Responsive Market Regulation in Environmental Economics: An Agent-Based Stochastic Equilibrium Approach

Dissertation
zur Erlangung des Doktorgrades Dr. rer. nat.

der Fakultät für Mathematik
der Universität Duisburg-Essen

vorgelegt von
Dipl.-Math. Sascha Kollenberg
aus Aachen

Essen
März 2018

Erstgutachter: Prof. Dr. Rüdiger Kiesel, Universität Duisburg-Essen
Zweitgutachter: Prof. Dr. Ralf Werner, Universität Augsburg

Datum der Disputation: 25.07.2018

Ich versichere an Eides statt durch meine Unterschrift, dass ich die vorstehende Arbeit selbstständig und ohne fremde Hilfe angefertigt und alle Stellen, die ich wörtlich oder inhaltlich aus Veröffentlichungen entnommen habe, als solche kenntlich gemacht habe, mich auch keiner anderen als der angegebenen Literatur und keinem anderen als der genannten Hilfsmittel bedient habe. Die Arbeit hat in dieser oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen. Es hat kein vorangegangenes Promotionsverfahren gegeben.

Sascha Kollenberg
März 2018

German Summary / Deutsche Zusammenfassung

In der vorliegenden Arbeit beschäftigen wir[1] uns mit der stochastischen Modellierung von Märkten für regulierte Ressourcen. Dabei verwenden wir den Begriff der *regulierten Ressource* (bzw. der *regulierten Commodity*) im Sinne eines Gutes dessen Gesamtverbrauch in einer vorher festgelegten Periode durch eine staatliche, zwischenstaatliche oder private, zentrale Entität gesteuert wird. Zudem soll ein etwaiger Zufluss der Ressource zum Markt durch ebenjene Entität einer (dynamischen) Regulierung unterworfen werden können. Indem der Zufluss der Ressource zum Markt derart reguliert wird, selbige jedoch anschließend frei durch die Marktteilnehmer gehandelt werden kann, können etwaige Externalitäten aus dem Verbrauch der Ressource eingepreist werden, ohne dass der (volks-) wirtschaftliche Schaden aus dem Verbrauch der Ressource selbst quantifiziert werden muss.

Ein solches Handelssystem stellt klassischerweise eine Alternative zur Besteuerung als Mittel der Einschränkung des (verbrauchenden) Einsatzes einer Ressource dar. Die Frage, ob ein wie oben beschriebenes *mengenbasiertes* Instrument oder ein *preisbasiertes* Instrument (wie bspw. eine Steuer) im Einzelfall vorzuziehen ist, wurde bereits von [Weitzman, 1974] behandelt, dessen Betrachtungen auf den Vergleich der Steigung marginaler Nutzen- und Kostenkurven bzgl. der Ressourceneinsparung hinausliefen. Weitzman bemerkt jedoch ebenfalls, dass ein ideales Instrument ein dynamisches Signal auf Basis des aktuellen Zustandes der Volkswirtschaft senden würde, für die oder in der die Regulierung gilt. In der vorliegenden Arbeit greifen wir diesen Gedanken Weitzmans auf und untersuchen, wie die Kosteneffizienz einer Regulierung durch ein dynamisches Signal gesteigert werden kann. Zudem stellen wir dar, wie dynamische Regulierungen zur Abbildung eines ganzen Spektrums an Instrumenten verwendet werden können, welche zwischen den Extremen eines reinen mengen- und eines reinen preisbasierten Instruments liegen. Die vorliegende Arbeit füllt damit eine nicht unwesentliche konzeptionelle Lücke in der ökonomischen Literatur.

Eine Einschränkung oder Regulierung der Förderung bzw. des Einsatzes einer Ressource kann im kollektiven Interesse einer Gruppe von Wirtschaftssubjekten sein, oder aber im Interesse der Gesellschaft, welche den Folgen des Handelns jener Wirtschaftssubjekte ausgeliefert ist. Während sich Beispiele für Ersteres vornehmlich im Bereich des Kartellwesens bewegen (bspw. Förderung von begrenzten natürlichen Ressourcen wie fossiler Brennstoffe), kommt Letzteres zum Tragen, wenn schwerwiegende gesellschaftliche Folgen zu erwarten sind, sollte die Verwendung einer Ressource gewisse Grenzen überschreiten. Beispiele für die Lösung einer solchen Problematik finden sich in der Fischerei in Form sogenannter Individual Transferrable Quotas[2] und in der Forstwirtschaft[3], wesentlich prominentererweise jedoch auch im Handel von Emissionsrechten, welchen wir als wesentliches Anwendungsbeispiel kurz vorstellen wollen.

Unser wesentliches Anwendungsbeispiel für oben beschriebene regulierte Ressourcen sind klimarelevante Commodities, allen voran CO_2-Emissionsrechte. Solche Emissionsrechte werden in der Europäischen Union seit 2005 unter dem sogenannten European Union Emissions Trading System (EU ETS) gehandelt. Hierbei werden dem Markt in regelmäßigen Auktionen sowie durch freie Allokationen Emissionsrechte in Form sogenannter EUAs (European Emission Allowances) zugeführt, welche anschließend frei und zum durch den Markt (endogen) festgelegten Preis gehandelt werden können. Die unter das EU ETS fallenden Agenten müssen nun für jedes Kalenderjahr eine Menge von EUAs vorweisen können, welche ihrem Treibhausgasaustoß in Tonnen CO_2-Äquivalent (CO_2 equivalent, CO_2e) entspricht. Zur Anreizsetzung fällt für jede nicht auf diese Weise abgedeckte Tonne CO_2e eine fixe Strafzahlung von derzeit 100 EUR an. (Es sei angemerkt, dass jeder unter die Regulierung fallende Marktteilnehmer auch nach Zahlung der Strafgebühr dazu verpflichtet ist, die entsprechende Menge an EUAs in der nächsten Abrechnungsperiode nachzureichen.)

Nach der Wirtschaftskrise 2007/08 trat auf dem europäischen Markt für Emissionsrechte ein Preisverfall ein, dessen Folgen bis heute andauern. Das dem Preisverfall zugrundeliegende Überangebot an EUAs wird landläufig weitestgehend einer Produktionseinbuße der unter die entsprechende Regulierung fallenden Wirtschaftssubjekte zugeschrieben. Als Folge der oben genannten Entwicklungen kamen seitens einiger gesellschaftlicher und politischer Interessengemeinschaften Zweifel an der Zulänglichkeit des Regelwerks auf, welches dem EU ETS bis dato zugrunde lag. Es wurde insbesondere moniert, das derzeitige System wäre nicht in der Lage, auf sich ändernde ökonomische Rahmenbedingungen angemessen zu reagieren.

Aufbauend auf einer auf diese Problematik abzielenden Konsultation von externen Experten und Stakeholdern wurde im November 2012 durch die Europäische Kommission ein Ergebnispapier[4] veröffentlicht, welches eine Reihe möglicher Maßnahmen zur Behebung der oben genannten Unzulänglichkeiten des Systems umriss. Unter den genannten Maßnahmen war die Einführung einer sogenannte Marktstabilitätsreserve (MSR), welche den Zufluss von EUAs zum Markt statt wie bislang statisch, nunmehr dynamisch gestalten sollte: Die Ausschreibungsmengen sollten sich gemäß dieses Vorschlags an der Menge der sich im Markt befindlichen (also insbesondere nicht bereits zur Abdeckung von Emissionen verwendeten) EUAs orientieren. Zudem sollte diese Zuflussregulierung nach festen und transparenten Regeln geschehen, um politische Risiken im Sinne

[1] Ich verwende in der vorliegenden Arbeit stets die Formulierung 'wir' anstelle von 'ich' (bzw. Englisch 'we' anstelle von 'I') etc. im Sinne einer sprachlichen Einbindung des Lesers in den Verlauf und Aufbau der Arbeit. Die Eigenständigkeit meiner zugrundeliegenden Leistung bleibt hiervon unberührt.

[2] Siehe bspw. [Grafton, 1996].

[3] Siehe bspw. [Teeguarden, 1969].

[4] Siehe [The European Commission, 2012].

der Marktteilnehmer einzudämmen. Die MSR wurde im Jahr 2016 ratifiziert und wird nach derzeit geltender Rechtsprechung ab 2020 aktiv.

In der vorliegenden Arbeit konstruieren wir ein zeitstetiges stochastisches Gleichgewichtsmodell eines Marktes für regulierte Ressourcen, mit besonderem Augenmerk auf obigem wesentlichem Anwendungsbeispiel. Wir legen hierbei insbesondere Wert auf eine mathematisch rigorose Abbildung von Risikoaversion auf Seiten der Marktteilnehmer, sowie eine detaillierte Ausarbeitung einer Lösung des Gleichgewichtsproblems in geschlossener Form. Es gelingt uns insbesondere, das für jeden einzelnen Marktteilnehmer auftretende Optimierungsproblem, zusammen mit der damit assoziierten Hamilton-Jacobi-Bellman-Gleichung explizit zu lösen. Anschließend konstruieren wir einen Mechanismus zur dynamischen Kontrolle des Ressourcenzuflusses, welcher, aufbauend auf unserem Gleichgewichtsmodell, eine Lösung des Marktgleichgewichtes unter jener Regulierung zulässt. Es gelingt uns damit auch, die Wechselwirkung zwischen der dynamischen Regulierung (deren Signal auf der Marktsituation beruht) und der Marktdynamik (dessen Entwicklung von der dynamischen Regulierung abhängt) darzustellen und eine entsprechende Gleichgewichtslösung zu finden. Dadurch gelingt es uns wiederum, die Kosteneffizienz des Systems explizit zu quantifizieren womit wir zeigen, wie ein solches dynamisches System hinsichtlich der Kosteneffizienz zu optimieren ist. Wir stellen dabei fest, dass ein dynamisches System durch Reduktion der Kosten, welche seitens der Marktteilnehmer durch Anpassung an stochastische Einflüsse entstehen, einem statischen System überlegen sein kann – und zwar insbesondere auch dann, wenn dem Markt in Erwartung weniger Ressourceneinheiten zur Verfügung gestellt werden als dies ohne eine dynamische Angebotsanpassung der Fall wäre.

Durch die Allgemeinheit unseres Vorgehens gelingt es uns außerdem, das eingangs erwähnte Spektrum zwischen reinen mengen- und reinen preisbasierten Instrumenten zur Regulierung des Ressourceneinsatzes abzubilden: Mithilfe einer einfachen Parametrisierung der Responsivität des Mechanismus offenbart sich eine vollständige abstrakte Darstellung des Kontinuums von Instrumenten. Damit adressieren wir mit unserer Analyse nicht nur die Frage nach einem optimalen dynamischen Allokationsmechanismus, sondern vielmehr auch die Frage nach einer optimalen Politikgestaltung durch gemischte Preis- und Mengensysteme. Damit schließen wir eine wesentliche konzeptionelle Lücke in dem durch [Weitzman, 1974] eingeleiteten Zweig der ökonomischen Literatur.

Unsere Arbeit gliedert sich in vier Kapitel. In Kapitel 1 beschreiben wir unsere Motivation und ordnen unsere Arbeit in den Kontext der ökonomischen Literatur ein. Es folgen in Kapitel 2 einige allgemeine Bemerkungen und Ergebnisse zu unserem stochastischen Modellierungsansatz. In Kapitel 3 erfolgt nun die Konstruktion unseres stochastischen Gleichgewichtsmodells sowie dessen explizite Lösung. Anschließend wird in Kapitel 4 der Mechanismus zur dynamischen Allokation spezifiziert und das resultierende Marktgleichgewicht dargelegt. Wir legen hier außerdem dar, wie die aggregierten Kosten berechnet und minimiert werden können. Wir gehen hier ebenfalls detailliert auf das eingangs genannte Spektrum von Instrumenten ein, und wie es anhand der Parametrisierung unseres dynamischen Mechanismus abgebildet werden kann.

Einige der in den Kapiteln 3 und 4 präsentierten Ergebnisse wurden von mir in kürzerer und vereinfachter Form in [Kollenberg and Taschini, 2016a] sowie [Kollenberg and Taschini, 2016b] veröffentlicht. Dies beinhaltet insbesondere die im letzten Kapitel gezeigten Abbildungen[5] *4.1, 4.2 und 4.3.*

[5] Der den Abbildungen zugrundeliegende Programmcode wurde von mir eigenständig in der zu MathWorks Matlab gehörenden, proprietären Sprache geschrieben. Die genannten Abbildungen habe ich mittels Matlab unter einer gültigen Lizenz selbständig erstellt.

English Summary

The present work is concerned with stochastic modelling of markets for regulated resources. Here, we[6] use the term *regulated resource* (or *regulated commodity*) in the sense of some tradable good, the total aggregate deployment of which is subject to control by some public or private, central entity. More precisely, we are interested in regulatory systems that can control the inflow of commodity units. Such control systems allow the regulator to force potential externalities, arising from using the resource, to be priced in on the market, without explicitly deciding on the respective premium or explicitly quantifying the economic damage avoided for each unit of the commodity.

Such (quantity-based) trading systems are traditionally viewed as alternatives to conventional price-based systems such as taxes. The question whether a quantity-based or price-based system is more advantageous has first been studied by [Weitzman, 1974], whose deliberations came down to comparing the relative slopes of the marginal abatement cost curve and marginal benefit curve (of avoiding emissions). However, Weitzman appreciated that an ideal instrument would implement a contingency message, indicating what policy would be implemented, based on the current state of the system. In the present work, we pick up on this idea and examine how the costs efficiency of controlling the use of a resource can be improved by such dynamic control systems. What's more, we construct a quantitative framework by which we can represent an entire *spectrum* of policies between 'pure' price-based and 'pure' quantity-based instruments. Thusly we fill a significant gap in the pertinent literature.

Limitation or regulation of the inflow of commodity units to some economy can be in the collective interest of a group of economic entities, or in the interest of the society which is subjected to those entities' externalities. While examples for the former are mainly found in the context of cartels, the latter is relevant whenever societal damages are the result of the consumption of a commodity – e.g. through overfishing or non-sustainable forestry. Another, particularly topical example are climate relevant commodities such as fossil fuels or emissions allowances.

The main inspiration for our modelling work are Emissions Trading Systems (ETSs), the currently largest of which is the European Union ETS (EU ETS). Therein, emissions allowances (EUAs – European Emission Allowances) are made available by the regulator through regular auctions and free allocations. Those allowances can then be traded freely among market participants at the price dictated only by supply and demand on this secondary market. After a given calendar year, agents whose facilities are subject to the regulation of the EU ETS have to present a number of EUAs equal to their total emissions of greenhouse gases, measured in (metric) tonnes of CO_2-equivalent (CO_2e), during the respective year. For each tonne not covered in this way, each agent who is short in allowances has to pay a penalty of (currently) 100 Euros. (It shall be mentioned that paying the penalty does not absolve the agent from having to present the respective allowances in the next year.)

