, $\delta_t h_s(w^s) = w_s$, does not depend on time, we get

$$\partial_t \left(\delta_t h_s(w^s)\right) = 0.$$

36

The Malliavin derivative is also simple as it measures the sensitivity of w^s to the shock w_t :

$$D_t\left(\delta_t h_s\left(w^s\right)\right) = D_t w_s = 1,$$

and the the stochastic derivative term of the future marginal costs becomes

$$E\left[D_t \int_t^T \delta_t h_s(w^s)ds \middle| \mathcal{F}_t\right] = T - t.$$

This has a natural interpretation. The stochastic derivative measures the impact of the stochastic shock dw_t on the future marginal effects of policies. Those are represented by the integrals $\int_t^T w_s ds$. A shock at time t affects the future marginal effect of policy in each period s as $D_t w_s = 1$. Since there are $(T-t)$ future periods, the shock dw_t has the effect $(T-t)$. In other words, the stochastic shocks for early periods t have longer lasting impact than for the later periods.

We then gather the present and the future terms to determine the dynamics of the first order process in (14):

$$dC_t = -\left(\underbrace{\Delta_t \partial_{c_t} h_t(w^t)}_{=w_t} dt + \frac{1}{2}\underbrace{\partial^3_{c_t} h_t(w^t)}_{=0}dt + \underbrace{\partial^2_{c_t} h_t(w^t)}_{=0}dw_t\right) -$$
$$-\left(\underbrace{-\delta_t h_t(w^t)}_{=w_t} dt + \underbrace{E\left[\int_t^T \partial_t \delta_t h_s(w^s)ds \middle| \mathcal{F}_t\right]}_{=0} dt + \underbrace{E\left[\int_t^T D_t \delta_t h_s(w^s)ds \middle| \mathcal{F}_t\right]}_{=T-t} dw_t\right),$$

or

$$dC_t = -\left(T - t\right)dw_t.$$

Remark 3. There are two alternative ways, without using our methodology, to derive the result. The first is to note that $E\left[\int_t^T w_s ds \middle| \mathcal{F}_t\right] = (T-t)w_t$ and then

$$dC_t = -w_t dt - \left((T-t)dw_t - w_t dt\right) = -\left(T-t\right)dw_t.$$

The second approach that works in a variety of other circumstances is in extracting a martingale and using the Clark-Ocone formula to provide its explicit characterization in terms of Malliavin derivatives. We can rewrite

$$E\left[\int_t^T w_s ds \middle| \mathcal{F}_t\right] = E\left[\int_0^T w_s ds \middle| \mathcal{F}_t\right] - \int_0^t w_s ds,$$

37

and

$$C_t = - \left(\int_0^t w_s ds + E \left[\int_0^T w_s ds \middle| \mathcal{F}_t \right] - \int_0^t w_s ds \right) = -E \left[\int_0^T w_s ds \middle| \mathcal{F}_t \right].$$

The Clark-Ocone formula gives the representations of the martingale where $M_t = E \left[\int_0^T w_s ds \middle| \mathcal{F}_t \right]$

as

$$dM_t = E \left[D_t \int_0^T w_s ds \middle| \mathcal{F}_t \right] dw_t = E \left[\int_t^T \underbrace{D_t w_s}_{=1} ds \middle| \mathcal{F}_t \right] dw_t = (T - t) dw_t.$$

We expand on this class of examples in Boulatov, Riabov, and Tsyvinski (2020) where we analyze a more general class of environments which can be solved by extracting a martingale and not using the main tools of this paper – the functional Itô formula and the total derivative formula.

6.4 Hysteresis with time and past dependency

Consider the functional $h_t(c^t) = h(c_t, \int_0^t a_{t,s} c_s ds)$, it describes the interaction of the policy at the present moment of time, and average of past values. The averaging is defined by the smooth kernel $a_{t,s}$ which measures how policy in period s affects period t. For example, $a_{t,s} = \frac{1}{t}$ corresponds to the usual averaging. Importantly, the kernel depends on current time t.

This example captures three important parts of history-dependent policies. First, there is an effect of the present which is represented by the first argument, c_t. Second, there is an effect of the past which is represented by the second argument, $\int_0^t a_{t,s} c_s ds$. Here, the past enters as the integral of the path of the previous consumptions where the path enters through the kernel $a_{t,s}$ that depends both on the current policy t and the path of previous policies.[13] Third, there is a joint dependence $h(.,.)$ between the present and the past.

The Fréchet derivative of h_t that appears in the first order conditions (10) is given by changing the whole path c^t by a variation z^t:

$$h_t(c + z) - h_t(c) = h \left(c_t + z_t, \int_0^t a_{t,s} (c_s + z_s) ds \right) - h_t(c) =$$

$$= z_t h_1' \left(c_t, \int_0^t a_{t,s} c_s ds \right) + \int_0^t a_{t,s} h_2' \left(c_t, \int_0^t a_{t,r} c_r dr \right) z_s ds + o(||z||).$$

Correspondingly, the present marginal effects of policy are given by the integral of the path of

---

[13] One can easily modify this part to be some more complicated path-dependent object – for example, $\int_0^t \alpha \left(c_{[0,s]} \right) ds$, where α is a functional of the path or to have arbitrary interaction of the past effects; or allow for $h_t(c) = \int_0^t \int_0^t g_t(c_s, c_r) ds dr$, ... where now there is a joint dependence on the past values c_s and c_r via a repeated integral.

previous policies

$$\partial_{c_t} h_t(c) = h_1' \left(c_t, \int_0^t a_{t,s} c_s ds \right).$$

The marginal effect of policy in time t n the past period s ($s \le t \le T$) is given by:

$$\delta_s h_t(c^t) = a_{t,s} h_2' \left(c_t, \int_0^t a_{t,r} c_r dr \right).$$

In particular, we see that $h_t \in \mathcal{A}_t$ and $\delta_s h_t \in \mathcal{A}_t$.

The first order condition (10) is

$$w_t = c_t^{(\epsilon)} + \epsilon h_1' \left(c_t, \int_0^t a_{t,s} c_s ds \right) + \epsilon E \left[\int_t^T a_{s,t} h_2' \left(c_s, \int_0^s a_{s,r} c_r dr \right) ds \bigg| \mathcal{F}_t \right];$$

and the stochastic elasticity (11) is given by, recalling that $c_t^* = w_t$:

$$C_t = \underbrace{-h_1' \left(w_t, \int_0^t a_{t,s} w_s ds \right)}_{\text{Present effect}} - \underbrace{E \left[\int_t^T a_{s,t} h_2' \left(w_s, \int_0^s a_{s,r} w_r dr \right) ds \bigg| \mathcal{F}_t \right]}_{\text{Future effect}}.$$

Note that in the future effects it is $\delta_t h_s(c^t) = a_{s,t} h_2' \left(w_s, \int_0^s a_{s,r} w_r dr \right)$ that enters as the kernel $a_{s,t}$ measures how policy in period t ($t \le s \le T$) affects future period s.

The conditional expectation $E \left[\int_t^T a_{s,t} h_2' \left(w_s, \int_0^s a_{s,r} w_r dr \right) ds \bigg| \mathcal{F}_t \right]$ is intractable for direct computations, as it requires infinite-dimensional integration over the distribution of the process $(w(r) - w(t))_{t \le r \le s}$. However, the semimartingale decomposition is obtained immediately after an application of the total derivative formula (8).

We now turn to calculation of the derivatives needed for the application of the Dupire's functional Itô formula (12). These derivatives are used in determining the decomposition of the present effects of the path $\partial_{c_t} h_t(c)$ in the first order conditions of the perturbed problem (10). Recall that the horizontal and two vertical derivatives are needed.

The horizontal derivative of the present effects of the path $\partial_{c_t} h_t(c)$ is given by freezing the path at time t and extending it with time:

$$\Delta_t \partial_{c_t} h_t(c) = \Delta_t h_1' \left(w_t, \int_0^t a_{t,s} w_s ds \right) = h_{12}'' \left(w_t, \int_0^t a_{t,s} w_s ds \right) \left(a_{t,t} w_t + \int_0^t (\partial_t a_{t,s}) w_s ds \right).$$

Vertical differentiation of the present effects of the path $\partial_{c_t} h_t(c)$ is given by bumping the path. It

reduces to the differentiation in the first argument:

$$\partial_{c_t}^2 h_t(c) = \partial_{c_t}\left(h_1'\left(w_t, \int_0^t a_{t,s}w_s ds\right)\right) = h_{11}''\left(w_t, \int_0^t a_{t,s}w_s ds\right),$$

$$\partial_{c_t}^3 h_t(c) = \partial_{c_t}^2\left(h_1'\left(w_t, \int_0^t a_{t,s}w_s ds\right)\right) = h_{111}'''\left(w_t, \int_0^t a_{t,s}w_s ds\right).$$

Note that since we are already finding the derivatives of the marginal effect $\partial_{c_t}h_t(c)$, the mixed time-derivative and the second and third order space derivatives appear (rather than just the time and first and second order derivatives).

We next calculate the derivatives needed for the application of the total derivative formula (13) to characterize the conditional expectation of the marginal effects of policy. The time derivative is given by

$$\partial_t\left(\delta_t h_s(w^s)\right) = \partial_t\left(a_{s,t}h_2'\left(w_s, \int_0^s a_{s,r}w_r dr\right)\right) = (\partial_t a_{s,t})h_2'\left(w_s, \int_0^s a_{s,r}w_r dr\right).$$

The Malliavin derivative is given by

$$D_t\left(\delta_t h_s(w^s)\right) = D_t\left(a_{s,t}h_2'\left(w_s, \int_0^s a_{s,r}w_r dr\right)\right) =$$

$$= a_{s,t}h_{12}''\left(w_s, \int_0^s a_{s,r}w_r dr\right)\underbrace{D_t w_s}_{=1} + a_{s,t}h_{22}''\left(w_s, \int_0^s a_{s,r}w_r dr\right)\underbrace{D_t\left(\int_0^s a_{s,r}w_r dr\right)}_{=\int_t^s a_{s,r}D_t w_r dr = \int_t^s a_{s,r}dr}.$$

Gathering the terms we find the differential of the first order process:

$$dC_t = -h_{12}''\left(w_t, \int_0^t a_{t,s}w_s ds\right)\left(a_{t,t}w_t + \int_0^t (\partial_t a_{t,s})w_s ds\right)dt-$$

$$- \left(h_{11}''\left(w_t, \int_0^t a_{t,s}w_s ds\right)dw_t + \frac{1}{2}h_{111}'''\left(w_t, \int_0^t a_{t,s}w_s ds\right)dt\right)-$$

$$- \left(-\underbrace{a_{t,t}h_2'\left(w_t, \int_0^t a_{t,r}w_r dr\right)}_{\delta_t h_t(w^t)}dt + \underbrace{E\left[\int_t^T (\partial_t a_{s,t})h_2'\left(w_s, \int_0^s a_{s,r}w_r dr\right)ds\bigg|\mathcal{F}_t\right]}_{E\left[\int_t^T \partial_t \delta_t h_s(w^s)ds\big|\mathcal{F}_t\right]}dt\right)+$$

$$+ \underbrace{E\left[\int_t^T\left(a_{s,t}h_{12}''\left(w_s, \int_0^s a_{s,r}w_r dr\right) + a_{s,t}h_{22}''\left(w_s, \int_0^s a_{s,r}w_r dr\right)\int_t^s a_{s,r}dr\right)ds\bigg|\mathcal{F}_t\right]}_{E\left[D_t\int_t^T \delta_t h_s(w^s)ds\big|\mathcal{F}_t\right]}dw_t.$$

Remark 4. The presented calculation can be done without the use of the total derivative formula when the occurrence of t under the integral in the conditional expectation can be removed, e.g. when $a_{s,t} = \tilde{g}(t)g(s)$, and a martingale can be extracted. Indeed, then future effects can be represented as

$$\tilde{g}(t)E\left[\int_t^T g(s)h_2'\left(w_s, \int_0^s a_{s,r}w_r dr\right)ds\Big|\mathcal{F}_t\right] = -\tilde{g}(t)\int_0^t g(s)h_2'\left(w_s, \int_0^s a_{s,r}w_r dr\right)ds + \tilde{g}(t)M_t,$$

where
$$M_t = E\left[\int_0^T g(s)h_2'\left(w_s, \int_0^s a_{s,r}w_r dr\right)ds\Big|\mathcal{F}_t\right].$$

Now the semimartingale decomposition immediately follows from the Itô formula. This type of examples (with exponential functions $g, \tilde{g}$) where studied by Detemple and Zapatero (1991) in the context of asset prices under habit formation.

6.5 Deterministic hysteresis

In this section, we consider a deterministic case. Assume that the underlying process shock θ is deterministic: $d\theta_t = b(\theta_t)dt$. Then optimal policies c^* and c^ϵ are deterministic as well:

$$c_t^* = \theta_t.$$

Our assumption of h_t belonging to the class $\mathcal{A}_t$ allows us to neatly decompose its effects into those of the effects of the present and the past history

$$c_t^\epsilon = \theta_t - \epsilon\left(\partial_{c_t}h_t(c^\epsilon) + \epsilon\int_t^1 \delta_t h_s(c^\epsilon)ds\right).$$

Differentiating the latter relation in ϵ at $\epsilon = 0$ we get

$$C_t = -\left(\partial_{c_t}h_t(c^*) + \int_t^1 \delta_t h_s(c^*)ds\right).$$

Now, apply the functional Itô formula (7) to $\partial_{c_t}h_t(c^*)$ and note that there are no second order vertical derivatives of $\partial_{c_t}h_t(c^*)$ due to the absence of stochasticity:

$$d\left(\partial_{c_t}h_t(c^*)\right) = \left(\Delta_t\partial_{c_t}h_t(c^*) + \partial_{c_t}^2 h_t(c^*)b(c_t^*)\right)dt.$$

The future marginal effects of policy are also deterministic and we can simply differentiate $\int_t^1 \delta_t h_s(c^*) ds$ in time.

$$dC_t = -\left(\underbrace{\Delta_t \partial_{c_t} h_t(c^*) + \partial_{c_t}^2 h_t(c^*) b(c_t^*)}_{d(\partial_{c_t} h_t(c^*))} \right) dt - \left(\underbrace{-\delta_t h_t(c^*) + \int_t^1 \frac{\partial}{\partial t} \delta_t h_s(c^*) ds}_{d(\int_t^1 \delta_t h_s(c^*) ds)} \right) dt.$$

The present effects of the policy $\partial_{c_t} h_t(c^*)$ change with time as the horizontal derivative Δ_t and with the movement in the path of the policy as the vertical derivative ∂_{c_t}. The cumulative future marginal effects of the policy $\int_t^1 \delta_t h_s(c^*) ds$ change as the relative difference between the time change in the cumulant of the future marginal effects $\frac{\partial}{\partial t} \delta_t h_s(c^*) ds$ relative to today's marginal effect of the policy $\delta_t h_t(c^*)$. This case is interesting as a stand alone result that allows to focus on history dependence without stochasticity.

6.6   A "tipping point"

The main premise of the literature on tipping points discussed in 3.2 is that there is some level of emissions the crossing of which leads to a different behavior of the climate system. For example, upon reaching the threshold the damages become larger or become irreversible. We now show an example where instead of considering a threshold we focus on the time when the climate variable achieves its maximum upto any given period of time.

Let θ_t be the time when the Brownian motion w achieves its maximum over $[0, t]$:

$$\theta_t = \arg \max_{s \in [0,t]} w_s$$

and it is known that θ_t is a.s. unique. Let $f(x)$ be an absolutely continuous function such that $f(x) = 0$ for $x \le 0$. Consider the objective function

$$E\left[\int_0^T \left(-\frac{1}{2}(c_t - w_t)^2 - \int_0^t f(s - \theta_t) c_s ds \right) dt \right].$$

One can think of this setting as follow. In each period t, we consider the time θ_t when the maximal level of w has been achieved – say, the time of the temperature record. The damages then are counted as the weighted by $f(s - \theta_t)$ emissions from the time of the record. When the new record is achieved, the weighting restarts. This example can be significantly expanded by having the weighting functions change with time or the weighting functions that weigh both the time prior

42

to θ_t and the time after θ_t (a form of highlighting the "salience" of the record time) but we chose to present the simple form here. That is, there is a salient "tipping" or reference point following which the damages change their behavior.