In the wake of the economic crisis of 2007/08, we saw a strong and persistent erosion of EUA prices, which is mainly attributed to an oversupply resulting from an unanticipated drop in production. It was this price erosion, along with the ETS's lack of provisions to adapt to such circumstances, that brought about doubt in the efficacy of the instrument altogether. More precisely, a number of experts in academia, politics and some stakeholders believed that the system was unable to 'function in an orderly fashion' [The European Commission, 2012] in the presence of such an enormous oversupply. In particular, the system lacked provisions to adapt the allocation scheme to changes in economic circumstances.

Based on public consultations, the European Commission released a report in November 2012 (see [The European Commission, 2012]), presenting a number of reform options for the EU ETS, which aimed at relieving the current oversupply and, potentially, making the system more responsive to changes in economic circumstances. One such measure is the introduction of the so-called Market Stability Reserve (MSR), which aims at making the supply of allowances dynamic and dependent on the state of the system, rather than static and rigid, as it was. Furthermore, this supply adjustment system is designed based on a transparent set of rules, rather than discretionary interventions. The MSR has been ratified in 2016 and is set to start operating in 2020.

In the present work, we construct a continuous-time stochastic equilibrium model of a market for an abstract regulated commodity, with special interest in applications to climate-relevant resources such as emissions allowances. In doing so, we put special emphasis on a mathematically rigorous representation of stochasticity and risk-aversion on behalf of market participants, as well as on a detailed (agent-based) derivation of the equilibrium dynamics in closed-form. In the process, we solve each agent's optimisation problem, along with the associated Hamilton-Jacobi-Bellman Equation, for which we are able to find an explicit solution. Furthermore, we construct a generic regulatory mechanism of dynamic resource allocation, similar in spirit to the Market Stability Reserve. Our dynamic regulatory mechanism adjusts the resource allocation based on the current system state. The system state, however, is dependent on the allocation scheme, which implies the existence of a feed-back loop between the market and the regulation. We are, notably, able to solve this inter-dependency and derive the equilibrium under this regulation. This in turn allows us to quantify the costs associated to any given parameterisation of the mechanism, adjusting its responsiveness in order to maximise the cost-efficiency of the system. What's more, by employing an adjustable responsiveness, we are able to represent an entire spectrum of policies between pure price instruments and pure quantity instruments. Thereby, we fill a significant gap in the stream of literature pioneered by [Weitzman, 1974].

The present text is divided into four chapters. In Chapter 1 we describe our motivation along with the literature context of

[6]Throughout the text, I use the term *we* instead of *I* etc. in the sense of 'guiding the reader through the text'. This formulation does not imply that I had any unwarranted assistance for my work. Accordingly, the independence of my efforts remains unaffected by such or similar formulations.

our work. In Chapter 2 we present some general concepts and observations on stochastic equilibria which will prove useful for our modelling approach. In Chapter 3 we derive our equilibrium in closed-form, before we construct our dynamic regulatory mechanism in Chapter 4. Therein, we furthermore compute total aggregate abatement costs as a function of the responsiveness of the system. This allows us to present means for the construction of a mechanism that maximises the policy's cost-efficiency. We also provide an in-depth analysis of how the above-mentioned spectrum of policies can be represented by our mechanism's parameterisation.

Some of the results presented in this text have been published by the author in a more compact and somewhat simplified fashion in [Kollenberg and Taschini, 2016a] and [Kollenberg and Taschini, 2016b]. This pertains in particular to Chapters 3 and 4; and specifically to Figures[7] *4.1, 4.2 and 4.3.*

[7] The code by which the figures have been generated is written by me in MathWorks Matlab's proprietary language. I have used Matlab to generate these figures under a valid licence.

Contents

Chapter 1

Motivation and Context

In this chapter, we will present the motivation for our quantitative work and its context in climate economics. As we shall see in Chapters 2 and 3, our model supports a wide variety of applications in the context of regulated commodities. However, we draw our inspiration specifically from commodities which are climate relevant due to the greenhouse gas emissions originating from their deployment in the production of goods and services. More specifically, we address the question of how the deployment of such resources can be limited cost-efficiently by implementing a responsive system of regulatory control.[1]

1.1 Responsive Policies in Environmental Economics

In the wake of the crisis that started in 2007 and shook the financial system to its core, we saw a surge in academic interest in how the risk of economic shocks should be incorporated in climate change policies. This particularly concerns instruments which aim at charging economic entities for the burden they impose onto society due to their greenhouse gas emissions. In economic terms, all such instruments have in common that they implicitly or explicitly put a price on that externality. Since greenhouse gases are commonly measured in terms of their contribution to climate change, relative to that of tonnes of CO_2, we refer to any such measure as a *carbon pricing instrument*.

Two such carbon pricing instruments have been of particular interest in the pertinent literature. These are Emissions Trading Systems (ETSs) on the one hand, and simple taxation on the other hand. We shall see that in many ways, these two options represent two poles at the edges of an entire spectrum of possible instruments: Situated on what we will refer to as the *quantity extreme* of this policy spectrum lies a *pure* ETS; i.e. an emissions trading system with a fixed cap[2] and a fixed allocation schedule that will not be altered ex-post. The idea of an ETS exhibits a strong conceptual appeal which makes it particularly popular among economists: In *any* ETS, the regulator makes a number of allowances available at different points in time. This may be done via free allocation, regular auctions, or a combination of the two. In a *pure* ETS, the total number of allowances that will be made available is finite and pre-determined by the regulator. At predefined intervals, the regulated entities have to present a number of allowances that suffice to cover their emissions during the respective timeframe. By means of banking and borrowing provisions[3], the regulated entities are provided with some degree of temporal flexibility, in addition to the spacial flexibility from being able to trade allowances amongst each other.[4] If adequately enforced (e.g. by sufficient penalties for non-compliance), the total number of allowances available (the cap) ensures that total emissions remain below a fixed ceiling. This relative certainty of ecological efficacy, paired with the cost-efficiency provided by spacial and temporal flexibility, make for the strong appeal of an ETS – in particular among economists. With the introduction of the European Union Emissions Trading System (EU ETS) in 2005, which is still the largest system of its kind, the concept of an ETS has found a strong flagship implementation, albeit suffering criticism and some considerable opposition.

On the other end of the policy spectrum, which we will refer to as the *price extreme*, lies pure carbon taxation; i.e. a fixed price or price trajectory is put on carbon, whereas the total amount of emissions is theoretically unbounded and subject to the taxed entities' decision making: The regulated agents simply assess their marginal benefit from emissions versus the stipulated tax, but no definitive maximum on their aggregate emissions is prescribed. However, this instrument has the advantage of being more straightforward in its concept and realisation. Furthermore, the idea of a carbon tax has a somewhat strong backing, not least from environmentalist groups, seeking a more direct and controllable form of a financial burden to be imposed on 'culprit companies'.

Pure quantity-based instruments such as emissions-trading systems with a fixed cap, and pure price-based instruments such as fixed taxes have been juxtaposed in a seminal paper by [Weitzman, 1974]. These two policies are characterised by either imposing a fixed total reduction quantity (via the cap) with a fully floating price (meaning the price is determined by the market and there is no ex-post adjustment to the cap), or by imposing a fixed price for the commodity which can (in theory) be set to induce any desired (expected) reduction quantity. In a deterministic world, the outcome of each of those instruments can be replicated by the other with an appropriate parameterisation (cap or tax level): In a stylised setting, where marginal abatement cost curves and marginal benefit curves are known and certain, a tax could be set to induce any desired emissions level, which in turn could be set to induce any desired price. However, under uncertainty, the two instruments are not necessarily perfectly equivalent and the choice of one over the other is essentially reduced to comparing the relative slopes of the marginal benefit curve (of reducing emis-

[1] Some of the results presented in this text have been published by the author in a more compact and somewhat simplified fashion in [Kollenberg and Taschini, 2016a] and [Kollenberg and Taschini, 2016b]. This pertains in particular to Chapters 3 and 4 and the figures presented therein.

[2] As is conventional, we refer to the total number of commodity units (e.g. emissions allowances) issued over the entire regulated timeframe simply as the *cap* of the regulated system.

[3] As is conventional, we often use the term *(net-)banking* to describe an agent's decision to keep commodity units allotted or purchased at one point of time in order to use them later-on. The term *(net-) borrowing*, on the other hand, describes an agent's act of using more allowances at one point in time than are currently in the agent's stock. We will provide more details on how agents can *use* commodity units in Chapter 3, where we introduce the notions of *deployment* and *abatement*.

[4] We will provide more details on the flexibilities provided to regulated agents and their significance in Section 1.4.

sions) and of the marginal abatement cost curve [Weitzman, 1974]. Whats more, the two instruments are taken to be non-responsive in the sense that the cap or tax rate is not adjusted in response to the ex-post observed outcome over time. In our analysis, we move away from a Weitzmanian comparison between individual pure instruments, i.e. price vs. quantity, and focus our attention towards the notion of responsiveness in ETSs and hybrid systems. Our approach is based on the understanding that, in analogy to monetary policy, a rigid ex-ante ETS cap may turn out to be too lenient during an economic recession and too strict in times of economic expansion. Accordingly, we add to the literature by confirming this intuition and showing in a rigorous quantitative way, how to reveal the 'sweet spot' in terms of expected cost-efficiency, when time, uncertainty, and risk-aversion are taken into account.

We will go into more detail on the topic of responsiveness and its relation to price- vs. quantity instruments in Section 1.3. However, in order to illustrate our motivation and the topicality of the subject matter, we first consider the present state of the currently largest ETS. This will, in particular, shed light on why the matter of responsiveness gained some considerable interest within the realm of climate economics.

1.2 Topical Inspiration: The European Carbon Market

The economic efficiency and ecological effectiveness of climate policy instruments can be greatly influenced by business cycles and unanticipated technological developments. Furthermore, overlapping policies can hinder each other in their efficacy, evidence of which can be found in the EU ETS: The erosion of prices for European Emissions Allowances (EUAs) in recent years has, at least partially, been attributed to concurrent policies such as feed-in tariffs for renewables. In addition, the system has been shown to suffer from a systematic unresponsiveness towards large-scale economic shocks, as impressively evidenced during the financial crisis and officially recognised by [The European Commission, 2012]. These two effects are what the drop and continuing low level of allowances prices since 2008 have primarily been attributed to. ([Grosjean et al., 2014] and [Ellerman et al., 2015a]).

To address the current situation on the European carbon market, which European authorities described as a 'structural supply-demand imbalance' [The European Commission, 2012], consultants to the EU regulators have proposed a number of plans, two of which have been made applicable legislation. The first such measure, labeled 'back-loading', temporarily offsets a portion of the issued allowances from the allocation schedule, to be definitively re-introduced later-on. In 2016, a second measure, called the 'Market Stability Reserve' (MSR) has been put into place and will start operating in 2020: Aiming to make the system "more resilient to supply-demand imbalances so as to enable the ETS to function in an orderly fashion" [The European Parliament and Council, 2015], the MSR is a much broader approach to binding the allowance supply to the economic development: The MSR will adjust the allowances allocation based on the current number ('bank') of unused allowances in circulation. More precisely, the MSR will withdraw a number corresponding to 12% of the current bank from next year's auction quantity if the current year's bank lies above 833 Million allowances. These allowances will then be placed in a so-called *reserve* and will be re-introduced to the market in batches of 100 million per year, should the bank fall below 400 million allowances. Notwithstanding these system-state-based allocation adjustments, the MSR may not be able to make the system *fully responsive* to external shocks, since the total cap over the regulated timeframe may be perceived to remain unaltered in expectation ([Perino and Willner, 2015] and [Kollenberg and Taschini, 2016b]). In this sense, the MSR can be seen as a partially responsive system, closely related, but not identical to systems with a fully floating cap. This is because the MSR does provide for some responsiveness, albeit potentially with only temporal effects: Even though there is no definitive timeline as to *when* allowances that have been withdrawn are re-introduced, eventually the reserve may well be expected to run out, reinstating the ex-ante cap.

Inspired by these shortcomings, we will consider a *spectrum* of policies, from non-responsive to fully responsive, where the cap may be altered, in order to study how different grades of responsiveness can benefit the system in terms of cost-efficiency.

1.3 Responsiveness and Price- vs. Quantity-Instruments

Taking the EC MSR as an inspiration, we will construct a generic supply control mechanism which is applicable to both physical resources and tradable externalities (such as emissions allowances): Regarding physical resources, the regulator may, for example, control the extraction of a resource by setting yearly quotas, based on the market-wide bank. In fact, the application of (discretionary) quotas is common practice in the regulation of renewable resources, as evidenced by individual transferable quotas in fishery [Grafton, 1996] or cutting quotas in forestry [Teeguarden, 1969]. Extraction quotas for non-renewable resources such as oil, on the other hand, are often subject to cartel interests which may not cover ecological objectives [Bain, 1948]. The present text abstracts away from the specifics of an individual class of resources but focuses on the matter of responsively regulating a group of agents who may deploy and trade a resource based on their individual incentives.

The concept of a responsive policy represents a well-established idea among economists. In his seminal paper on the subject of quantity- vs. price-based instruments, [Weitzman, 1974] appreciates that an optimised instrument of controlling a scarce resource or externality would implement a contingency message, indicating which realisation of the policy would be taken into action, depending on the realised state of the market and the economy. For example, the number of tradable units of the regulated commodity that are made available to the market, may be made dependent on some reference quantity that reflects the state of the system with respect to economic or ecological aspects. Often times, though, such contingent policies may be hard to bring forward in the legislative procedure. Nevertheless, in other than climate-related areas of quantity-based regulation, this impediment does not exist across the board: Central banking, for example, relies strongly on the adjustment of interest rates to be set in response to economic cycles. Seen from this perspective, a system-state-dependent regulation is not at all just an idealising concept, but very much a concrete example of applicable economic theory. In appreciation of supposed barriers of implementing contingent mechanisms and in light of the novelty of his approach, [Weitzman, 1974] concentrated his analysis on single order policies. In the present text, we pick up on his ideas and examine a spectrum of policies between the pure-price and pure-quantity extremes, simultaneously covering responsiveness over time and across different possible states of the system. In doing so, we take into account the issues of uncertainty and risk-aversion, which will prove vital to the development of an ideal policy.