The first-order conditions give us the optimal policy in a closed form is given by

$$c_t = w_t - E\left[\int_t^T f(t - \theta_s)ds \middle| \mathcal{F}_t\right].$$

The process $\xi_t = \int_t^T f(t - \theta_s)ds$ is absolutely continuous and square integrable. Indeed,

$$\xi_t = \int_t^T \int_0^t f'(x - \theta_s)dxds.$$

Let us represent ξ_t using the Clark-Ocone formula

$$\xi_t = \mathbb{E}\xi_t + \int_0^T g_{t,s}dw_s.$$

By the total derivative formula of Proposition 1, we immediately find that c is an Itô process and

$$dc_t = \left(1 - g_{t,t}\right)dw_t - \mathbb{E}[\partial_t\xi_t|\mathcal{F}_t]dt. \tag{16}$$

Remark 5. Let us try to get this result using standard methods. We compute the process c_t explicitly as a functional of the Wiener process w. At first we find the conditional expectation

$$E[f(t - \theta_s)|\mathcal{F}_t] = E\left[\int_0^\infty f'(x)1_{x<t-\theta_s}dx \middle| \mathcal{F}_t\right] =$$

$$= \int_0^t f'(x)P(\theta_s < t - x|\mathcal{F}_t)dx = \int_0^t f'(x)P\left(\max_{[0,t-x]} w > \max_{[t-x,s]} w \middle| \mathcal{F}_t\right)dx =$$

$$= \int_0^t f'(x)1_{\max_{[0,t-x]} w > \max_{[t-x,t]} w} P\left(\max_{[0,t-s]} w < z\right)\Bigg|_{z=\max_{[0,t-x]} w - w_t} dx =$$

$$= \int_0^t f'(x)1_{\max_{[0,t-x]} w > \max_{[t-x,t]} w}\left(2\Phi\left(\frac{\max_{[0,t-x]} w - w_t}{\sqrt{t-s}}\right) - 1\right)dx.$$

Hence,

$$
\begin{aligned}
c_t &= w_t - \int_0^t f'(x) 1_{\max_{[0,t-x]} w > \max_{[t-x,t]} w} \int_t^T \left(2\Phi \left(\frac{\max_{[0,t-x]} w - w_t}{\sqrt{t-s}} \right) - 1 \right) ds\, dx = \\
&= w_t - \int_0^{t-\theta_t} f'(x) \int_t^T \left(2\Phi \left(\frac{\max_{[0,t]} w - w_t}{\sqrt{t-s}} \right) - 1 \right) ds\, dx = \\
&= w_t - f(t - \theta_t) \int_t^T \left(2\Phi \left(\frac{\max_{[0,t]} w - w_t}{\sqrt{t-s}} \right) - 1 \right) ds.
\end{aligned}
$$

We get the representation of the form

$$
c_t = w_t - f(t - \theta_t) g(\max_{[0,t]} w - w_t).
$$

Now one can try two standard approaches to get the semimartingale decompostion. The first is to apply the Itô formula. In order to do this, we see that

$$
c_t = F(t, w_t, \theta_t, M_t),
$$

where $M_t = \max_{[0,t]} w$, and $F(t, x, y, z) = x - f(t - y) g(z - x)$. Hence the Itô formula will lead to a semimartingale decomposition that contains terms $d\theta_t$, dM_t, and the sum over jumps of the process θ_t. It is not obvious that the process c_t is an Itô process.

The second approach is to apply the functional Itô formula. In order to do this the functional c_t must be extended to a $C_b^{1,2}$-functional (in the sense of Cont and Fournie) on the space of cadlag paths. However, if such extension is possible, vertical derivatives of c_t must be equal to zero and the functional Itô formula would lead to semimartingale decomposition $dc_t = \alpha_t dt$, which is not the case.

Instead, the application of our methodology give a straightforward and compact answer of the equation (16).

7 Discussion

In this section, we discuss some of the more technical issues behind the results.

7.1 Why do we need both the functional Itô formula and the total derivative formula?

The stochastic elasticity C_t^h is described in (11) via two terms: a vertical derivative $\partial_{c_t} h_t(c)$ of the perturbation functional h_t and the conditional expectation $\mathbb{E}[\int_t^1 \delta_t h_s(c^*) ds | \mathcal{F}_t]$. Both are functionals of the path $(c_s)_{0 \leq s \leq t}$ but we need to use different methods to study them.

We assume that the functional $\partial_{c_t} h_t(c)$ satisfies conditions from Dupire (2019). This is a natural assumption, since $\partial_{c_t} h_t(c)$ is a well-defined time-dependent functional of the path $(c_s)_{0 \leq s \leq t}$.

However, in the case of the functional $\mathbb{E}[\int_t^1 \delta_t h_s(c^*) ds | \mathcal{F}_t]$ the same assumption is not generally suitable. The application of the Dupire's functional Itô formula requires that

$$\mathbb{E}\left[\int_t^1 \delta_t h_s(c^*) ds | \mathcal{F}_t\right] = q_t(m),$$

where m is a certain semimartingale and $(q_t)_{t \in [0,T]}$ is smooth in the sense of Dupire (2019) and Cont and Fournie (2013) family of path-dependent functionals. It may be a difficult stand-alone problem even to verify that such representation holds. Consider, for example, the case when

$$\mathbb{E}\left[\int_t^1 \delta_t h_s(c^*) ds | \mathcal{F}_t\right] = \int_0^t f_t(c_s^*) dc_s^*,$$

and f_t is smooth in t. The stochastic integral is defined for almost all realizations of c^* only and it is unclear how we can extend it smoothly to all continuous paths.[14]

Instead we assume that the expression under the conditional expectation is smooth in t and then apply the total derivative formula in Proposition 1 to prove directly that C^h is an Itô process and to find its semimartingale decomposition as well.

In other words, the functional Itô formula is useful when the functional of the path is already well-defined. When faced with a conditional expectation process such as the one we considered here or that frequently occurs in a variety of other economic problems, the total derivative formula allows to straightforwardly calculate its semimartingale decomposition.

7.2 Assumptions on smoothness of the functionals

Our results are valid for path-dependent functionals $h = (h_t)_{0 \leq t \leq T}$ such that: (1) $h_t \in \mathcal{A}_t$; (2) $t \to \partial_{c_t} h_t$ is horizontally and twice vertically differentiable; (3) $t \to \delta_t h_s(c^s)$ is absolutely continuous;

---

[14]Recall that the functional Itô formula is applicable to functionals defined on a larger space of cádlág paths.

(4) $w \to \delta_t h_s(c^s(w))$ is Malliavin differentiable.

The first restriction that h_t belongs to the class $\mathcal{A}_t$ allows to calculate the first order conditions in a tractable form. The second restriction is a typical condition needed for the functional Itô formula to be valid. Last two conditions are imposed for the ease of presentation and also they are satisfied in all our examples. They are needed for the total derivative formula to be applicable for the process

$$t \to E[\int_t^T \delta_t h_s(c^s)ds|\mathcal{F}_t].$$

However, these conditions can be considerably relaxed. It is enough to find an absolutely continuous square integrable process ξ_t, such that

$$E[\int_t^T \delta_t h_s(c^s)ds|\mathcal{F}_t] = E[\xi_t|\mathcal{F}_t].$$

As a simple example consider the process

$$\delta_t g_s(c_s) = w_t.$$

This process is not absolutely continuous, however, $E[w_t|\mathcal{F}_t] = E[w_T|\mathcal{F}_t]$ and one can take $\xi_t = (T-t)w_T$.