Recent contributions to the literature on climate policy have helped a great deal in understanding responsiveness and the lack thereof in current legislative systems. We refer to [Doda, 2016] for a rather comprehensive review of the literature on state-based policies. In anticipation of our modelling sections, we point out that our generic mechanism takes the number of unused commodity units as its reference state-dimension. Other such dimensions may include e.g. intensity targets[5] or indices that reflect the state of the entire economy in the form of some quantitative aggregator (for example the gross domestic product of the regulated economy). Regarding climate policy cyclability, we refer to [Heutel, 2012] and [Golosov et al., 2014], while indexing rules have been studied by [Jotzo and Pezzey, 2007], [Ellerman and Wing, 2003] and [Newell and Pizer, 2008], among others.

Other approaches to establishing responsiveness in an ETS include the combination of quantity- and price-based measures into so-called 'hybrid systems', a possibility first pointed out by [Roberts and Spence, 1976]. As for an example for such combined instruments, price ceilings and floors can be introduced to an ETS by adjusting the amount of allowances supplied by the regulator in response to changes in the price level. Strict enforcement of such price bounds of course requires the regulator to be able to adjust the cap: Theoretically, if market fundamentals where able to drive the price upwards illimitably, then upwards-cap adjustments would need to be unlimited in order to adhere to a strict price ceiling. Conversely, strict price floors could only be established if the regulator were given the authority to completely remove any unused allowances from the system. A number of publications revealed that hybrid systems can offer a benefit in terms of welfare over pure price and pure quantity instruments.[6] In the present text, we focus on quantity-based allocation adjustment mechanisms, which are similar in spirit to hybrid systems that offer brice-based adjustments of the allocation quantity. We will see that quantity-based allocation adjustment mechanisms can outperform both fixed-cap and fixed-price instruments (in terms of expected compliance costs).

The present work ties up the literature on both responsiveness as well as price- vs quantity instruments by filling in the blanks between the *pure* varieties of these two instruments – with special attention to the role of uncertainty. By considering a generic mechanism, the cap-stringency[7] of which can be adjusted, a continuous spectrum of responsiveness can be achieved: This spectrum lies between a pure quantity instrument (a standard ETS with a fixed cap and a fully floating price) and a pure price instrument (a fixed carbon tax trajectory with a fully floating cap). On the one end of the spectrum lies a mechanism that is completely non-responsive towards external shocks and leaves the cap unaltered. In other words, it exhibits full cap stringency and we end up with a standard ETS – a *pure* quantity instrument. On the other end of the spectrum lies a mechanism that is fully responsive and completely offsets external shocks, ultimately inducing a fixed price trajectory at the expense of altering the ex-ante cap to a potentially unlimited extend. Hence, we end up with a de-facto tax – a *pure* price instrument. We will see that when designing a regulatory framework to the end of limiting the consumption of a scarce resource or restricting an externality (at minimal economic costs), our approach reveals an inherent trade-off between the flexibility provided by an ETS and the reduction of uncertainty that elements of price-based policies can provide. We will go into more detail on this concept in Section 1.4.

1.4 The Trade-Off Between Flexibility and Reduced Uncertainty

Compliance cost minimisation on behalf of agents under an ETS regulation is achieved in two ways: On the one hand, *spacial flexibility* is provisioned by the agents' ability to freely trade allowances amongst each other, which lets each agent adjust the amount of allowances to be taken from one point in time to the next without having to provide the necessary difference in its bank only by emissions abatement and deployment of allowances. By means of trading allowances, agents can thus benefit from their individual cost differences at any point in time, allowing them to perform an *instantaneous* optimisation. On the other hand, and connected to this provision, an ETS typically provides some *temporal flexibility*. That is, agents can, to some extend, choose the *when* of their abatement and trading decisions, within the bounds dictated by compliance requirements. In terms of their number of unused allowances, the regulator typically provides some range within which banking and borrowing of allowances is permitted. In that sense, each agent can optimise its decisions over the entire regulated period by choosing an optimal banking path, subject to regulatory constraints, and the agent's own instantaneous optimisation.

When risk-averse agents are faced with uncertainty, their inter-temporal optimisation problem is affected by how risky an investment in carbon-reduction is. Accordingly, the agents' respective banking curves will be skewed, compared to what would be optimal under risk-neutrality. This can have a potentially dramatic effect on the cost-efficiency of any system where the aggregate banking curve is not fixed. Fixing the banking curve, on the other hand, limits the inter-temporal optimisation opportunities on behalf of the regulated agents.

The above phenomenon inspires us to consider a generic mechanism that adjusts the allocation of allowances based on the current aggregate bank of unused allowances. In doing so, we leave a degree of freedom that allows us to represent the entire policy spectrum between pure price- and pure quantity-instruments: Each point on the policy spectrum will be characterised in terms of a parameter which we will call the allowance allocation *adjustment rate*. This parameter will deter-

[5] In this context, the term *intensity* usually refers to carbon-intensity of economic output. For example, we may consider a proxy for the tonnes of CO_2e emitted per EUR of gross domestic product.

[6] See e.g. [Unold and Requate, 2001], [Pizer, 2002], [Hepburn, 2006], [Fell and Morgenstern, 2010], [Grüll and Taschini, 2011], and [Fell et al., 2012].

[7] We use the term *cap-stringency* when referring in a qualitative sense to how strictly the regulator maintains the ex-ante cap. The *quantitative* parameterisation of the policy will be done through what we will refer to as an *adjustment rate*, which will quantify the level of responsiveness of a given policy. As it will turn out, the terms *cap-stringency* and *responsiveness* represent complementary notions. We refer to later chapters for details.

mine the rate by which new allowances are made available or unused allowances are removed from the allocation scheme, based on the current number of unused allowances in the system. In doing so, the rate will alter the stochastic distribution of the bank of unused allowances, tightening it towards one end of the spectrum (towards a pure price instrument) and loosening it towards the other end (towards a pure quantity instrument).

Based on our model developed in Chapter 3, we subsequently devote Chapter 4 to analysing in depth, how different levels for the adjustment rate affect the agents' banking behaviour along with aggregate costs. However, in order to motivate our approach, we briefly sketch the relevant intuition in the following.

When the adjustment rate is increased in favour of a more responsive policy, agents' inter-temporal cost-saving opportunities are reduced in favour of a market price less susceptible to external shocks. In other words, if the system moves away from a pure quantity instrument, the free-floating property of the bank is restrained and shocks are absorbed by the regulator to a larger degree. In particular, when the cap-stringency is fully relaxed and the system is fully responsive, exogenous shocks are perfectly offset by changes to the allocations and agents' expectations about their future required abatement do not change throughout the regulated period. As a result, a fixed price trajectory is induced and we obtain a de-facto tax-like system: When market shocks are perfectly offset by changes to the cap, expectations about future required abatement do not change due to exogenous shocks.

With this in mind, a responsive mechanism which adjusts allowance allocation based on the current aggregate bank of allowances may deeply affect the behaviour of regulated agents: In fact it may constrain agents' banking behaviour to such an extend that the benefit it provides by lowering the impact of external shocks is hypercompensated by the limitation it imposes on inter-temporal cost-saving opportunities.

In Chapters 3 and 4, we quantify this relationship to the end of exposing a trade-off along the spectrum and showing that there is indeed an optimal policy, which lies not necessarily on the one nor the other end of the policy spectrum but may imply a mixed policy to be ideal. We shall see that the time component as well as uncertainty is vital to this insight as well as to our conclusions in terms of an optimal policy.

Before we develop our model and derive our policy implications, we will devote Chapter 2 to making some general assumptions and observations.

Chapter 2

Market Equilibria

Practitioners and academics in economics and financial mathematics are often interested in finding projections of future developments on markets for specific commodities and how these developments may be affected by regulatory intervention. Such regulation problems are particularly relevant in environmental economics, where an unregulated economy may consume more of a commodity, or impose more externalities onto society, than is in society's own interest, both today and in the distant future.

On the quantitative side, markets pertaining to a specific commodity or externality are typically modelled with the intention of inferring projections of future developments under different regulatory schemes. From a regulatory perspective, we may for example be interested in finding the best available trade-off between the benefit we obtain from the present consumption of a commodity on the one hand, and the damage it does to future generations on the other hand. Here, the damage may be inflicted either by reducing the stock of the commodity available for future consumption, or by the environmental damages incurred from its consumption today. This perspective has been taken in a stream of literature pioneered by [Hotelling, 1931], whose main area of interest lay in the optimal temporal distribution of consumption from a societal perspective.

This societal inter-temporal trade-off breaks down, however, in in the context of of anthropogenic climate change: One the one hand, catastrophic damages are expected to occur, should the cumulated 'consumption' of a commodity, namely an atmosphere with acceptable levels of greenhouse gases, exceed a certain threshold: By the special nature of climate-relevant commodities, it is imperative and without alternatives to observe certain limits on this 'consumption of clean air'. On the other hand, CO_2 and similar greenhouse gases are what is commonly referred to as a *stock pollutant*, meaning in particular, that the marginal societal benefit of emitting less greenhouse gases can be regarded as negligibly small, compared to the economic costs of reducing the same amount (as long as catastrophic damages are avoided).[1]

The present work is therefore set to explore how the deployment of resources entailing the emission of greenhouse gases can be regulated with minimal expected economic costs. Furthermore, our work incorporates particularly relevant challenges faced by the regulator, such as long timeframes and severe economic uncertainties.

A natural and common approach to modelling markets for any type of commodity relies on the notion of a (Nash-type) market equilibrium; that is, a collection of strategies, each for one agent, such that none of the agents has an incentive to individually deviate from its strategy. As simple examples[2] suffice to show, equilibria are, in general, by no means a guarantor of cost-efficiency or collective optimisation. In that sense, the entirety of agents may have an incentive to deviate from its collection of individual strategies, while each individual may not. The Nash-type concept of market equilibria therefore reflects how social and economic catastrophes can be the result of individual optimisation - a phenomenon which may strongly suggest regulatory intervention, be it discretionary or based on a transparent set of rules.

Besides their general conceptual appeal, Nash-type equilibria are a straightforward choice in particular for approaching regulatory problems, since each agent typically observes a market state which can be affected by regulation. In the present work, we will use the classical notion of a Nash-type equilibrium, garnished with some useful complications such as continuous time and stochasticity. Thereby, we can observe the reaction a specific policy may induce on behalf of market participants. In Chapter 3 we will develop such an equilibrium model and solve it in closed-form, where the system state (process) may be affected by regulatory intervention. In Chapter 4 we will impose a generic regulatory policy to obtain the equilibrium under regulation. Therein, the regulatory policy is made dependent on the (system) state process itself, which implies the existence of a feed-back loop between the market and its regulation. In Chapter 4, we solve this interdependence and draw some policy conclusions which are particularly relevant in light of the recent reform of the European Union Emissions Trading System (EU ETS).

A very common approach to equilibrium modelling is based in one way or another on the notion of a representative agent, i.e. one whose characteristics are in some way representative of the entirety of agents on the considered market.[3] Albeit usually not explicitly stated, we may read this common notion as based on the assumption that any agent on the market is characterised by a set of parameters, and that these parameters determine the representative agent's characteristics simply as their average of some form. For example, we may consider a number of agents in perfect competition who face production costs for some commodity. These costs may arise from extracting a resource in situ and delivering it to the market, or generating emissions allowances by abatement to sell them on the secondary market. More importantly, these costs may depend on each individual agent's instantaneous production volume. In any case, the common understanding is that, given perfect competition and convex (in terms of production volume) production costs, one should observe an equalisation of *average* production costs across all agents on the one hand and the price for the produced commodity on the other hand. This idea is based on the sentiment that if the average price were higher than average production costs, those agents on the long end of this imbalance, i.e. those with production costs lower than the price at which they may sell part of their commodity stock, were to expand their production until that imbalance ceases: There is a free lunch in for each those agents, since

[1] Cf. [Newell and Pizer, 2008].

[2] A well-established example of this is the 'prisoner's dilemma' where the separation of agent's decision making leads to a collectively sub-optimal outcome. See for example [Poundstone, 1992].

[3] We refer to [Hartley and Hartley, 2002] for an overview and critique on the subject.

the non-existence of market power (due to the assumption of perfect competition) and the resulting lack of a non-negligible price-impact makes producing *more* of the commodity profitable from an individual perspective. When those agents on the long end expand their production, their per-unit production costs would increase. Conversely, agents on the short end have an incentive to *decrease* their production, since effectively, those agents incur net losses for each unit produced. One should thus observe that the production pattern instantaneously changes to a state where each individual agent has either no incentive to change their production or no way of doing so while still finding a counterparty on the commodity's market.

The notion of a representative agent and the conclusions drawn from the simple equilibrium argument outlined above often fully suffice to derive or describe economic principles and phenomena, such as Hotelling's classical rule for the value of in-situ resources[4]. However, here we are interested in each individual agent's strategy and find that we are able to obtain quite a satisfying result: For each agent, we are able to obtain a closed-form solution to the Hamilton-Jacobi-Bellman Equation, associated to the agent's optimisation problem. In connection with the notions of weak and strong equilibria introduced in the next chapter, this will be of great technical value – in addition to its conceptional appeal.

The rest of this chapter is devoted to introducing some general assumptions and concepts that will put our model into context. In particular, we will define the notions of *weak* and *strong* equilibria under different measures, which are deeply connected to our representation of risk-aversion.

[4] See [Hotelling, 1931].

2.1 First- and Second-Level Uncertainty

We conceptualise our model based on the notions of *first-level* and *second-level* uncertainty, which we will establish in the following.

Concerning first-level uncertainty, let $(I, \mathcal{I}, \mathbb{I})$ denote a probability space, where the set I represents all 'conceivable agents', $\mathcal{I}$ is a σ-field on I, and $\mathbb{I}$ is a probability measure on $(I, \mathcal{I})$. We can think of the randomness modelled here as a representation of the uncertainty all agents and the regulator face, regarding the *characteristics* of (other) agents on the market, a term we will use to refer to any quantity that depends on i: Any $\mathbb{I}$-measurable random variable on $(I, \mathcal{I})$ shall be labelled an *(agent) characteristic*.

These characteristics include, for example, any agent i's (exogenous) initial resource endowment B_0^i: An agent's initial endowment is defined as a realisation B_0^i of the random variable $B_0 : I \to \mathbb{R}$, measurable with respect to $\mathbb{I}$, where $\mathbb{R}$ is equipped with the Borel-σ-field. Accordingly, we consider any agent's (endogenous) time-t stock of unused commodity units (henceforth labelled *bank*) B_t^i, on which we will provide more details further below.

The second exogenous characteristic we consider is each agent's time-t exposure σ_t^i to systemic risks, subject to the technological profile and overall market share of agent i: For example, a company that relies heavily on assets that are based on non-renewable resources will be more affected by the price and regulation of these resources than a company that invested heavily in renewables-based assets. The parameter σ_t captures how the impact of events affecting the entire economy is distributed across regulated agents.

The notion of first-level uncertainty introduced above expands the notion of a 'continuum of agents', which is conventionally used in economics to capture the idea of a large mass of agents, each with negligible market power. The lack of realism in the that approach is usually not relevant as long as one is solely interested in results pertaining to a representative agent; i.e. one whose characteristic are in some way 'average'.

By construction, our approach is more general than an approach based on a representative agent, in the sense that 'average' characteristics can be obtained from (first-level-) random agents by simple integration: Given any characteristic X_t, we define the *aggregate* X_t as $\int_{\mathbb{I}} X_t^i d\mathbb{I}(i)$. That is, all realisations are weighted with the probability of their respective occurrence and integrated. Where appropriate, we will also use the notation $X_t^I = \int_{\mathbb{I}} X_t^i d\mathbb{I}(i)$ for aggregate quantities.

In the following we will, for any $\mathbb{I}$-measurable quantity, assume that is is integrable with respect to $\mathbb{I}$. Furthermore, we assume that $\mathbb{I}(i) = 0$ for all $i \in I$; i.e. any particular *manifestation* of some characteristic has a negligible impact on the *aggregate* of that characteristic; but it is rather the probabilistic distribution of characteristics across I that governs the state and dynamics of the system.

Our approach has the conceptual appeal that we can consider strategies and results for individual *realised* agents, which depend on their realised characteristics. In our setup, each agent then faces *second-level uncertainty* with respect to exogenous factors:

We model the uncertainty with respect to those exogenous (in particular economic and technological) factors throughout a regulated timeframe $[0, T]$, $T < \infty$, as an additional complication: To this end, consider a second (filtered) probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, where $\mathcal{F}$ is a σ-field on Ω and $(\mathcal{F}_t)_{t\in[0,T]}$ is a filtration of $\mathcal{F}$. As is conventional, we will refer to $\mathbb{P}$ as the *objective measure*. This terminology will become relevant when we consider risk-averse agents and change to a risk-neutral measure $\mathbb{Q} \sim \mathbb{P}$ on $(\Omega, \mathcal{F})$,[5] in order to derive the equilibrium dynamics in closed-form.

In analogy to integration with respect to $\mathbb{I}$, we often integrate with respect to $\mathbb{P}$ (resp. $\mathbb{Q}$), where we use the standard term ($\mathbb{P}$-, resp. $\mathbb{Q}$-) *expectation*. In this context, we adopt for any $t \in [0, T]$ the following notation for the time-t (conditional) expectation with respect to $\mathbb{P}$: We put $\mathbb{E}_t^{\mathbb{P}}[\cdot] = \mathbb{E}^{\mathbb{P}}[\cdot|\mathcal{F}_t]$, and $\mathbb{E}^{\mathbb{P}}[\cdot] = \mathbb{E}_0^{\mathbb{P}}[\cdot]$; and accordingly for $\mathbb{Q}$. Furthermore, when integrating with respect to the Lebesgue-measure over the time interval $[0, T]$, we refer to the result as *total*: For example, if α_t^i is agent i's abatement (of emissions or physical commodity deployment), we refer to $\int_0^T \alpha_s^i ds$ as i's *total* abatement.

We will revisit agents' characteristics in Chapter 3, where we specify each agent's optimisation problem and provide a derivation of the equilibrium.

As for now, we will simply let $\mathbb{Q}$ denote some probability measure on $(\Omega, \mathcal{F})$. In Section 2.2 we will introduce the notions of *strong* and *weak* equilibria. As we shall see, these notions and their relation are intimately linked to the respectively applicable measure. As the term 'risk-neutral measure' suggests, the notions of weak and strong equilibria (and the respectively applicable measures) are related to the concept of risk-aversion and its correspondence in our modelling approach. We will establish this link in Section 2.3. These expositions will lay the ground for our analysis and equilibrium solution in the later sections.

2.2 Strong and Weak Equilibria

We want to extend the conventional nomenclature in the context of arbitrage strategies to market equilibria in the context of regulated commodities. To this end, consider some adapted stochastic price process $P = (P_t)_{t\in[0,T]}$. Furthermore, consider for each agent $i \in I$ a cost functional $J(\gamma^i)$, where γ^i denotes agent i's strategy, element of some set of *attainable strategies* $\mathcal{G}^i$, which is specific to the applied model (and defined for our model in the next chapter). Note that agent i's costs $J(\gamma^i)$ also depend on the price process P. However, in the following we will notationally omit the dependence of $J(\gamma^i)$ on P to improve legibility. Furthermore, we will leave the technical specifications of J for the next chapter to keep our discussion general in this context. At this point, it suffices to assume that $J(\gamma^i)$ is integrable with respect to $\mathbb{P}$ (and, where applicable, with respect to $\mathbb{Q}$) for all $\gamma^i \in \mathcal{G}^i$; and that each γ^i is a measurable stochastic process, adapted to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$.

We call a (first-level) random variable $\gamma : i \mapsto \gamma^i$ a *strategy bundle*, which is the notion in the center of our attention regarding equilibria. The cost functional J will capture the costs associated to complying with regulations with respect to the regulated commodity.

[5] As is conventional we write $\mathbb{Q} \sim \mathbb{P}$ (verbally: '$\mathbb{Q}$ is equivalent to $\mathbb{P}$'), when $\mathbb{Q}(A) = 0$ if and only if $\mathbb{P}(A) = 0$, for all $A \in \mathcal{F}$.

Consider some *benchmark* strategy $\tilde{\gamma}^i \in \mathcal{G}^i$ for some $i \in I$. We then say that a strategy $\gamma^i \in \mathcal{G}^i$ *strongly outperforms* $\tilde{\gamma}^i$ with respect to $\mathbb{P}$, and given P, if it $\mathbb{P}$-a.s. doesn't increase costs and decreases costs with positive $\mathbb{P}$-probability, compared to the benchmark $\tilde{\gamma}^i$ (Definition 1).

Strong Outperformance, Weak Optimality

Definition 1. *(i) We say that a strategy $\gamma^i \in \mathcal{G}^i$* **strongly outperforms** *a benchmark strategy $\tilde{\gamma}^i \in \mathcal{G}^i$ with respect to $\mathbb{P}$, and given P, if*

$$\mathbb{P}\left[J(\gamma^i) \leqslant J(\tilde{\gamma}^i)\right] = 1$$
$$\text{and} \quad \mathbb{P}\left[J(\gamma^i) < J(\tilde{\gamma}^i)\right] > 0.$$

(ii) If no strategy $\gamma^i \in \mathcal{G}^i$ exists that strongly outperforms $\tilde{\gamma}^i$ with respect to $\mathbb{P}$, and given P, then $\tilde{\gamma}^i$ is called **$\mathbb{P}$-weakly optimal** *in $\mathcal{G}^i$, given P.*

Note that in the BAU case, i.e. in case of no regulation, it is natural to consider vanishing costs as the benchmark scenario. That is, it is natural to consider a strategy $\tilde{\gamma}^i$ for which $J(\tilde{\gamma}^i) = 0$, $\mathbb{P}$-a.s., and simply measure performance in terms of the costs themselves (as opposed to relative to a benchmark). In this case, we thus reobtain the conventional notion of a *strong arbitrage strategy*, i.e. a strategy for which a non-negative profit (non-positive costs) is gained $\mathbb{P}$-a.s. and a positive profit (negative costs) is obtained with positive $\mathbb{P}$-probability.

We also want to consider a notion of *weak outperformance* (Definition 2), in analogy to the notion of a *weak arbitrage strategy* in the conventional terminology. We consider such *weak outperformance* under a different measure $\mathbb{Q}$ to the end of deriving sufficient conditions for the existence of such a measure $\mathbb{Q}$, equivalent to $\mathbb{P}$, such that weak outperformance is not attainable for some strategy $\tilde{\gamma}^i$.

Weak Outperformance, Strong Optimality

Definition 2. *(i) We say that a strategy $\gamma^i \in \mathcal{G}^i$* **weakly outperforms** *a benchmark strategy $\tilde{\gamma}^i \in \mathcal{G}^i$ with respect to $\mathbb{Q}$, and given P, if*

$$\mathbb{E}^{\mathbb{Q}}\left[J(\gamma^i)\right] < \mathbb{E}^{\mathbb{Q}}\left[J(\tilde{\gamma}^i)\right].$$

(ii) If no strategy $\gamma^i \in \mathcal{G}^i$ exists that weakly outperforms $\tilde{\gamma}^i \in \mathcal{G}^i$, with respect to $\mathbb{P}$, and given P, then $\tilde{\gamma}^i$ is called **$\mathbb{Q}$-strongly optimal** *in $\mathcal{G}^i$, given P.*

Note that for a benchmark strategy $\tilde{\gamma}^i$ that yields vanishing costs, this notion coincides with the conventional notion of an *in-expectation arbitrage strategy*. The following lemma justifies the terms *weakly* and *strongly* in the above definitions.

The Relation Between Strong and Weak Outperformance and Optimality

Lemma 3. *We have with respect to any measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ and any adapted price process P:*

(i) Any strategy γ^i that strongly outperforms $\tilde{\gamma}^i$ also weakly outperforms $\tilde{\gamma}^i$.

(ii) If $\tilde{\gamma}^i$ is strongly optimal, then it is also weakly optimal.

Proof. In the following, $\mathbb{Q}$ denotes some probability measure on $(\Omega, \mathcal{F})$.

(i) Let γ^i strongly outperform $\tilde{\gamma}^i$ and assume γ^i does not weakly outperform $\tilde{\gamma}^i$. Then, by definition of weak outperformance, we have

$$\mathbb{E}^{\mathbb{Q}}\left[J(\gamma^i)\right] \geqslant \mathbb{E}^{\mathbb{Q}}\left[J(\tilde{\gamma}^i)\right].$$

Since γ^i strongly outperforms $\tilde{\gamma}^i$ we have $\mathbb{Q}[J(\gamma^i) \leqslant J(\tilde{\gamma}^i)] = 1$ and hence

$$0 \leqslant \mathbb{E}^{\mathbb{Q}}\left[J(\gamma^i) - J(\tilde{\gamma}^i)\right] = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{J(\gamma^i) \leqslant J(\tilde{\gamma}^i)\}}\left(J(\gamma^i) - J(\tilde{\gamma}^i)\right)\right] \leqslant 0,$$

from which we obtain

$$\begin{aligned} 0 =& \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{J(\gamma^i) \leqslant J(\tilde{\gamma}^i)\}}\left(J(\gamma^i) - J(\tilde{\gamma}^i)\right)\right] \\ =& \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{J(\gamma^i) < J(\tilde{\gamma}^i)\}}\left(J(\gamma^i) - J(\tilde{\gamma}^i)\right)\right] + \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{J(\gamma^i) = J(\tilde{\gamma}^i)\}}\left(J(\gamma^i) - J(\tilde{\gamma}^i)\right)\right]. \end{aligned}$$

It follows that $\mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{J(\gamma^i) < J(\tilde{\gamma}^i)\}}\left(J(\gamma^i) - J(\tilde{\gamma}^i)\right)\right] = 0$, which holds if and only if $\mathbb{Q}\left[J(\gamma^i) < J(\tilde{\gamma}^i)\right] = 0$. Hence γ^i did not strongly outperform $\tilde{\gamma}^i$ to begin with.

(ii) Let $\tilde{\gamma}^i$ be strongly optimal and assume $\tilde{\gamma}^i$ is not weakly optimal. Then, by definition, there exists a strategy γ^i that strongly outperforms $\tilde{\gamma}^i$. But then γ^i also weakly outperforms $\tilde{\gamma}^i$ and hence $\tilde{\gamma}^i$ could not have been strongly optimal to begin with. $\square$

Q-Strongly Optimal and P-Weakly Optimal Strategies

Lemma 4. *If there exists a $\mathbb{Q} \sim \mathbb{P}$ such that $\tilde{\gamma}^i$ is $\mathbb{Q}$-strongly optimal, given P, then $\tilde{\gamma}^i$ is $\mathbb{P}$-weakly optimal, given P.*

Proof. Let $\mathbb{Q} \sim \mathbb{P}$ and let $\tilde{\gamma}^i$ be $\mathbb{Q}$-strongly optimal, given P. By Lemma 3 we know that $\tilde{\gamma}^i$ is also $\mathbb{Q}$-weakly optimal. Now assume $\tilde{\gamma}^i$ was not $\mathbb{P}$-weakly optimal. By definition, there would then exist a strategy γ^i that $\mathbb{P}$-strongly outperformed $\tilde{\gamma}^i$. In particular, this would imply that

$$\mathbb{P}\left[J(\gamma^i) \leqslant J(\tilde{\gamma}^i)\right] = 1.$$

Since $\mathbb{Q} \sim \mathbb{P}$ we would then obtain that

$$\mathbb{Q}\left[J(\gamma^i) \leqslant J(\tilde{\gamma}^i)\right] = 1.$$

But since $\tilde{\gamma}^i$ is $\mathbb{Q}$-weakly optimal, it is not $\mathbb{Q}$-strongly outperformed and hence

$$\mathbb{Q}\left[J(\gamma^i) < J(\tilde{\gamma}^i)\right] = 0,$$

and, again since $\mathbb{Q} \sim \mathbb{P}$, this would imply

$$\mathbb{P}\left[J(\gamma^i) < J(\tilde{\gamma}^i)\right] = 0.$$

Thus, γ^i does not $\mathbb{P}$-strongly outperform $\tilde{\gamma}^i$ to begin with and hence $\tilde{\gamma}^i$ is $\mathbb{P}$-weakly optimal. □

Based on our observations in Lemma 4, we can now introduce the notions of *strong and weak equilibria* (Definition 6), in correspondence to the notions of strong and weak optimality. To this end, we first define the *market clearing* property with respect to price processes (Definition 5).

Market Clearing Price Process

Definition 5. *Consider some strategy bundle γ, whose individual strategies $\gamma^i = (\alpha^i, \beta^i)$ may depend on a price process P. We interpret the process β^i as agent i's trading strategy; i.e. β^i_t denotes the number of commodity units sold by agent i at time t. If $\int_I \beta^i_t d\mathbb{I}(i) = 0$ for all $t \in [0, T]$, we say that the process P is* market clearing.

Strong and Weak Equilibria

Definition 6. *For a strategy bundle γ^*, let there be a unique market clearing price P. If, for all $i \in I$, $\gamma^{i,*}$ is strongly (weakly) optimal with respect to some measure, and given P, we call γ^* a* strong (weak) equilibrium strategy bundle *and P its* strong (weak) equilibrium price process*, with respect to the same measure.*

As we will expound in Section 2.3, the notions of *strong* and *weak equilibria* are particularly helpful when we are dealing with risk-aversion. To this end, we establish in Proposition 7, how the existence of a strong equilibrium under one measure can imply the existence of a weak equilibrium under a different measure.