7.3 Relationship to the Clark-Ocone formula

It is instructive to compare Proposition 1 to the well-known Clark-Ocone formula. Recall that $\mathcal{F} = (\mathcal{F}_t)_{t \in [0,T]}$ is a filtration generated by Wiener process $(w_t)_{t \in [0,T]}$. Every $\mathcal{F}$-martingale $(M_t)_{t \in [0,T]}$ can be represented as a conditional expectation process. Indeed, from the definition of a martingale we get $M_t = \mathbb{E}[M_T|\mathcal{F}_t]$. Conversely, every process of the form

$$Z_t = \mathbb{E}[\xi|\mathcal{F}_t],$$

where ξ is an integrable random variable, is a martingale. This follows from the basic properties of conditional expectations:

$$\mathbb{E}[Z_{t+h}|\mathcal{F}_t] = \mathbb{E}[\mathbb{E}[\xi|\mathcal{F}_{t+h}]|\mathcal{F}_t] = \mathbb{E}[\xi|\mathcal{F}_t] = Z_t.$$

Moreover, every $\mathcal{F}$-martingale $Z_t = \mathbb{E}[\xi|\mathcal{F}_t]$ is an Itô process and if ξ is Malliavin differentiable, the Clark-Ocone formula gives

$$Z_t = Z_0 + \int_0^t \mathbb{E}[D_s\xi|\mathcal{F}_s]dw_s,$$

and the differential of Z :

$$dZ_t = \mathbb{E}[D_t\xi|\mathcal{F}_t]dw_t.$$

The processes that we consider in the Proposition 1 are of the different type - they are of the form

$$X_t = \mathbb{E}[\xi_t|\mathcal{F}_t]$$

with $\xi = (\xi_t)_{t \in [0,T]}$ being an absolutely continuous process. Such process are not martingales in general. The simplest example is provided with deterministic non-constant process ξ, e.g. if we take $\xi_t = t$, then $X_t = t$ and this is obviously not a martingale.

The Clark-Ocone formula is then a partial case of Proposition 1 corresponding to constant in t process $(\xi_t)_{t \in [0,T]}$. For the proof and general versions of Proposition 1 we refer to the Appendix.

7.4   Discussion of the optimal control approach

A natural question is whether we can use optimal control and dynamic programming to simplify the analysis.

One of the standard approaches for solving optimization problems is the dynamic programming principle. It can be used for maximization of functionals of the type (see see Fabbri, Gozzi, and Swiech 2017 for the exposition of infinitely dimensional problems):

$$V(a) = E \int_0^T l(t, x^t, a(t))dt,$$

where a is the control process, the process x satisfies certain SDE whose coefficients depend on the control a as well. This framework is not well suited for our problem, as in our case the dependence of the expression under the integral is a path-dependent functional of the policy c:

$$V(c) = E \int_0^T \left(-\frac{1}{2}(c_t - w_t)^2 - h_t(c^t) \right) dt.$$

Hence, to fit the framework of the optimal control we must consider the whole path c^t as the value

of the control $a(t)$. That is, the control space becomes the set of paths $\cup_{t \in [0,T]} C[0,t]$. This leads to the consideration of the cost functional

$$V(t,x,c^t) = \max_{a \in C[t,T]} E_{t,x} \int_t^T \left(-\frac{1}{2}(a_s - w_{s-t} - x)^2 - h_s(c^t \otimes a^s) \right) ds.$$

The HJB equation then becomes a path-dependent PDE that contains infinite-dimensional optimization over trajectories of the control. Thus the analysis may be as complicated as our original problem. For example, Cosso et. al (2018) use functional Itô calculus to write the path-dependent HJB equation, prove that the value function is a viscosity solution and prove a partial comparison principle.

8 Conclusion

Path-dependent policies are a feature of a number of economic models. Doing a lot of sports only when young leads to a very different health outcome than exercising lightly throughout life, even though the cumulative lifetime exercise may be the same. Moreover, the importance of the previous pattern of exercise may change with time. For example, intensively training several times a week is best when young while a steady pattern of light exercise during a week may be preferable when older. Missing a credit card payment twice in a row in a ten year period leads to a very different credit history than missing a payment once every five years. The consequence of the same sequence of non-payments for the credit history may be quite different in a recession versus a boom. These policies are not just a function of the current state or time but depend on the whole trajectory of the past actions. Moreover, the dependency itself is changing with time.

We developed a methodology for analysis of general class of policy with path-dependent effects in an uncertain environment with forward looking agents. The primary difficulty that arises in such models is that the optimizing agents foresee such path dependency and the actions they take incorporate the expectation of the future effects. We show that three ingredients are needed for the analysis of such problems. First, we introduced a general class of path-dependent functionals that allows to tractably write the first order conditions for the problem. Second, the recently developed functional Itô calculus allows to describe the dynamics of the present effect of the past choices. Third, the total derivative formula that we develop in this paper allows to derive the dynamics of the conditional expectation processes of the future effects of today's choices. Our analysis shows that even when policy has very small contemporaneous effects, it may have large effects

due to either interaction with the past choices or due to the expected future effects. The effects of the past are represented by the magnitude of horizontal derivatives and higher order vertical derivative. The future expected effects are determined by the magnitude of the time derivative and by the stochastic derivatives. We believe that the methodology we develop in, on purpose, a stark underlying environment will facilitate analysis of a wide class of other economic problems.

9   Appendix

9.1   Remark 1 in Section 4.1

The objective function (3) can be written as follows:

$$
\mathbb{E}\int_0^T (-\frac{1}{2}\left(c_t - \theta_t\right)^2 - \epsilon c_{\frac{t}{2}})dt = \mathbb{E}\int_0^T (-\frac{1}{2}\left(c_t - \theta_t\right)^2)dt - \mathbb{E}\int_0^T \epsilon c_{\frac{t}{2}}dt =
$$
$$
= \mathbb{E}\int_0^T (-\frac{1}{2}\left(c_t - \theta_t\right)^2)dt - 2\mathbb{E}\int_0^{\frac{T}{2}} \epsilon c_s ds =
$$
$$
= \mathbb{E}\int_0^{\frac{T}{2}} (-\frac{1}{2}\left(c_t - \theta_t\right)^2 - 2\epsilon c_t)dt + \mathbb{E}\int_{\frac{T}{2}}^T (-\frac{1}{2}\left(c_t - \theta_t\right)^2)dt.
$$

The first-order conditions become

$$
\begin{cases}
c_t = \theta_t - 2\epsilon, \ t < \frac{T}{2} \\
c_t = \theta_t, \ t > \frac{T}{2},
\end{cases}
$$

which implies that the optimal policy has a jump at $t = \frac{T}{2}$.

9.2   An example with contemporaneous policies

In this section, we develop a simple example of introducing additional effects of policies that can be easily handled by the usual Itô formula. That is, there is no hysteresis and policy only has contemporaneous effects.

Consider an additional effect of policies given by $f(c_t)$, where $f : \mathbb{R} \to \mathbb{R}$ is a smooth function:

$$
\max_c \mathbb{E}\int_0^T (-\frac{1}{2}\left(c_t - w_t\right)^2 - \epsilon f(c_t))dt,
$$

and ϵ is a parameter. The first order conditions are given by

$$
c_t^\epsilon = w_t - \epsilon f'\left(c_t^\epsilon\right), \tag{17}
$$

and equate the marginal benefit of the policy tracking the process w_t with the additional marginal effects $\epsilon f'\left(c_t^\epsilon\right)$.

9.2.1 Stochastic elasticity and its dynamics

We are particularly interested in the case of $\epsilon \to 0$. One can think about this as a small parameter asymptotics for the problem. Let C_t^f be a first variation process:

$$c_t^\epsilon = c_t^* + \epsilon C_t^f + o(\epsilon).$$

The process C_t^f represents how optimal policy process changes locally in response to the introduction of a small effect of policy f. One can think of C_t^f as a notion of stochastic elasticity that represents the change in the whole process when policy has additional effects f.