The Relation between Q-Strong and P-Weak Equilibria

Proposition 7. *If there exists a measure $\mathbb{Q} \sim \mathbb{P}$ such that γ^* is a $\mathbb{Q}$-strong equilibrium strategy bundle then γ^* is a $\mathbb{P}$-weak equilibrium strategy bundle.*

Proof. This is a direct consequence of Lemma 4. □

In general terms, Proposition 7 can be of great use when the stochastic properties of a model are such that a $\mathbb{P}$-strong equilibrium cannot readily be established. For example, and in anticipation of our approach, let $W^{\mathbb{P}}$ denote a $\mathbb{P}$-Brownian motion driving the stochasticity behind our model. If we can show that there exists a measure $\mathbb{Q} \sim \mathbb{P}$ such that $W^{\mathbb{Q}} = W^{\mathbb{P}} + \Lambda$ is a $\mathbb{Q}$-Brownian motion for some process Λ, then we may be able to circumvent the stochastic intractabilities of the model with the help of the process Λ, and show that there exists a $\mathbb{Q}$-strong equilibrium that is also a $\mathbb{P}$-weak equilibrium by Proposition 7. Apart from thus being useful on its own and from a technical perspective, Proposition 7 holds some appealing value for our intuition and interpretation, which we shall explore in the next section.

2.3 Risk-Aversion

Assume some regulatory requirements which stipulate limitations on the agents' inflow and deployment of some commodity. At time t, we then assess how many units of the commodity have to be conserved by each agent over the remaining regulated timeframe. Let R^i_t denote this time-t *expected residual abatement requirement*[6]. When each agent tries to optimise its strategy in response to the regulation, the (second-level) uncertainty to which $R^i = (R^i_t)_{t\in[0,T]}$ is subjected becomes an important factor.

In our model, we will be dealing with two types of investments that any agent can pursue in order to adhere to those regulatory requirements: These are reduction of deployment or generation of commodity units (i.e. abatement) on the one hand, and purchase of additional units on the other hand. Abatement, on the one hand, acts on the quantity dimension and inflicts costs based on alterations of an agent's production process. On the other hand, the purchase of commodity units is a simple financial act, the costs of which emerge from market conditions such as the equilibrium price. Ultimately, compliance with the regulatory requirements can be attained with any mix of quantity investments and financial investments.

However, both of these types of investment are subject to the

[6]We will specify the notion of *(expected) residual abatement requirement* more precisely in Chapter 3, along with the necessary modelling assumptions.

same investment-specific risks, since generation and purchase both yield the same asset for the investor in terms of its net-position R_t^i: namely some Δ_t^i additional commodity units in agent i's stock, which reduce the expected residual abatement requirement R_t^i by the same amount Δ_t^i.

We denote the time-t instantaneous costs borne by agent i by $v(Z_t^i, \gamma_t^i)$, where γ^i denotes agent i's strategy (i.e. a control process) and Z^i denotes agent i's *state process*, to be specified in Chapter 3: At each time t, agent i observes a number of variables relevant to his decision making, comprising the time-t state Z_t^i. We will specify the above costs along with tractability assumptions later in the text but keep the discussion general in the context of the preliminary discussion in this section.

We assume that there exists an alternative investment[7] for speculative agents with negligible risk. As is conventional, we refer to such an asset as *risk-free*. We denote the (exogenous) rate of return of this risk-free asset by $\mu > 0$ and let μ be constant in order to improve clarity. It will be straightforward to see that the generality of the results is not affected by this simplification.

The above-mentioned asset-specific risks are taken into account by any risk-averse agent, when assessing the time-value of the investment. Accordingly, the time-t rate of return of the investment should include a risk-premium ζ_t on top of the money-market interest rate μ. In Section 3.7 of the next Chapter, we will demonstrate how this risk-premium is connected to the notions of strong and weak equilibria and the associated change of measure from $\mathbb{Q}$ to $\mathbb{P}$. In particular, we will find that the intuition behind a risk-premium aligns nicely with our analytical approach – something that we shall exploit when we examine the intuition of policy implications in Chapter 4.

If agents were risk-neutral, we could characterise their $\mathbb{P}$-strong optimisation problem simply as trying to obtain a strategy $\gamma^{i,*}$ such that expected total costs, discounted at the risk-free interest rate $\mu > 0$, are minimised; i.e.

$$\gamma^{i,*} = \underset{\gamma^i \in \mathcal{G}^i}{\operatorname{argmin}}\ \mathbb{E}^{\mathbb{P}} \left[\int_0^T e^{-\mu t} v(Z_t^i, \gamma_t^i)\ dt \right], \tag{2.1}$$

where $\mathcal{G}^i$ is some set of *attainable* strategies, which will be specified in the next chapter. However, when firms are risk-averse, the minimisation problem in Equation (2.1) is no longer valid, since the risk associated to any investment in the commodity should be taken into account on behalf of the agents. However, if we can identify a measure $\mathbb{Q} \sim \mathbb{P}$, such that the rate of return of P under $\mathbb{Q}$ consists only of the risk-free rate μ, we can consider the problem

$$\gamma^{i,*} = \underset{\gamma^i \in \mathcal{G}^i}{\operatorname{argmin}}\ \mathbb{E}^{\mathbb{Q}} \left[\int_0^T e^{-\mu t} v(Z_t^i, \gamma_t^i)\ dt \right], \tag{2.2}$$

which takes risk-aversion into account by re-weighting the probabilities of events concerning the uncertainty in P (and R). Equation (2.2) corresponds to the problem of finding a $\mathbb{Q}$-*strongly optimal* strategy in the sense of Definition 2. Proposition 7 then yields that if there exists a measure $\mathbb{Q} \sim \mathbb{P}$ and a $\mathbb{Q}$-strong equilibrium strategy bundle γ^* (based on the above cost functional), then γ^* is also a $\mathbb{P}$-weak equilibrium strategy bundle.

In Chapter 3, we will first posit the existence of such a measure $\mathbb{Q} \sim \mathbb{P}$ and identify a $\mathbb{Q}$-strong equilibrium strategy bundle γ^*. In particular, we will obtain a closed-form solution for the associated equilibrium price process P under $\mathbb{Q}$. This will in turn allow us to present a sufficient condition on the functional representation of risk-aversion, such that the existence of $\mathbb{Q} \sim \mathbb{P}$ is guaranteed. In effect, we will thus prove that, under this condition, γ^* is a $\mathbb{P}$-weak equilibrium strategy bundle, in line with Proposition 7.

[7] Note that we do not exclude the possibility of agents for whom the terminal condition is already satisfied from the outset. Hence, speculative investment in the commodity is covered in our model.

Chapter 3

A Stochastic Equilibrium Model

In this chapter, we will present a partial equilibrium model where agents each receive, use, and possibly generate, as well as trade units of some commodity, the consumption of which the regulator seeks to reduce up to a finite time-horizon T. Our prime example for this are emissions allowances, each of which securitises the right to emit one (metric) tonne of greenhouse gases, measured in CO_2-equivalents (CO_2e).

We assume that for each agent, the total number of units used must be met by an equal number of units held at the end T of the regulated timeframe. For example, the finiteness of our time-horizon T can be taken as based upon climate negotiations, which typically impose the reduction of greenhouse gas emissions to a certain level up to a pre-determined date. Notwithstanding technical and political considerations as to how this timeframe should ideally be set, we take it as given that due to political or technical circumstances, such timeframes typically exist and their finiteness can safely be assumed. We will provide more details on the relevance of T with respect to our model and its interpretations below.

3.1 Deployment, Abatement and Trading

The objective of this chapter is to give a comprehensive derivation of a stochastic equilibrium in continuous time on a market for a single regulated commodity, with a finite time-horizon. Each agent faces a problem of optimal *deployment* (or *abatement* thereof), as well as trading of the same commodity over the same period $[0, T]$, $T < \infty$: In what follows, we will narrow down these notions and introduce the key quantities necessary to describe each agent's decision problem.

The term *deployment* refers to using the resource to profitably produce any number of goods or services, or assigning a specific amount of the resource to use in production at a fixed future date. More precisely, we let the term refer to the non-retrievable portion in the amount of the commodity used in the production process. Particularly relevant examples for those resources are oil, coal, natural gas and CO_2 emissions allowances, used for example for the generation of electrical power. Other relevant commodities include storable, non-renewable resources like precious metals or rare earths, required in the production of consumer electronics. Due to limitations of current recycling technology (in case of physical resources) or due to the producer emitting a stock-pollutant (such as CO_2), at least part of the resources deployed here are irretrievably *consumed* in the production process.[1]

We use the term *abatement* when referring to the difference between the Business-as-Usual (BAU) deployment (i.e. deployment if the market were unregulated) and the deployment obtained in a regulated market, each pertaining to the same period of time and state of the world[2]. When the resource is regulated, agents may, for example, have an incentive to extract more or less of a natural resource in-situ, such as oil or lignite. Similarly, agents may have an incentive to reduce their deployment of the resource, e.g. by scaling down production or by modernising their infrastructure to require less emissions allowances to produce the same output, which is commonly referred to as emissions abatement. Since we are only interested in the difference in the deployment of the resource (i.e BAU vs. regulated market), we subsume any of the above forms of deployment reduction under one term, simply labelled *abatement*.

The rules and intuition for the finite time horizon T are as follows: At time T, we require all agents to cover their cumulated deployment of the resource (comprised of its BAU deployment, net its abatement) with an equal number of units obtained from the regulator or purchased on the market. Depending on the application, we may think of T as the production launch of those production units to which the agent assigned the resource units; or as the end of a regulated timeframe, where each agent's cumulated deployment is measured against its stock in the commodity. While the former mainly applies to physical resources that are purchased before the respective units are produced, the latter applies in particular to emissions allowances, where compliance is required in regular intervals (e.g. one year or one EU ETS phase[3]) and with respect to the emissions in that period.

Regarding the BAU deployment, we assume that its relevant cumulated volume is uncertain only up to some deterministic point in time $\theta \in [0, T]$, such that there is always at least some time left for each agent to close its remaining imbalance. In case of the EU ETS, this may, for example, be interpreted as corresponding to a 'recording period' of one calendar year $[0, \theta]$, the emissions during which must be covered with an equal number of allowances in May of the subsequent year (i.e. $T > \theta$). This correspondence applies similarly, when we interpret $[0, \theta]$ as a range of calendar years and T as May of the year subsequent to the last year in question. As for physical resources assigned to later production, it is also natural to assume that any unanticipated variations to the demand for the goods or services offered by the agent, are also only taken into account some time before production commences. (That is, production plans are not changed 'last minute'.)

Note that we endogenise the agents' deployment and trading behaviour with respect to the regulated resource only. That is, we consider a partial equilibrium and leave, in particular, customers' demand for any goods produced by the agents exogenous. Furthermore, we only consider optimal deployment in

[1] In this context, atmospheric pollution can be colloquially characterised as the 'consumption of clean air'.

[2] That is, when we compare *BAU* versus the *regulated market*, each notion pertains to the same points in time and the same realisation of all relevant random variables.

[3] The EU ETS is temporally divided into several *phases*, each a number of years long. Between these phases, changes to the regulatory framework may be made.

the sense of controlling the amount of a resource used or sold at any point in time rather than an optimisation of resource assignment across an agent's production units. Thereby, we observe a conventional and sensible principle in the economics of climate change, namely that the social cost of deployment is fully equivalent across all spacial distributions; in other words: In terms of environmental impact, it does not matter *where* emissions of greenhouse gases (GHG) occur - the climate impact is only determined by the cumulated amount of GHG in the atmosphere.

Measurability Assumptions: For $t \in [0,T]$ let $\mathcal{B}([0,t])$ denote the Borel-σ-field on $[0,t]$ and let $\mathbb{L}^{[0,t]}$ denote the Lebesgue-measure on $\mathcal{B}([0,t])$. We then note that $\mathbb{P}$, $\mathbb{I}$ and $\mathbb{L}^{[0,t]}$ are, in particular, σ-finite and hence, by Carathéodory's Theorem, the product measure $\mathbb{L}^{[0,t]} \otimes \mathbb{P} \otimes \mathbb{I}$ on the product-σ-field $\mathcal{B}([0,t]) \otimes \mathcal{F}_t \otimes \mathcal{I}$, exists and is unique. Therefore, Fubini's Theorem applies for all functions that are integrable with respect to this product measure and integration with respect to each individual of the above measures can be conveniently swapped in the sense of that Theorem.

We refer to any process that that is adapted to the filtration $(\mathcal{F}_t \otimes \mathcal{I})_{t\in[0,T]}$ simply as being *adapted*. Accordingly, we refer to any process $X = (X_t)_{t\in[0,T]}$ as being *progressively measurable*, if it is progressively measurable with respect to $(\mathcal{F}_t \otimes \mathcal{I})_{t\in[0,T]}$, i.e. if for all $t \in [0,T]$, the function $(s,\omega,i) \mapsto X^i_s(\omega)$ is measurable on $[0,t] \times \Omega \times I$, equipped with the product-σ-field $\mathcal{B}([0,t]) \otimes \mathcal{F}_t \otimes \mathcal{I}$.

Throughout the text, whenever we consider a stochastic process that is progressively measurable with respect to $(\mathcal{F}_t)_{t\in[0,T]}$, we assume that it is also progressively measurable with respect to $(\mathcal{F}_t \otimes \mathcal{I})_{t\in[0,T]}$. Furthermore, if a process is integrable with respect to $\mathbb{L}^{[0,t]} \otimes \mathbb{P}$ for any $t \in [0,T]$, we assume that it is integrable with respect to $\mathbb{L}^{[0,t]} \otimes \mathbb{P} \otimes \mathbb{I}$.
The above measurability assumptions are taken accordingly when $\mathbb{P}$ is replaced by $\mathbb{Q}$.

As for the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{Q})$, we assume the following 'usual conditions': We assume that $(\mathcal{F}_t)_{t\geqslant 0}$ is right-continuous, i.e. $\bigcap_{s\geqslant t} \mathcal{F}_s = \mathcal{F}_t$, $\forall t \in [0,T]$ and that it is complete, i.e. $\mathcal{F}_0$ contains the negligible sets of the σ-field $\sigma\left(\bigcup_{0\leqslant t\leqslant T} \mathcal{F}_t\right)$. The intuition behind the first condition (right-continuity) is that negligible knowledge is gained during any infinitely small step in time, implying that shocks to the system are free from jumping behaviour but instead unfold gradually over time. As for the second condition (completeness), the intuition is simply that if an event 'definitely' occurs or not occurs, then that fact is known at time 0.

Notice that $(\mathcal{F}_t \otimes \mathcal{I})_{t\in[0,T]}$ is an increasing sequence of σ-fields and is hence a filtration of the σ-field $\mathcal{F} \otimes \mathcal{I}$. We assume that this filtration is right-continuous and complete.

3.2 Required Abatement and the Bank of Commodity Units

For each agent i, we will take as given an adapted cumulated BAU deployment process $(E^i(0,t))_{t\in[0,T]}$, representing for example agent i's use of non-renewable resources, such as oil, natural gas or coal or the 'destructive consumption of clean air' by emission of greenhouse gases.

For example, we may consider a market for emissions allowances which securitise the right to emit one metric tonne of CO_2 or other greenhouse gases, measured in CO_2-equivalents (CO_2e), within the regulated period. In this example, each agent's stock of the commodity (emissions allowances), is subject to a BAU outflow $dE^i(0,t)$ and can be controlled by reducing emission (e.g. by reducing output, changing fuels or through technical improvements) and, additionally, by generation of additional allowances though compatible compliance mechanisms[4].

For other commodities, *deploying* the resource may be interpreted more directly as physical use, such as burning a type of fuel like coal, oil or natural gas; or purchasing units of such resources for later production.

We will model each agent's abatement as an adapted control process $\alpha^i = (\alpha^i_t)_{t\in[0,T]}$, representing the amount by which the instantaneous deployment is reduced ($\alpha^i_t > 0$) or increased ($\alpha^i_t < 0$) at any point in time t, compared to the BAU case. We denote by $dE^i(0,t)$ the agent's instantaneous BAU deployment, in line with the conventional notation for the increment of diffusion processes. Accordingly, $dE^i(0,t) - \alpha^i_t dt$ is agent i's net-deployment at time t which may be positive or negative, as well as larger or smaller than the BAU deployment $dE^i(0,t)$: When, for example we interpret $dE^i(0,t)$ as the BAU emissions of some relevant greenhouse gases, $\alpha^i_t > 0$ corresponds to a net emissions reduction. Conversely, the agent may increase its emissions, $\alpha^i_t < 0$, e.g. by increasing overall output or using cheaper, but more emissions-intensive means of production. Note that by allowing instantaneous abatement to exceed instantaneous BAU deployment, we take into account technologies like carbon-capture and storage or, in case of an ETS, compatible compliance mechanisms allowing agents to be credited additional allowances.

Each agent is characterised, in particular, by its initial bank of commodity units B^i_0 and its cumulated BAU deployment process $(E^i(0,t))_{t\in[0,T]}$. At any point in time t, the agent may receive ψ^i_t units of the resource free of charge, which is henceforth labelled *allocation*. This inflow may, for example, originate from physical sources such as oil wells or natural gas deposits, in which case we ignore the costs of extraction and storage. Here, we assume that the extraction is subject to regulatory control.[5] In emissions trading systems, however, the inflow consists simply of free allowances allocation. We denote agent i's cumulated allocation process by $(A^i(0,t))_{t\in[0,T]}$, where $A^i(0,t) = \int_0^t \psi^i_s ds$.

Note that the allocation process may be subject to an allocation adjustment mechanism such as constructed in Chapter 4. Therein, a *responsive* mechanism will adjust instantaneous allocations in response to changes in the aggregate bank, subject in particular to changes in the aggregate cumulated BAU deployment $E^I(0,t)$. In 2016, the EU ETS was equipped with a similar mechanism, in the form of the so-called Market Stability Reserve (MSR), taking effect in 2020. Note that the term

[4]Under the EU ETS, regulated agents can generate additional units through carbon reduction via overseas projects such as the construction of wind parks.

[5]While such systems are common in cartel structures, here we suggest that equivalent measures may be taken as a course of action in environmental policy.

Market Stability Reserve is occasionally used in the literature to refer to any rules-based mechanism that takes the aggregate bank of unused allowances as an indicator. Where appropriate, we use the term *EC MSR* to refer to the specific implementation brought forth by the European Commission (EC).

With a responsive mechanism in place, agents' strategies will be affected by changes to the time-t expectation $\mathbb{E}_t^{\mathbb{Q}}[A^i(t,T) - E^i(t,T)]$ of *future* cumulated allocations minus BAU deployment. We will see later that the uncertainty in the interrelated processes above is fundamental to the regulated agents' intertemporal optimisation.
Recall that we require that the time-T cumulated BAU deployment $E(0,T)$ be smaller than total abatement plus received allocation and purchase. In this context we implement the convention to use the term *bank* of commodity units, to refer to the cumulated received allocation plus abatement, net the BAU deployment and commodity units sold (Definition 8).

Bank of Commodity Units

Definition 8. *At all points in time t, agent i's* bank *of commodity units is given by*

$$B_t^i = B_0^i + A^i(0,t) - E^i(0,t) + \int_0^t \alpha_s^i \, ds - \int_0^t \beta_s^i \, ds,$$

where α_s^i denotes instantaneous abatement and $|\beta_s^i|$ is the number of commodity units sold ($\beta_s^i > 0$) or bought ($\beta_s^i < 0$) at times $s \in [0,t]$.

The time-t aggregate bank of allowances, i.e. the aggregate volume $B_t^I = \int_{\mathcal{I}} B_t^i d\mathbb{I}(i)$, will be the quantity indicator based on which the regulator will adjust the allocation. Note that this is identical to the quantity indicator used by the EC MSR ([The European Parliament and Council, 2015]), therein referred to as the 'total number of allowances in circulation'.

As a compliance constraint, we now require that $B_T^i = c$ for all agents i, where $c \in \mathbb{R}$ is some constant. In many applications, the constant c may be set to zero; i.e. the bank should neither be negative (which would indicate that total deployment exceeds total abatement plus allocation and purchase) nor should it be positive: When, for example, we interpret T as the time of production launch of the units for which units of the resource have been assigned throughout $[0,T]$, no additional resources should be assigned to that batch of produced units. Similarly, when T is the end of the regulated timeframe in an ETS, allowances are typically only valid throughout $[0,T]$ and loose their entire value after T. Notwithstanding, the regulator may require all agents to hold a fixed positive bank at T, or, equivalently, reduce the initial bank by c.

Note that we assume that the compliance constraint is only binding at time T and that hence we explicitly allow for banking and borrowing during the regulated period.[6] We are willing to accept this limitation in realism concerning some existing commodity markets in order to analyse the policy spectrum as anticipated in the previous chapters.[7]

Given our terminal constraint, the key state variable observed by any agent is its expected (residual) required abatement R_t^i; i.e. the number of units of the commodity, the agent expects (at a given time) to still be required in order to adhere to the terminal constraint $B_T^i = c$ (Definition 9). Accordingly, the terminal constraint is equivalent to the requirement that $R_T^i = 0$.

Expected Residual Required Abatement

Definition 9. *The residual required abatement is given by*

$$R_t^i := \mathbb{E}_t^{\mathbb{Q}}\left[E^i(t,T) - A^i(t,T)\right] - B_t^i + c$$

for all $t \in [0,T]$.

Key to our understanding of the impact any supply adjustment mechanism has on agents' behaviour, is how R_t^i changes in response to (1) shocks to expected future deployment $\mathbb{E}_t^{\mathbb{Q}}[E^i(t,T)]$ and (2) responsive adjustments on behalf of the mechanism through $\mathbb{E}_t^{\mathbb{Q}}[A^i(t,T)]$. Adjustments to the allocation schedule that may alter the cap have both a direct and an indirect effect on $\mathbb{E}_t^{\mathbb{Q}}[A^i(t,T)]$: While the direct impact is self-explanatory, the indirect effect emerges from the fact that any changes made to the allocation schedule impact agents' inter-temporal optimisation. Their behaviour in the remaining time $[t,T]$ is thus altered through the adjustments, which may entail a different expected adjustment scheme later-on. This reveals an important and interesting feature of policy design: How can we solve the model in order to include this feed-back loop between the market and the regulation? We will show that with our model, these dynamics can in fact be solved in closed form, which is of particular interest for the design of an optimal policy.

We can express the expected future residual abatement requirement as $R_t^i = \mathbb{E}_t^{\mathbb{Q}}[Y^i(t,T)]$, where $Y^i(t,T)$ is given by

$$Y^i(t,T) = Y^i(0,T) - \int_0^t \alpha_s^i \, ds + \int_0^t \beta_s^i \, ds,$$

and where agent i's total abatement requirement is given by

$$Y^i(0,T) = E^i(0,T) - A^i(0,T) - B_0^i + c.$$

We note that R^i has the dynamics

$$dR_t^i = d\mathbb{E}_t^{\mathbb{Q}}[Y^i(t,T)] = (\beta_t^i - \alpha_t^i)dt + d\mathbb{E}_t^{\mathbb{Q}}[Y^i(0,T)].$$

[6] Concerning carbon markets, note that under the current EU ETS Directive, borrowing is, only to a limited extend, implicitly possible within a given trading phase due to a given year's allocation starting in advance of the submission date for allowances that cover the preceding year's emissions. In the context of the EU ETS, our model is therefore applicable when either of the following cases holds: (1) The timeframe $[0,T]$ corresponds to one year in the EU ETS setup and we assume a fixed 'hedging requirement' at T for the following years, or (2) the timeframe corresponds to one trading phase of the EU ETS and we abstract from borrowing limitations.

[7] This is particularly relevant since when the bank equals zero and borrowing constraints are binding, the inter-temporal optimisation problem breaks down: Agents simply cover part of their emissions with their allocation and, in aggregate terms, abate the remainder themselves. Hence, the analysis of our policy spectrum is only relevant before the occurrence of such a break-down. We refer to our paper [Kollenberg and Taschini, 2016a] in which we showcase a heuristic approach to transfer our model from a setup without borrowing limitations and a fixed time-horizon to a no-borrowing setup over an infinite time horizon. In the latter setup, the timing of the inter-temporal breakdown becomes endogenous. In the context of the present work, however, this transfer has no immediate value for our analysis of the policy stringency spectrum, which lies in the centre of our attention. Accordingly, we stick with a fixed time-horizon and no borrowing limitations in order to make our analysis tractable and transparent.

Let the random shocks to the time-t expected aggregate total abatement requirement $\mathbb{E}_t[Y^I(0,T)]$ be governed by a driftless diffusion

$$d\mathbb{E}_t^{\mathbb{Q}}[Y^I(0,T)] = \sigma_t^I dW_t^{\mathbb{Q}},$$

where $W^{\mathbb{Q}}$ is a Brownian motion under $\mathbb{Q}$ and where σ_t^i is bounded and deterministic for all i. Furthermore, let $\sigma_t^I = 0$ for $t \in [\theta, T]$, according to our earlier assumption that the uncertainty in BAU deployment is only relevant up to time θ.

The processes $\sigma_t^i, i \in I$ describe how changes in the expected total required abatement are distributed across the set of agents I. We abstract away from specific assumptions on the distribution of σ_t^i across $i \in I$. However, we note that it is reasonable to assume different σ_t^i for different agents, since BAU deployment levels and resource allocations can vary depending on the type of industry or services.

In Chapter 4, we will analyse how σ_t^I is affected by a mechanism that adjusts the resource allocation based on the aggregate bank.

The degree of impact from changes to BAU deployment and allocations can vary across agents. However, all agents are subject to systemic shocks. Hence, we consider the same Brownian motion $W^{\mathbb{Q}}$ for all $i \in I$, whereas differences in market share, technology etc. are represented by the distribution of σ_t^i across I. Accordingly, shocks to $\mathbb{E}_t^{\mathbb{Q}}[Y^i(0,T)]$ are represented by

$$d\mathbb{E}_t^{\mathbb{Q}}[Y^i(0,T)] = \sigma_t^i dW_t^{\mathbb{Q}}.$$

3.3 Instantaneous and Inter-Temporal Optimisation

We assume that the market for the regulated commodity is perfectly competitive. As for a price impact, we assume the slight complication of a possibly non-negligible instantaneous impact on the price paid or received; i.e. the price received per unit sold at any point in time decreases with the volume sold, albeit not affecting the *market price* or the price received for any unit sold at a later date. (And the analogue holds for buying units of the commodity.) This assumption will improve tractability by allowing us to identify a unique optimum in terms of instantaneous optimisation without affecting the generality of our result but rather improving realism.[8] Note that, in particular, the assumption of perfect competition is not affected by the above assumption.

Each Agent's Cost Functional

At each point in time $t \in [0,T]$, every agent $i \in I$ on the market faces an inter-temporal as well as an instantaneous optimisation problem: Each agent has to decide (1) by what amount Δ_t^i it wants to change its stock in the resource and (2) what percentage of Δ_t^i shall be realised by trading and how much shall be realised by abatement. The terminal condition (i.e. the necessity to end up with a time-T residual abatement requirement of $R_T^i = 0$) puts each agent's residual abatement requirement at the heart of that agent's optimisation problem: Given its time-t cumulated BAU deployment $E_t^i(0,t)$ and cumulated allocations $A^i(0,t)$, as well as its expectations about future BAU deployment and allocations, respectively, each agent observes an amount R_t^i, which represents how many units of its requirement needs to be either cut down by abatement or covered by purchase or generation.

We will now introduce our assumptions on the cost components each agent faces: At each point in time $t \in [0,T]$, the agent chooses a level of abatement α_t^i for which the agent incurs costs that are increasing in the absolute value of α_t^i. That is, the more the agent reduces or increases its deployment compared to the BAU case, the higher are the costs borne by the agent. This rather conventional setup reflects the assumption that the agent's Business-as-Usual deployment plan was optimal to begin with and any deviation from this planning increases per-units costs proportionally to the amount of deviation. What's more, it is intuitive that a *quick* change in deployment and, accordingly, a quick change of the production planning is more costly than a slow and progressive change. We make the conventional assumption that the *per-unit (i.e. marginal)* abatement costs increase linearly in abatement. Accordingly, let $Q_t + 2\varrho\alpha_t^i$ be the marginal costs per abated unit, measured in time-t-EUR/unit[9] (e.g. time-t-EUR/tonne of CO_2e). Here, the parameter Q_t is the minimum per-unit cost at each point in time, measured in time-t-EUR/unit. And ϱ is the proportionality by which per-unit costs increase in the amount abated, measured in time-t-EUR/square-unit. The function $\alpha_t^i \mapsto Q_t + 2\varrho\alpha_t^i$ is conventionally called Marginal Abatement Cost Curve (MACC), whereas the image of some amount α_t^i under the MACC is simply called the Marginal Abatement Costs (MAC), for α_t^i. Given some abatement quantity α_t^i, the agent then incurs costs from abatement at the amount of

$$\int_0^{\alpha_t^i} Q_t + 2\varrho a \, da = Q_t\alpha_t^i + \varrho(\alpha_t^i)^2 \quad \text{(in time-}t\text{-EUR).} \tag{3.1}$$

The costs incurred from abating α_t^i units thus increases quadratically in α_t^i.