Differentiating the first order condition (17) with respect to ϵ and evaluating at $\epsilon = 0$, we get

$$C_t^f = f'(c_t^*).$$

Without any restrictions on f, the process C_t^f may be quite arbitrary, for example, discontinuous. However, since we assumed smoothness of f, Itô's lemma implies that the process C_t^f is a semimartingale. Moreover, its decomposition is given as (using also that $c_t^* = w_t$) in closed form:

$$dC_t^f = \alpha(w_t)dt + \beta(w_t)dw_t,$$

where

$$\alpha(w_t) = \frac{1}{2} f'''(w_t),$$
$$\beta(w_t) = f''(w_t). \tag{18}$$

9.2.2 Dynamics of optimal policy

We now turn to characterizing the optimal policy c_t^ϵ rather than the first variation process. Let the process c_t^ϵ have the form

$$dc_t^\epsilon = \alpha^\epsilon\left(c_t^\epsilon\right) dt + \beta^\epsilon\left(c_t^\epsilon\right) dw_t.$$

Applying Itô formula to (17), we can find the dynamics of the process c_t^ϵ:

$$dc_t^\epsilon = dw_t - \epsilon \left(f''(c_t^\epsilon) \, dc_t^\epsilon + \frac{1}{2} f'''(c_t^\epsilon) (\beta_t^\epsilon)^2 \, dt \right),$$

$$\alpha_t^\epsilon dt + \beta_t^\epsilon dw_t = dw_t - \epsilon \left(f''(c_t^\epsilon) \, \alpha_t^\epsilon + \frac{1}{2} f'''(c_t^\epsilon) (\beta_t^\epsilon)^2 \right) dt - \epsilon f''(c_t^\epsilon) \beta_t^\epsilon dw_t.$$

The drift and diffusion coefficients are then given by collecting terms at dt and dw_t:

$$\alpha_t^\epsilon (c_t^\epsilon) = -\epsilon \frac{\frac{1}{2} f'''(c_t^\epsilon) (\beta_t^\epsilon)^2}{1 + \epsilon f''(c_t^\epsilon)},$$
$$\beta_t^\epsilon (c_t^\epsilon) = \frac{1}{1 + \epsilon f''(c_t^\epsilon)}. \tag{19}$$

The coefficients in (19) are given as a system of coupled equations and themselves depend on the process c_t^ϵ. This is in contrast to the coefficients of the stochastic elasticity given in (18) which are given in closed form as they are evaluated at the process $c_t^* = c_t^{\epsilon=0} = w_t$.

The key to tractability in the simple setting of this section is Itô lemma. However, Itô lemma only applies to functions and not the functionals. That is, it does not apply to the path-dependent effects of policies studying which is the main goal of the rest of this paper.

It is useful to summarize these results in terms of how the magnitude of the additional effects influences policy. This parallels the discussion of Dixit (1991), Reis (2006) and Alvarez, Lippi, and Paciello (2011, 2016) who show that in the environments with uncertainty and adjustment frictions the costs up to the fourth order may have first order effects on the dynamics of optimal policies. In our setting, since the additional effects f are smooth, we show that the effects up to the third order, that is the second and the third derivative of f will have the first-order effects. The first and second derivative of the marginal effects (that is, the second and third derivative of f) matter for the drift of the optimal policy – this is the evolution of optimal policy with respect to time. The first derivative of the marginal effects matters for the diffusion coefficients – this is response of optimal policy to stochastic shocks. Importantly, the dynamics of the optimal policy is not directly influenced by either the past or the future evolution of the policies.

9.3 Vertical derivatives and the class $\mathcal{A}_t$ of functionals

The functional $h : C[0, t] \to \mathbb{R}$ is said to be in the class $\mathcal{A}_t$ if for each path $c \in C[0, t]$ there exists a number $\partial_{c_t} h(c)$ and an integrable function $(\delta_s h(c))_{s \in [0,t]}$, such that the following asymptotic relation

holds for any $c \in C[0, t]$:

$$h(c + z) = h(c) + \partial_{c_t} h(c) \cdot z_t + \int_0^t \delta_s h(c) \cdot z_s ds + o(||z||), \ z \to 0.$$

The number $\partial_{c_t} h(c)$ is the derivative of h along the value of the path $(c_s)_{s \in [0,t]}$ at time t (i.e. the derivative along the present value). The function $(\delta_s h(c))_{s \in [0,t]}$, represents the integrated influence of the past on the variation of the functional.

Lemma 5. *Assume that the functional $h \in \mathcal{A}_t$ admits a continuous extension to the space $D[0, t]$ of cadlag paths (equipped with the Skorokhod topology). Then the functional h is vertically differentiable at every $c \in C[0, t]$, and the vertical derivative coincides with $\partial_{c_t} h(c)$.*

Proof. Fix $\alpha > 0$. There exists $\delta > 0$ such that for any $z \in C[0, t]$ with $||z|| \leq \delta$,

$$\left| h(c + z) - h(c) - \partial_{c_t} h(c) \cdot z_t - \int_0^t \delta_s h(c) \cdot z_s ds \right| \leq \alpha ||z||.$$

The function $e_t \in D[0, t]$ can be approximated in the Skorokhod topology by continuous functions $z^{(n)}$, such that $0 \leq z^{(n)} \leq 1$, $z_s^{(n)} = 0$ for $s \leq t - \frac{1}{n}$, $z_t^{(n)} = 1$. We have for all $\epsilon \leq \delta$

$$\left| h(c + \epsilon z^{(n)}) - h(c) - \partial_{c_t} h(c) \cdot \epsilon - \epsilon \int_{t-\frac{1}{n}}^t \delta_s h(c) \cdot z_s^{(n)} ds \right| \leq \alpha \epsilon ||z^{(n)}|| = \alpha \epsilon.$$

Dividing by ϵ we get

$$\left| \frac{h(c + \epsilon z^{(n)}) - h(c)}{\epsilon} - \partial_{c_t} h(c) \right| \leq \alpha + \int_{t-\frac{1}{n}}^t |\delta_s h(c)| ds$$

Taking $n \to \infty$,

$$\left| \frac{h(c + \epsilon e_t) - h(c)}{\epsilon} - \partial_{c_t} h(c) \right| \leq \alpha.$$

Since $\alpha > 0$ is arbitrary, this proves that the vertical derivative of h exists and is equal to $\partial_{c_t} h(c)$. $\square$

9.4   Proof of Claim 2

We recall that the class $\mathcal{A}_T$ consists of functionals $g : C[0, T] \to \mathbb{R}$ such that for all $c, z \in C[0, T]$

$$g(c + \epsilon z) = g(c) + \epsilon \partial_{c_t} g(c) z_t + \epsilon \int_0^T \delta_s g(c) z_s ds + o(\epsilon), \ \epsilon \to 0.$$

For any function $c \in C[0,T]$ introduce the transformation

$$(R^{(n)}c)_t = e^{-n(T-t)}c_T + n \int_t^T e^{-n(s-t)}c_s ds.$$

The function $R^{(n)}c$ solves the problem

$$\begin{cases} \frac{dR^{(n)}c}{dt} = -n(c_t - (R^{(n)}c)_t) \\ (R^{(n)}c)_T = c_T \end{cases}$$

Hence, $R^{(n)}c \to c$ uniformly on $[0,T]$. Given any Frechet differentiable functional $g : C[0,T] \to \mathbb{R}$, consider

$$g^{(n)}(c) = g(R^{(n)}c).$$

Fix $c \in C[0,T]$ and let the measure μ be the Frechet derivative of g at $R^{(n)}c$.

$$g^{(n)}(c + \epsilon z) = g(R^{(n)}c + \epsilon R^{(n)}z) = g(R^{(n)})(c) + \epsilon \int_0^T (R^{(n)}z)_s \mu(ds) + o(\epsilon) =$$

$$= g^{(n)}(c) + \epsilon \int_0^T \left(e^{-n(T-s)}z_T + n \int_s^T e^{-n(r-s)}z_r dr \right) \mu(ds) + o(\epsilon) =$$

$$= g^{(n)}(c) + \epsilon \left(\int_0^T e^{-n(T-s)}\mu(ds) \right) z_T + \epsilon \int_0^T \left(\int_0^r e^{-n(r-s)}\mu(ds) \right) z_r dr + o(\epsilon)$$

So, $g^{(n)} \in \mathcal{A}_T$. Every differentiable functional is a pointwise limit of functionals from the class $\mathcal{A}_T$.

9.5   Total derivative formula: a general version of Proposition 1

Let $\zeta = (\zeta_t)_{t \in [0,T]}$ be a measurable process. We will say that ζ is square integrable, if $\mathbb{E} \int_0^T \zeta_s^2 ds < \infty$.