We make the rather natural assumption that the intercept Q_t increases at the risk-free rate μ. More precisely, we assume that the costs of generating the first (marginal) unit are constant in terms of time-0 Euros. In order to simplify notation, we furthermore assume that the parameter ϱ is constant and identical across all agents.[10]

As for the costs of trading, we denote by P_t the time-t market price for the commodity, as observed by agent i. When selling an amount β_t^i, we assume that the agent again bears *per-unit* costs, which are proportional to the traded position: These additional costs comprise transaction costs from trading smaller or larger positions in a less-than perfectly liquid market. In a highly liquid market, these costs can be rather small (such as simply given by the observed spread between best-bid and best-ask) or rather substantial: When large positions are traded on the market, other participants may take notice of one agent's intentions and adjust their bids accordingly, leading to a large price impact from trading the position. In any case, we limit the complexity of the problem by only considering an *instantaneous* price impact borne *only* by the trading agent; i.e. neither the instantaneous nor future *market prices* are immediately affected by a single agent i trading $|\beta_t^i|$ units but

[8] In the context of emissions trading systems, we refer to [Frino et al., 2010] and [Medina et al., 2014], who document non-negligible transaction costs in the EU ETS.

[9] We adopt the notation of time-t-EUR for the value of one Euro at time 0, compounded by the risk-free interest rate from 0 to t.

[10] We refer to [Landis, 2015] for a detailed discussion on marginal abatement costs of greenhouse gases, and their calibration.

are determined by the distribution of β_t^i across I. More precisely recall that $\mathbb{I}(i) = 0$ for all $i \in I$, which implies

$$\int_{I\setminus\{i\}} \beta_t^j \, d\mathbb{I}(j) = \int_I \beta_t^j \, d\mathbb{I}(j) \quad \text{for all } i \in I \text{ and } t \in [0,T].$$

However, the per-unit profit of selling $\beta_t^i > 0$ units is assumed to be given by $P_t - 2\nu\beta_t^i$ time-t-EUR/unit, where the parameter ν is assumed to be constant and identical across all agents. Conversely, buying $-\beta_t^i > 0$ units incurs per(-negative)-unit costs of $-P_t + 2\nu\beta_t^i$ time-t-EUR/unit. Given a traded amount β_t^i, the agent thus incurs costs (profit when negative) at the amount of

$$\int_0^{\beta_t^i} -P_t + 2\nu b \, db = -P_t\beta_t^i + \nu(\beta_t^i)^2 \quad \text{time-}t\text{-EUR}. \tag{3.2}$$

Given the instantaneous costs of abating and trading in Equations (3.1) and (3.2), respectively, the agent's instantaneous costs in time-0-EUR are thus given by

$$e^{-\mu t} v(Z_t^{i,0,Z_0}, \alpha_t^i, \beta_t^i) = e^{-\mu t}\Big(Q_t\alpha_t^i + \varrho(\alpha_t^i)^2 - P_t\beta_t^i + \nu(\beta_t^i)^2\Big),$$

where $Z^{i,0,Z_0}$ denotes agent i's *state process* starting at time 0, which we will specify in the following. As for the above, recall that μ denotes the risk-free interest rate, which we assume to be constant for the sake of simplifying notation.
The processes R^i, P and Q together comprise agent i's state process as summed up in Definition 10.

Agent i's State Process

Definition 10. *The agent observes the state process $Z^i = (R^i, P, Q)$, where $dQ_t = \mu Q_t dt$,*

$$\begin{aligned} dR_t^i &= (\beta_t^i - \alpha_t^i)dt + d\mathbb{E}_t^{\mathbb{Q}}[Y^i(0,T)] \\ &= (\beta_t^i - \alpha_t^i)dt + \sigma_t^i \, dW_t^{\mathbb{Q}}, \end{aligned}$$

for all $t \in [0,T]$, and where P will be subject to the equilibrium dynamics.

For $(t,z) \in [0,T] \times \mathbb{R}^3$ we want to denote by $Z^{i,t,z} = (Z_s^{i,t,z})_{s\in[t,T]}$ firm i's state process with time-t value $Z_t^{i,t,z} = z$. In the following, we will narrow down and specify this notion, since it will become relevant in the more technical Section 3.5.

Let $W^{\mathbb{Q}} = (W_t^{\mathbb{Q}})_{t\in[0,T]}$ denote a one-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbb{Q})$, and let $(\mathcal{F}_t)_{t\in[0,T]}$ denote the filtration generated by $W^{\mathbb{Q}}$. Based on our modelling approach, we will construct for each $i \in I$ a (system) state process $Z^i = (Z_t^i)_{t\in[0,T]}$, valued in $\mathbb{R}^3$, and governed by

$$dZ_t^i = \Gamma^i(t, Z_t^i, \gamma_t^i)dt + \Sigma^i(t, Z_t^i, \gamma_t^i)dW_t^{\mathbb{Q}}. \tag{3.3}$$

The control processes $\gamma^i = (\gamma_t^i)_{t\in[0,T]}$ are assumed to be progressively measurable with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ and valued in $\mathbb{R}^2$ for all $i \in I$. Note that we can allow the functions Γ^i and Σ^i to depend on time since we are working with a finite time-horizon and hence do not require the stationarity of the optimisation problem that we want to consider. This is natural since we will impose a terminal constraint at time $T < \infty$, and hence the problem should intuitively not be stationary and the i's value function should depend on time.
We will need the functions $\Gamma^i : [0,T] \times \mathbb{R}^3 \times \mathbb{R}^2 \to \mathbb{R}^3$ and $\Sigma^i : [0,T] \times \mathbb{R}^3 \times \mathbb{R}^2 \to \mathbb{R}^3$ to satisfy a uniform Lipschitz condition in their second argument: More precisely, we require that there exists a constant $K \geqslant 0$ such that for all $t \in [0,T]$, all $x, y \in \mathbb{R}^3$, and all $(a,b) \in \mathbb{R}^2$ we have

$$\begin{aligned} |\Gamma^i(t,x,a,b) - \Gamma^i(t,y,a,b)| + |\Sigma^i(t,x,a,b) - \Sigma^i(t,y,a,b)| \\ \leqslant K|x-y|, \end{aligned}$$

where $|\cdot|$ denotes the Euclidian norm. We denote by $\tilde{\mathcal{G}}^i$ the set of control processes γ^i such that

$$\mathbb{E}^{\mathbb{Q}}\left[\int_0^T |\Gamma^i(t,x,\gamma_t^i)|^2 + |\Sigma^i(t,x,\gamma_t^i)|^2 \, dt\right] < \infty,$$

where $x \in \mathbb{R}^3$ can be chosen as an arbitrary value for the state diffusion process due to the above Lipschitz condition. The usual conditions on the filtration $(\mathcal{F}_t)_{t\in[0,T]}$, together with the uniform Lipschitz condition above ensure for all $\gamma^i \in \tilde{\mathcal{G}}^i$ and for any initial condition $(t,z) \in [0,T] \times \mathbb{R}^3$ the existence and uniqueness of a strong solution to the SDE (3.3) on $[t,T]$, with initial value $Z_t^{i,t,z} = z$. (See [Pham, 2009], Section 3.2 in Chapter 3.) We then denote by $Z^{i,t,z} = (Z_s^{i,t,z})_{s\in[t,T]}$ this solution with a.s. continuous paths.

We also note that under these conditions on Γ^i, Σ^i and γ^i, we have that the statements in Proposition 11 hold, which will be quite useful for the proof of our Verification Theorem in Section 3.5.3.

Tractable State Process Properties

Proposition 11. *Under the above conditions we have*

$$\mathbb{E}^{\mathbb{Q}}\left[\sup_{s\in[t,T]} |Z_s^{i,t,z}|^2\right] < \infty,$$

and

$$\lim_{h\searrow 0} \mathbb{E}^{\mathbb{Q}}\left[\sup_{s\in[t,t+h]} |Z_s^{i,t,z} - z|^2\right] = 0.$$

Proof. As these are fairly standard results, we refer to [Krylov, 2008], who provides a proof for more general estimates which yield our claim as a direct consequence. □

We define the sets $\mathcal{G}^i(t,z)$ of admissible strategies for agent i, given the time-t initial state z, as the space of strategies $\gamma^i \in \tilde{\mathcal{G}}^i$ such that time-t expected future costs under the measure $\mathbb{Q}$ are finite (Definition 12).

The Sets $\mathcal{G}^i(t,z)$ of Agent i's Admissible Strategies

Definition 12. *Let the processes P and Q be progressively measurable. We then denote by $\mathcal{G}^i(t,z)$ the set of admissible strategies for agent $i \in I$, given $(t,z) \in [0,T] \times \mathbb{R}^3$; defined as the set of progressively measurable processes $\gamma^i = (\alpha^i, \beta^i) : \Omega \times [0,T] \to \mathbb{R}^2$ in $\tilde{\mathcal{G}}^i$ such that*

$$\mathbb{E}_t^{\mathbb{Q}}\left[\int_t^T e^{-\mu s}\left|v(Z_s^{i,t,z}, \gamma_s^i)\right| \, ds\right] < \infty.$$

In the special case of $t = 0$, we simply write $\mathcal{G}^i$ for $\mathcal{G}^i(0,z)$.

Note that $\mathcal{G}^i(t,z)$ depends on the dynamics of P and Q. However, for the sake of brevity, we omit this dependence in our notation. Furthermore, recall that since $\mathcal{G}^i(t,z)$ is a subset of $\tilde{\mathcal{G}}^i$, any $\gamma^i \in \mathcal{G}^i(t,z)$ satisfies

$$\mathbb{E}^{\mathbb{Q}}\left[\int_0^T |\Gamma^i(t,z,\gamma_t^i)|^2 + |\Sigma^i(t,z,\gamma_t^i)|^2 \, dt\right] < \infty,$$

and hence agent i's system state process SDE has a unique solution, given γ^i and some time-0 initial system state.

The Inter-Temporal Optimisation Problem

Definition 13. *Given a price process $P = (P_s)_{s \in [t,T]}$ and some $(t,z) \in [0,T] \times \mathbb{R}^3$, each agent's dynamic cost minimisation problem during $[t,T]$ is to find*

$$\underset{\gamma^i \in \mathcal{G}^i(t,z):\ R_T^i = 0 \text{ a.s.}}{\text{arginf}} \ \mathbb{E}_t^{\mathbb{Q}}\left[\int_t^T e^{-\mu s} \ v(Z_s^{i,t,z}, \gamma_s^i) \, ds\right].$$

Note that the terminal constraint $R_T^i = 0$ has to be satisfied.

Standard optimal control problems with a finite time-horizon posit the existence of a bequest function which models the final payoff at T. In Section 3.5, however, we will take a different route in order to accommodate the terminal constraint. To this end, we will implement a singular terminal condition, the intuition of which will be that an agent faces an infinite penalty if the terminal condition $R_T^i = 0$ is not satisfied. It will thus suffice to require the agent's strategies to be admissible, which, in particular, imposes the limitation of obtaining finite expected total costs as per Definition 12. Before we provide a treatise of the stochastic equilibrium problem over $[0,T]$, we first consider the deterministic equilibrium problem over the period $[\theta, T]$, where we assume that $\sigma \equiv 0$. To this end, we first make some useful observations.

Let $t \in [0,T]$ and $s \in [t,T]$. Notice that the instantaneous cost function v given by

$$v(Z_s^{i,t,z}, \gamma_s^i) = Q_s\alpha_s^i + \varrho(\alpha_s^i)^2 - P_s\beta_s^i + \nu(\beta_s^i)^2$$

admits to isolate the *instantaneous* optimisation problem in the following sense: Let $\Delta_s^i = \beta_s^i - \alpha_s^i$. We then have

$$dR_s^i = \Delta_s^i ds + d\mathbb{E}_s^{\mathbb{Q}}[Y^i(0,T)].$$

Thus, agent i's control on R^i is determined only by the difference $\Delta^i = \beta^i - \alpha^i$, i.e. by the difference between sold and abated commodity units at each point in time, whereas the proportions between α^i and β^i have no influence on the rate of change in i's residual abatement requirement.

We can thus express the agent's inter-temporal optimisation problem as trying to obtain a process Δ^i that ultimately minimises expected total costs, given by

$$\mathbb{E}_t^{\mathbb{Q}}\left[\int_t^T \inf_{(a,b)\in G(\Delta_s^i)} e^{-\mu s} v(Z_s^{i,t,z}, a, b) \, ds\right],$$

where $G(\Delta_s^i) = \{(a,b) \in \mathbb{R}^2 : \ b - a = \Delta_s^i\}$ for $s \in [t,T]$.

We use the term *instantaneous optimisation problem* to describe each agent's problem of finding (α_s^i, β_s^i) such that

$$(\alpha_s^i, \beta_s^i) = \underset{(a,b)\in G(\Delta_s^i)}{\text{arginf}} \ e^{-\mu s} v(Z_s^{i,t,z}, a, b), \tag{3.4}$$

where Δ_s^i is *given*, at each point in time. The particular choice for the instantaneous cost function v makes solving the above minimisation problem rather tractable. In fact, the first- and second-order conditions associated to Equation (3.4) immediately yield

$$\alpha_s^i = \frac{P_s - Q_s - 2\nu\Delta_s^i}{2(\nu + \varrho)}, \quad \text{and} \quad \beta_s^i = \frac{P_s - Q_s + 2\varrho\Delta_s^i}{2(\nu + \varrho)}. \tag{3.5}$$

In this context, we define the sets $\mathcal{D}^i(t,z)$ associated to $\mathcal{G}^i(t,z)$ (Definition 14).

The Sets $\mathcal{D}^i(t,z)$

Definition 14. *For each $(t,z) \in [0,T] \times \mathbb{R}^3$ and each $i \in I$, we let $\mathcal{D}^i(t,z)$ denote the set of progressively measurable processes $\Delta^i = (\Delta_s^i)_{s\in[0,T]}$ for which all $\gamma^i = (\alpha^i, \beta^i) \in \tilde{\mathcal{G}}^i$, satisfying Equation* (3.5) *for all $s \in [t,T]$ are elements of $\mathcal{G}^i(t,z)$.*

3.4 Equilibrium in the Deterministic Case

In the following Theorem 17, we will solve for a strong equilibrium[11] in the deterministic case. More precisely, we will only consider the interval $[\theta, T]$ during which $\sigma \equiv 0$ and obtain an equilibrium strategy bundle $(\gamma_s^*)_{s\in[\theta,T]}$, and the associated equilibrium price process $(P_s^*)_{s\in[\theta,T]}$. To this end, we will first make some useful observations in Lemmas 15 and 16.