We say that a measurable stochastic process $\xi = (\xi_t)_{t \in [0,T]}$ is absolutely continuous, if ξ_0 is square integrable and there exists a square integrable process $(\zeta_t)_{t \in [0,T]}$, such that

$$\xi_t = \xi_0 + \int_0^t \zeta_s ds, \ 0 \le t \le T. \tag{20}$$

Observe that an absolutely continuous process ξ satisfies $\sup_{0 \le t \le T} \mathbb{E}\xi_t^2 < \infty$ and is a.s. continuous. It is important that an absolutely continuous process need not be adapted. Since the processes under consideration are square integrable and measurable with respect to a Wiener process w, their

values can be represented as stochastic integrals. Namely,

$$\xi_t = \mathbb{E}\xi_t + \int_0^T g_{t,s} dw_s, \ \zeta_t = \mathbb{E}\zeta_t + \int_0^T h_{t,s} dw_s, \ 0 \le t \le T, \tag{21}$$

where $g_t = (g_{t,s})_{s \in [0,T]}$ and $h_t = (h_{t,s})_{s \in [0,T]}$ are progressively measurable (in s) square integrable processes. We need one technical result on the regularity of the process g_t.

Lemma. *For each $t \in [0,T]$ and almost all $s \in [0,T]$*

$$g_{t,s} = g_{0,s} + \int_0^t h_{r,s} dr \ \ a.s.$$

In particular, for almost all $s \in [0,T]$ the process $t \to g_{t,s}$ has a continuous modification.

Proof. Let $u = (u_s)_{s \in [0,T]}$ be an arbitrary progressively measurable square integrable process. Using Itô's isometry we compute

$$\mathbb{E} \int_0^T g_{t,s} u_s ds = \mathbb{E} \int_0^T g_{t,s} dw_s \int_0^T u_s dw_s = \mathbb{E}\xi_t \int_0^T u_s dw_s =$$

$$= \mathbb{E}\xi_0 \int_0^T u_s dw_s + \int_0^t \left(\mathbb{E}\zeta_r \int_0^T u_s dw_s \right) dr =$$

$$= \mathbb{E} \int_0^T g_{0,s} u_s ds + \int_0^t \left(\mathbb{E} \int_0^T h_{r,s} u_s ds \right) dr = \mathbb{E} \int_0^T \left(g_{0,s} + \int_0^t h_{r,s} dr \right) u_s ds.$$

Since the latter holds for arbitrary u we deduce that $g_t = g_0 + \int_0^t h_r dr$ as elements of $L^2(\Omega \times [0,T])$. $\qquad \square$

Further we will always deal with a continuous in t modifications of processes $t \to g_{t,s}$.

Proposition. *Let $\eta = (\eta_t)_{t \in [0,T]}$ be a square integrable progressively measurable process. Then η is an Itô process if and only if it can be represented in the form $\eta_t = \mathbb{E}[\xi_t | \mathcal{F}_t]$ for some square integrable absolutely continuous process $\xi = (\xi_t)_{t \in [0,T]}$. In this case the semimartingale representation of η is given by*

$$d\eta_t = \mathbb{E}[\zeta_t | \mathcal{F}_t] dt + g_{t,t} dw_t,$$

where processes ζ and g_t are determined from (20), (21).

Proof. Assume that η is an Itô process. Then it can be written in the form

$$\eta_t = \eta_0 + \int_0^t \alpha_s ds + \int_0^t \beta_s dw_s, \ 0 \le t \le T.$$

55

Introduce the process

$$\xi_t = \eta_0 + \int_0^T \beta_s dw_s + \int_0^t \alpha_s ds.$$

Observe that the process ξ is an absolutely continuous process. Since stochastic integrals are martingales, we deduce that

$$\mathbb{E}\left[\xi_t | \mathcal{F}_t\right] = \eta_0 + \int_0^t \alpha_s ds + \mathbb{E}\left[\int_0^T \beta_s dw_s | \mathcal{F}_t\right] = \eta_0 + \int_0^t \alpha_s ds + \int_0^t \beta_s dw_s = \eta_t.$$

Conversely, assume that $\eta_t = \mathbb{E}[\xi_t | \mathcal{F}_t]$ with some absolutely continuous process ξ. Related processes ζ, g_t are defined in (20), (21). Consider the process

$$M_t = \eta_t - \eta_0 - \int_0^t \mathbb{E}[\zeta_r | \mathcal{F}_r] dr, \ 0 \le t \le T.$$

We verify that the process M is a martingale. Indeed, for $s < t$ we compute

$$\mathbb{E}[M_t | \mathcal{F}_s] = \mathbb{E}\left[\eta_t - \eta_0 - \int_0^t \mathbb{E}[\zeta_r | \mathcal{F}_r] dr | \mathcal{F}_s\right] =$$

$$= \mathbb{E}\left[\mathbb{E}[\xi_t | \mathcal{F}_t] - \eta_0 - \int_0^t \mathbb{E}[\zeta_r | \mathcal{F}_r] dr | \mathcal{F}_s\right] =$$

$$\mathbb{E}[\xi_t | \mathcal{F}_s] - \eta_0 - \int_s^t \mathbb{E}[\zeta_r | \mathcal{F}_s] dr - \int_0^s \mathbb{E}[\zeta_r | \mathcal{F}_r] dr =$$

$$= \mathbb{E}\left[\xi_s + \int_s^t \zeta_r dr | \mathcal{F}_s\right] - \eta_0 - \int_s^t \mathbb{E}[\zeta_r | \mathcal{F}_s] dr - \int_0^s \mathbb{E}[\zeta_r | \mathcal{F}_r] dr =$$

$$= \mathbb{E}[\xi_s | \mathcal{F}_s] - \eta_0 - \int_0^s \mathbb{E}[\zeta_r | \mathcal{F}_r] dr = \eta_s - \eta_0 - \int_0^s \mathbb{E}[\zeta_r | \mathcal{F}_r] dr = M_s.$$

As a Wiener martingale, the process M has a representation as a stochastic integral. Since $M_0 = 0$ the representation is of the form:

$$M_t = \int_0^t v_s dw_s$$

for some progressively measurable square integrable process $v = (v_s)_{s \in [0,T]}$. To find this process we use Itô's isometry. For arbitrary progressively measurable square integrable process $u = (u_s)_{s \in [0,T]}$,

we have

$$\mathbb{E} \int_0^T v_s u_s ds = \mathbb{E} \int_0^T v_s dw_s \int_0^T u_s dw_s = \mathbb{E} M_T \int_0^T u_s dw_s =$$

$$= \mathbb{E} \left(\xi_T - \eta_0 - \int_0^T \mathbb{E}[\zeta_r | \mathcal{F}_r] dr \right) \int_0^T u_s dw_s = \mathbb{E} \left(\xi_T \int_0^T u_s dw_s \right) - \int_0^T \mathbb{E} \left(\zeta_r \int_0^r u_s dw_s \right) dr =$$

$$= \mathbb{E} \int_0^T g_{T,s} u_s ds - \int_0^T \mathbb{E} \int_0^r h_{r,s} u_s ds dr = \mathbb{E} \int_0^T \left(g_{T,s} - \int_s^T h_{r,s} dr \right) u_s ds.$$

From the previous Lemma we deduce

$$v_s = g_{T,s} - \int_s^T h_{r,s} dr = g_{0,s} + \int_0^s h_{r,s} dr = g_{s,s}.$$

Finally, equality

$$\eta_t = \eta_0 + \int_0^t \mathbb{E}[\zeta_s | \mathcal{F}_s] ds + M_t = \eta_0 + \int_0^t \mathbb{E}[\zeta_s | \mathcal{F}_s] ds + \int_0^t g_{s,s} dw_s$$

implies the needed semimartingale representation of η. $\qquad\square$

Remark. For almost all $t \in [0, T]$ the process ζ_t is the time-derivative of the process ξ_t. If we assume that ξ_t is Malliavin differentiable, then the Clark-Ocone formula implies

$$g_{t,t} = \mathbb{E} \left[D_t \xi_t | \mathcal{F}_t \right].$$

In these terms the result of the theorem can be written as

$$d\mathbb{E} \left[\xi_t | \mathcal{F}_t \right] = \mathbb{E} \left[\partial_t \xi_t | \mathcal{F}_t \right] dt + \mathbb{E} \left[D_t \xi_t | \mathcal{F}_t \right] dw_t,$$

which is the motivation behind the term "total derivative formula".

9.6   Characterization of Itô processes: converse of Proposition 1

The total derivative formula states that for any square integrable absolutely continuous process $(\eta_t)_{t \in [0,T]}$, the process $\xi_t = E[\eta_t | \mathcal{F}_t]$ is an Itô process. In this section we give the converse statement.

Claim. For any square integrable Itô process $(\xi_t)_{t \in [0,T]}$ there exists an absolutely continuous process $(\eta_t)_{t \in [0,T]}$ such that

$$\xi_t = E[\eta_t | \mathcal{F}_t].$$

Proof. By the definition of the Itô process, there exist adapted square integrable processes $(\alpha_t)_{t \in [0,T]}$ and $(\beta_t)_{t \in [0,T]}$ such that

$$\xi_t = \xi_0 + \int_0^t \alpha_s ds + \int_0^t \beta_s dw_s.$$

We simply define

$$\eta_t = \xi_0 + \int_0^t \alpha_s ds + \int_0^T \beta_s dw_s.$$

The process η is absolutely continuous, and

$$E[\eta_t | \mathcal{F}_t] = \xi_0 + \int_0^t \alpha_s ds + E\left[\int_0^T \beta_s dw_s \bigg| \mathcal{F}_t\right] =$$

$$= \xi_0 + \int_0^t \alpha_s ds + \int_0^t \beta_s dw_s = \xi_t.$$

$\square$

References

Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hemous. "The environment and directed technical change." American economic review 102, no. 1 (2012): 131-66.

Acemoglu, Daron, Philippe Aghion, Lint Barrage and David Hemous "Climate change, directed innovation, and energy transition: The long-run consequences of the shale gas revolution." In 2019 Meeting Papers, no. 1302. Society for Economic Dynamics, 2019.

Acemoglu, Daron, Ufuk Akcigit, Douglas Hanley, and William Kerr. "Transition to clean technology." Journal of Political Economy 124, no. 1 (2016): 52-104.

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. Institutional Change and Institutional Persistence. No. w27852. National Bureau of Economic Research, 2020.

Aghion, Philippe, Antoine Dechezleprêtre, David Hemous, Ralf Martin, and John Van Reenen. "Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry." Journal of Political Economy 124, no. 1 (2016): 1-51.

Aghion, Philippe, Cameron Hepburn, Alexander Teytelboym, and Dimitri Zenghelis. "Path dependence, innovation and the economics of climate change." In Handbook on Green Growth. Edward Elgar Publishing, 2019.

Akerlof, George A., and Janet L. Yellen. "A near-rational model of the business cycle, with wage and price inertia." The Quarterly Journal of Economics 100, no. Supplement (1985): 823-838.

Alvarez, Fernando, and Francesco Lippi. "The Analytic Theory of a Monetary Shock." Einaudi Institute for Economics and Finance Mimeo, in preparation (2019).

Alvarez, Fernando E., Francesco Lippi, and Luigi Paciello. "Optimal price setting with observation and menu costs." The Quarterly journal of economics 126, no. 4 (2011): 1909-1960.

Alvarez, Fernando E., Francesco Lippi, and Luigi Paciello. "Monetary shocks in models with inattentive producers." The Review of economic studies 83, no. 2 (2016): 421-459.

Alvarez, Fernando, Herve Le Bihan, and Francesco Lippi. "The real effects of monetary shocks in sticky price models: a sufficient statistic approach." American Economic Review 106, no. 10 (2016): 2817-51.

Alvarez, Fernando E., Francesco Lippi, and Aleksei Oskolkov. The Macroeconomics of Sticky Prices with Generalized Hazard Functions. No. w27434. National Bureau of Economic Research, 2020.

Arrow, Kenneth J. "Increasing returns: historiographic issues and path dependence." The European Journal of the History of Economic Thought 7, no. 2 (2000): 171-180.

Bakkensen, Laura A., and Lint Barrage. Flood risk belief heterogeneity and coastal home price dynamics: Going under water?. No. w23854. National Bureau of Economic Research, 2017.

Baldwin, Elizabeth, Yongyang Cai, and Karlygash Kuralbayeva. "To build or not to build? Capital stocks and climate policy∗." Journal of Environmental Economics and Management 100 (2020): 102235.

Barnett, Michael, William Brock, and Lars Peter Hansen. "Pricing uncertainty induced by climate change." The Review of Financial Studies 33, no. 3 (2020): 1024-1066.

Barrage, Lint, and Jacob Furst. "Housing investment, sea level rise, and climate change beliefs." Economics letters 177 (2019): 105-108.

Bergemann, Dirk, and Philipp Strack. "Dynamic revenue maximization: A continuous time approach." Journal of Economic Theory 159 (2015): 819-853.

Berger, David W., Konstantin Milbradt, Fabrice Tourre, and Joseph Vavra. Mortgage prepayment and path-dependent effects of monetary policy. No. w25157. National Bureau of Economic Research, 2018.

Blanchard, Olivier J., and Lawrence H. Summers. "Hysteresis and the European unemployment problem." NBER macroeconomics annual 1 (1986): 15-78.Bhamra, Harjoat S., and Raman Uppal. "The effect of introducing a non-redundant derivative on the volatility of stock-market returns when agents differ in risk aversion." The Review of Financial Studies 22, no. 6 (2009): 2303-2330.

Borovička, J. and Hansen, L.P., 2016. Term structure of uncertainty in the macroeconomy. In Handbook of Macroeconomics (Vol. 2, pp. 1641-1696). Elsevier.

Borovička, Jaroslav, Lars Peter Hansen, and José A. Scheinkman. "Shock elasticities and impulse responses." Mathematics and Financial Economics 8, no. 4 (2014): 333-354.

Borovička, Jaroslav, Lars Peter Hansen, Mark Hendricks, and José A. Scheinkman. "Risk-price dynamics." Journal of Financial Econometrics 9, no. 1 (2011): 3-65.

Boucher, O., P. R. Halloran, E. J. Burke, M. Doutriaux-Boucher, C. D. Jones, J. Lowe, M. A. Ringer, E. Robertson, and P. Wu. "Reversibility in an Earth System model in response to CO2 concentration changes." Environmental Research Letters 7, no. 2 (2012): 024013.

Boulatov, Alexei, Georgii Riabov, and Aleh Tsyvinski. "Optimal trading with dynamic stochastic price impact", Working paper (2020).

Brian, Arthur, W. "Competing technologies, increasing returns, and lock-in by historical events."

The economic journal 99, no. 394 (1989): 116-131.

Brock, William A., and Lars Peter Hansen. "Wrestling with uncertainty in climate economic models." University of Chicago, Becker Friedman Institute for Economics Working Paper 2019-71 (2018).

Cai, Yongyang, and Thomas S. Lontzek. "The social cost of carbon with economic and climate risks." Journal of Political Economy 127, no. 6 (2019): 2684-2734.

Collins, Matthew, Reto Knutti, Julie Arblaster, Jean-Louis Dufresne, Thierry Fichefet, Pierre Friedlingstein, Xuejie Gao et al. "Long-term climate change: projections, commitments and irreversibility." In Climate Change 2013-The Physical Science Basis: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 1029-1136. Cambridge University Press, 2013.

Cont, Rama, and David-Antoine Fournie. "Functional Itô calculus and stochastic integral representation of martingales." The Annals of Probability 41, no. 1 (2013): 109-133.

Cosso, Andrea, Salvatore Federico, Fausto Gozzi, Mauro Rosestolato, and Nizar Touzi. "Path-dependent equations and viscosity solutions in infinite dimension." The Annals of Probability 46, no. 1 (2018): 126-174.

Cvitanić, Jaksa, and Semyon Malamud. "Equilibrium driven by discounted dividend volatility." Swiss Finance Institute Research Paper 09-34 (2009)

Cvitanić, Jakša, Dylan Possamaï, and Nizar Touzi. "Moral hazard in dynamic risk management." Management Science 63, no. 10 (2017): 3328-3346.

David, Paul A. "Clio and the Economics of QWERTY." The American Economic Review 75, no. 2 (1985): 332-337.

Detemple, Jerome B., and Fernando Zapatero. "Asset prices in an exchange economy with habit formation." Econometrica: Journal of the Econometric Society (1991): 1633-1657.

Dietz, Simon, James Rising, Thomas Stoerk, and Gernot Wagner. "Tipping Points in the Climate System and the Economics of Climate Change." In EGU General Assembly Conference Abstracts, p. 2959. 2020.

Dixit, Avinash. "Analytical approximations in models of hysteresis." The Review of Economic Studies 58, no. 1 (1991): 141-151.

Dixit, Avinash. "Investment and hysteresis." Journal of economic perspectives 6, no. 1 (1992): 107-132.

Dupire, Bruno. "Functional Itô calculus." 2009. Bloomberg Working Paper.

Dupire, Bruno. "Functional Itô calculus." Quantitative Finance 19, no. 5 (2019): 721-729.

Eggertsson, Gauti B., Neil R. Mehrotra, and Jacob A. Robbins. "A model of secular stagnation: Theory and quantitative evaluation." American Economic Journal: Macroeconomics 11, no. 1 (2019): 1-48.

Egorov, Georgy, and Konstantin Sonin. The Political Economics of Non-democracy. No. w27949. National Bureau of Economic Research, 2020.

Eliseev, Alexey V., Pavel F. Demchenko, Maxim M. Arzhanov, and Igor I. Mokhov. "Transient hysteresis of near-surface permafrost response to external forcing." Climate dynamics 42, no. 5-6 (2014): 1203-1215.

Fabbri, Giorgio, Fausto Gozzi, and Andrzej Swiech. "Stochastic optimal control in infinite dimension." Probability and Stochastic Modelling. Springer (2017).

Fouquet, Roger. "Path dependence in energy systems and economic development." Nature Energy 1, no. 8 (2016): 1-5.

Fournié, Eric, Jean-Michel Lasry, Jérôme Lebuchoux, Pierre-Louis Lions, and Nizar Touzi. "Applications of Malliavin calculus to Monte Carlo methods in finance." Finance and Stochastics 3, no. 4 (1999): 391-412.

Galí, Jordi. Insider-outsider labor markets, hysteresis and monetary policy. No. w27385. National Bureau of Economic Research, 2020.

Garbe, Julius, Torsten Albrecht, Anders Levermann, Jonathan F. Donges, and Ricarda Winkelmann. "The hysteresis of the Antarctic ice sheet." Nature 585, no. 7826 (2020): 538-544.

Gasser, Thomas, Mehdi Kechiar, Phillippe Ciais, E. J. Burke, Thomas Kleinen, Dan Zhu, Yuanyuan Huang, Altug Ekici, and M. Obersteiner. "Path-dependent reductions in CO_2 emission budgets caused by permafrost carbon release." Nature Geoscience 11, no. 11 (2018): 830-835.

Giglio, Stefano, Bryan T. Kelly, and Johannes Stroebel. Climate Finance. No. w28226. National Bureau of Economic Research, 2020.

Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski. "Optimal taxes on fossil fuel in general equilibrium." Econometrica 82, no. 1 (2014): 41-88.

Grubb, Michael, Rutger-Jan Lange, Nicolas Cerkez, Pablo Salas, Jean-Francois Mercure, and Ida Sognnaes. "Taking Time Seriously: Implications for Optimal Climate Policy." (2020).

Hassler, John, Per Krusell, Conny Olovsson, and Michael Reiter. "On the effectiveness of climate policies." IIES WP (2020).

Huang, Zangbo and Michael Sockin "Dynamic Taxation with Endogenous Skill Premia", 2018.

Hong, Harrison, Neng Wang, and Jinqiang Yang. Mitigating Disaster Risks in the Age of Climate Change. National Bureau of Economic Research, 2020.

Jordà, Òscar, Sanjay R. Singh, and Alan M. Taylor. The long-run effects of monetary policy. No. w26666. National Bureau of Economic Research, 2020.

Kotlikoff, L.J., Kubler, F., Polbin, A. and Scheidegger, S., 2020. Pareto-Improving Carbon-Risk Taxation (No. w26919). National Bureau of Economic Research.

Krasnosel'skii, Mark A., and Aleksei V. Pokrovskii. Systems with hysteresis. Springer Science & Business Media, 2012.

Lemoine, Derek. "The climate risk premium: how uncertainty affects the social cost of carbon." Journal of the Association of Environmental and Resource Economists 8, no. 1 (2021): 27-57.

Lemoine, Derek, and Christian Traeger. "Watch your step: optimal policy in a tipping climate." American Economic Journal: Economic Policy 6, no. 1 (2014): 137-66.

Li, Xin, Borghan Narajabad, and Ted Temzelides. "Robust dynamic energy use and climate change." Quantitative Economics 7, no. 3 (2016): 821-857.

Ljungqvist, Lars, and Thomas J. Sargent. Recursive macroeconomic theory. MIT press, 2018.

Lontzek, Thomas S., Yongyang Cai, Kenneth L. Judd, and Timothy M. Lenton. "Stochastic integrated assessment of climate tipping points indicates the need for strict climate policy." Nature Climate Change 5, no. 5 (2015): 441-444.

Mankiw, N. Gregory. "Small menu costs and large business cycles: A macroeconomic model of monopoly." The Quarterly Journal of Economics 100, no. 2 (1985): 529-537.

Makris, Miltiadis, and Alessandro Pavan. Taxation under learning-by-doing. Working Paper, 2017.

Makris, Miltiadis, and Alessandro Pavan. Wedge Dynamics with Endogenous Private Information: A General Recursive Characterization. mimeo, Northwestern University, 2018.

Meng, Kyle C. Estimating path dependence in energy transitions. No. w22536. National Bureau of Economic Research, 2016.

Nohara, Daisuke, Yoshikatsu Yoshida, Kazuhiro Misumi, and Masamichi Ohba. "Dependency of climate change and carbon cycle on CO2 emission pathways." Environmental Research Letters 8, no. 1 (2013): 014047.

Nordhaus, W.D., 1994. Managing the global commons: the economics of climate change (Vol. 31). Cambridge, MA: MIT press.

Page, Scott E. "Path dependence." Quarterly Journal of Political Science 1, no. 1 (2006):

87-115.

Pavan, Alessandro. "Dynamic mechanism design: Robustness and endogenous types." In Advances in Economics and Econometrics: Eleventh World Congress, pp. 1-62. 2017.

Pavan, Alessandro, Ilya Segal, and Juuso Toikka. "Dynamic mechanism design: A myersonian approach." Econometrica 82, no. 2 (2014): 601-653.

Reis, Ricardo. "Inattentive producers." The Review of Economic Studies 73, no. 3 (2006): 793-821.

Prescott, Edward C. "Should control theory be used for economic stabilization?." In Carnegie-Rochester Conference Series on Public Policy, vol. 7, pp. 13-38. North-Holland, 1977.

Rogelj, J., D. Shindell, K. Jiang, S. Fifita, P. Forster, V. Ginzburg, C. Handa, H. Kheshgi, S. Kobayashi, E. Kriegler, L. Mundaca, R. Séférian, and M.V.Vilariño, 2018: Mitigation Pathways Compatible with 1.5°C in the Context of Sustainable Development. In: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty [Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (eds.)]. In Press.

Sannikov, Yuliy. "Moral hazard and long-run incentives." Unpublished working paper, Princeton University (2014).

Serrat, Angel. "A dynamic equilibrium model of international portfolio holdings." Econometrica 69, no. 6 (2001): 1467-1489.

Summers, Lawrence H. "US economic prospects: Secular stagnation, hysteresis, and the zero lower bound." Business economics 49, no. 2 (2014): 65-73.

Temzelides, Ted. "Needed: Robustness in Climate Economics." Coping with the climate crisis (2016): 104.

Traeger, Christian P. "Ace–analytic climate economy (with temperature and uncertainty)." (2018).

Van Den Bremer, Ton, and Rick van der Ploeg. "Pricing Carbon Under Economic and Climactic Risks: Leading-Order Results from Asymptotic Analysis." (2018).

van der Ploeg, Frederick, and Aart de Zeeuw. "Climate tipping and economic growth: precautionary capital and the price of carbon." Journal of the European Economic Association 16, no. 5

(2018): 1577-1617.

Walsh, Conor. "Hysteresis Implies Scale Effects." (2020): Working paper.

Zickfeld, K., Arora, V.K. and Gillett, N.P., 2012. Is the climate response to CO2 emissions path-dependent?. Geophysical Research Letters, 39(5).