The Inter-Temporal Optimisation Problem Rephrased

Lemma 15. *Agent i's inter-temporal optimisation problem during $[\theta, T]$ in Definition 13 can be equivalently expressed as finding*

$$\Delta^{i,*} = \underset{\Delta^i \in \mathcal{D}^i(\theta, Z_\theta^{i,t,z}):\int_\theta^T \Delta_s^i ds = -R_\theta^i}{\arg\inf} \int_\theta^T e^{-\mu s} \left(\Delta_s^i - \left(\frac{Q_s}{2\varrho} + \frac{P_s}{2\nu} \right) \right)^2 ds.$$

Proof. Since $\beta_s^i = \alpha_s^i + \Delta_s^i$, we obtain from instantaneous optimisation, according to Equations (3.4) and (3.5):

$$\begin{aligned}
v\left(Z_s^{i,t,z}, \gamma_s(\Delta_s^i)\right) &= Q_s\alpha_s^i + \varrho(\alpha_s^i)^2 - P_s\beta_s^i + \nu(\beta_s^i)^2 \\
&= Q_s\alpha_s^i + \varrho(\alpha_s^i)^2 - P_s\alpha_s^i - P_s\Delta_s^i + \nu(\alpha_s^i)^2 + 2\nu\alpha_s^i\Delta_s^i + \nu(\Delta_s^i)^2 \\
&= (\nu+\varrho)(\alpha_s^i)^2 - (P_s - Q_s - 2\nu\Delta_s^i)\alpha_s^i - P_s\Delta_s^i + \nu(\Delta_s^i)^2 \\
&= -\,(\nu+\varrho)\,(\alpha_s^i)^2 - P_s\Delta_s^i + \nu(\Delta_s^i)^2 \\
&= -\frac{(P_s-Q_s)^2 - 4\nu\Delta_s^i(P_s-Q_s) + 4\nu^2(\Delta_s^i)^2 + 4(\nu+\varrho)P_s\Delta_s^i - 4\nu(\nu+\varrho)(\Delta_s^i)^2}{4(\nu+\varrho)} \\
&= -\frac{(P_s-Q_s)^2 + 4\nu\Delta_s^iQ_s + 4\varrho P_s\Delta_s^i - 4\nu\varrho(\Delta_s^i)^2}{4(\nu+\varrho)} \\
&= -\frac{(P_s-Q_s)^2 + 4\Delta_s^i(\nu Q_s + \varrho P_s) - 4\nu\varrho(\Delta_s^i)^2}{4(\nu+\varrho)} \\
&= \frac{\nu\varrho}{\nu+\varrho}\left(\Delta_s^i - \left(\frac{Q_s}{2\varrho} + \frac{P_s}{2\nu}\right)\right)^2 - \frac{\nu\varrho}{\nu+\varrho}\left(\frac{Q_s}{2\varrho} + \frac{P_s}{2\nu}\right)^2 - \frac{(P_s-Q_s)^2}{4(\nu+\varrho)}
\end{aligned} \tag{3.6}$$

Notice that only the first summand on the right-hand side in the above equation depends on Δ_s^i, which immediately proves our claim. ☐

The Hilbert Spaces $\mathcal{H}(t,z)$

Lemma 16. *Let $(t,z) \in [0,T] \times \mathbb{R}^3$. If non-empty, the space $\mathcal{H}(\theta, Z_\theta^{i,t,z})$ of processes $D = (D_s)_{s\in[\theta,T]}$ such that*

$$\int_\theta^T e^{-\mu s}\, D_s^2\, ds < \infty,$$

equipped with the inner product $\langle \cdot, \cdot \rangle$ defined by

$$\langle A, D \rangle = \int_\theta^T e^{-\mu s}\, A_s\, D_s\, ds$$

is a Hilbert space; i.e. it is a complete vector space with respect to the metric induced by the inner product defined above.

Proof. Recall that $\mathbb{L}^{[\theta,T]}$ denotes the Lebesgue-measure on $([\theta,T], \mathcal{B}([\theta,T]))$. The function $s \mapsto e^{-\mu s}$ is positive and finite on $[\theta,T]$. Therefore, there exists a measure $\mathbb{L}_\mu^{[\theta,T]}$ on $([\theta,T], \mathcal{B}([\theta,T]))$, unique up to null-sets, with Radon-Nikodým derivative with respect to $\mathbb{L}^{[\theta,T]}$ given by $d\mathbb{L}_\mu^{[\theta,T]}/d\mathbb{L}^{[\theta,T]}(s) = e^{-\mu s}$, $s \in [\theta,T]$. We find that $\mathcal{H}(\theta, Z_\theta^{i,t,z})$ is simply the L^2-space defined with respect to the this measure. By the Theorem of Riesz-Fischer, $\mathcal{H}(\theta, Z_\theta^{i,t,z})$ is thus a Hilbert space. ☐

[11] By the nature of the deterministic case, the notion of a *strong equilibrium* is independent of the measure being applied in this context.

Equilibrium in the Deterministic Case

Theorem 17. *Let $\sigma_t = 0$ for $t \in [\theta, T]$, where $\theta \in [0, T)$ and denote $z_\theta = Z_\theta^{i,t,z} \in \mathbb{R}^3$. For all $j \in I$, let agent j's strategy be given by $\gamma_t^j = (\alpha_t^j, \beta_t^j)$, for $t \in [\theta, T]$, where the components of γ_t^j are given by*

$$\alpha_t^j = \frac{P_t - Q_t - 2\nu\Delta_t^j}{2(\nu + \varrho)}, \quad \text{and} \quad \beta_t^j = \frac{P_t - Q_t + 2\varrho\Delta_t^j}{2(\nu + \varrho)}, \tag{3.7}$$

and where

$$\Delta_t^j = -\frac{\mu e^{\mu t}}{e^{\mu T} - e^{\mu\theta}} R_\theta. \tag{3.8}$$

Then $\gamma = (\gamma^j)_{j \in I}$ is a strong equilibrium strategy bundle on $[\theta, T]$ with respect to any probability measure on $(\Omega, \mathcal{F})$. The equilibrium price process P induced by γ is given on $[\theta, T]$ by $P_t = P_\theta e^{\mu(t-\theta)}$ for all $t \in [\theta, T]$ and we have

$$P_t = Q_t - 2\varrho\Delta_t^I = Q_t + \frac{2\varrho\mu e^{\mu t}}{e^{\mu T} - e^{\mu\theta}} R_\theta^I. \tag{3.9}$$

In particular, we obtain the intuitive result that the time-t equilibrium price equals the time-t aggregate *marginal abatement costs*

$$P_t = Q_t + 2\varrho\alpha_t^I.$$

Proof. We are going to prove our assertion by means of the following pattern: First, we consider some candidate strategy for all agents in I. Second, we fix some arbitrary $i \in I$ and use the fact that the market clearing price P is fully determined by the strategies of all $j \in I\backslash\{i\}$. We then solve agent i's optimisation problem, given P and show that the resulting strategy coincides with our candidate. Since i was arbitrary in I, this will show that our candidate is valid for all $i \in I$ and that hence it is indeed an equilibrium strategy bundle with induced price process P.

Let $t \in [\theta, T]$ and fix some $i \in I$. For all $j \in I\backslash\{i\}$, let those agents' strategies be given according to Equations (3.7) and (3.8). We then obtain from Equation (3.7), together with the market clearing condition $\int_I \beta_t^i d\mathbb{I}(i) = 0$, $\forall t \in [\theta, T]$, along with the fact that $\mathbb{I}(i) = 0$, that the time-t market clearing price is given by

$$P_t = Q_t - 2\varrho \int_I \Delta_t^j \, d\mathbb{I}(j) = Q_t - 2\varrho \int_{I\backslash\{i\}} \Delta_t^j \, d\mathbb{I}(j).$$

This yields, together with Equation (3.8), and that

$$P_t = Q_t + \frac{2\varrho\mu e^{\mu t}}{e^{\mu T} - e^{\mu\theta}} \int_{I\backslash\{i\}} R_\theta^j \, d\mathbb{I}(j) = Q_t + \frac{2\varrho\mu e^{\mu t}}{e^{\mu T} - e^{\mu\theta}} R_\theta^I,$$

where, again, we used that $\mathbb{I}(i) = 0$. Since $Q_t = Q_\theta e^{\mu(t-\theta)}$, we thus obtain $P_t = P_\theta e^{\mu(t-\theta)}$.

For our fixed agent $i \in I$, let $\Delta_t^i = \tilde{\Delta}_t^i + \chi_t^i$, where

$$\tilde{\Delta}_t^i = -\frac{\mu e^{\mu t}}{e^{\mu T} - e^{\mu\theta}} R_\theta^i.$$

The (sought-after) process Δ^i determines agent i's inter-temporal behaviour in the sense that $\beta_t^i - \alpha_t^i = \Delta_t^i$ for all $t \in [\theta, T]$ and hence

$$R_t^i = R_\theta^i + \int_\theta^t \Delta_s^i \, ds.$$

We will prove that $\tilde{\Delta}^i$ minimises agent i's costs by showing that Δ^i is cost-minimising if and only if $\chi^i \equiv 0$. To this end, first notice that $\int_\theta^T \tilde{\Delta}_t^i \, dt = -R_\theta^i$. Hence we have $\int_\theta^T \Delta_t^i \, dt = -R_\theta^i$ if and only if $\int_\theta^T \chi_t^i \, dt = 0$. We know that instantaneous optimisation commands that at time t, abatement and trading must be given by $\gamma_t^i(\Delta_t^i) = \left(\alpha_t^i(\Delta_t^i), \beta_t^i(\Delta_t^i)\right)$, where

$$\alpha_t^i(\Delta_t^i) = \frac{P_t - Q_t - 2\nu\Delta_t^i}{2(\nu + \varrho)}, \quad \text{and} \quad \beta_t^i(\Delta_t^i) = \frac{P_t - Q_t + 2\varrho\Delta_t^i}{2(\nu + \varrho)}.$$

Agent i's inter-temporal minimisation problem between θ and T is then to find

$$\Delta^{i,*}\big|_{[\theta,T]} = \underset{\Delta^i \in \mathcal{D}^i(\theta, z_\theta):\, \int_\theta^T \Delta_t^i dt = -R_\theta^i}{\operatorname{arginf}} \int_\theta^T e^{-\mu t} \, v\left(Z_t^{i,\theta,z}, \gamma_t^i(\Delta_t^i)\right) \, dt.$$

Since P and Q simply increase deterministically at the risk-free rate, we notice that by Lemma 15, the space $\mathcal{D}^i(\theta, z_\theta)$ is identical to the Hilbert space $\mathcal{H}(\theta, z_\theta)$: This is because $P, Q \in \mathcal{H}(\theta, z_\theta)$ and for any $A \in \mathcal{H}(\theta, z_\theta)$ we have $D \in \mathcal{H}(\theta, z_\theta)$ if and only if $D - A \in \mathcal{H}(\theta, z_\theta)$. Let $\mathcal{C}(\theta, z_\theta) \subseteq \mathcal{H}(\theta, z_\theta)$ denote the set

$$\mathcal{C}(\theta, z_\theta) = \left\{ \zeta \in \mathcal{H}(\theta, z_\theta) : \int_\theta^T \zeta_t \, dt = 0 \right\}.$$

Our minimisation problem is then equivalent to finding a process $\chi^{i,*} \in \mathcal{C}(\theta, z_\theta)$ such that

$$\chi^{i,*} = \operatorname*{arginf}_{\chi^i \in \mathcal{C}(\theta, z_\theta)} \int_t^T e^{-\mu t} \, v\left(Z_t^{i,\theta,z}, \gamma_t^i(\tilde{\Delta}_t^i + \chi_t^i) \right) \, dt.$$

By Lemma 15, we know that the minimisation problem above is equivalent to finding $\chi^{i,*}$ such that

$$\chi^{i,*} = \operatorname*{arginf}_{\chi^i \in \mathcal{C}(\theta, z_\theta)} \int_\theta^T e^{-\mu t} \, \left(\chi_t^i - \left(\frac{Q_t}{2\varrho} + \frac{P_t}{2\nu} - \tilde{\Delta}_t^i \right) \right)^2 dt.$$

This is equivalent to

$$\chi^{i,*} = \operatorname*{arginf}_{\chi^i \in \mathcal{C}(\theta, z_\theta)} \|\chi^i - \phi^i\|, \quad \text{where} \quad \phi^i = \frac{Q}{2\varrho} + \frac{P}{2\nu} - \tilde{\Delta}^i, \tag{3.10}$$

and where the $\|\cdot\|$ denotes the norm induced by the inner product $\langle \cdot, \cdot \rangle$, as defined as in Lemma 16. Notice that $\mathcal{C}(\theta, z_\theta)$ is a closed subspace of the Hilbert space $\mathcal{H}(\theta, z_\theta)$, since $\mathcal{C}(\theta, z_\theta)$ is the preimage of $\{0\} \subset \mathbb{R}$ under the function $\zeta \mapsto \int_\theta^T \zeta_t dt$, the continuity of which can be shown as follows: Let $\eta = (e^{\mu t})_{t \in [\theta, T]}$ and let $\zeta^n \in \mathcal{H}(\theta, z_\theta)$, $n \in \mathbb{N}$, converge to $\zeta \in \mathcal{H}(\theta, z_\theta)$ in the $\mathcal{H}(\theta, z_\theta)$-sense, for $n \to \infty$. We then have by the Cauchy-Schwarz Inequality that

$$\left| \int_\theta^T \zeta_t \, dt - \int_\theta^T \zeta_t^n \, dt \right| = \left| \int_\theta^T e^{-\mu t} \, e^{\mu t} \, (\zeta_t - \zeta_t^n) \, dt \right| = |\langle \eta, \zeta - \zeta^n \rangle| \leq \|\eta\| \, \|\zeta - \zeta^n\| \to 0.$$

We thus find by Hilbert's Projection Theorem, that a necessary and sufficient condition for $\chi^{i,*}$ to attain the infimum in Equation 3.10 is that $\chi^{i,*} - \phi^i \perp \mathcal{C}(\theta, z_\theta)$ with respect to $\langle \cdot, \cdot \rangle$; i.e.

$$0 = \langle \chi^{i,*} - \phi^i, \zeta \rangle = \langle \chi^{i,*}, \zeta \rangle - \langle \phi^i, \zeta \rangle, \quad \forall \zeta \in \mathcal{C}(\theta, z_\theta). \tag{3.11}$$

For the rightmost summand on the right-hand side of Equation (3.11), note that all three of the processes P, Q, and $\tilde{\Delta}^i$ increase deterministically at rate μ on $[\theta, T]$. Therefore, we have that $\phi_t^i = \phi_\theta^i e^{\mu(t-\theta)}$ and hence

$$\langle \phi^i, \zeta \rangle = \int_\theta^T \phi_\theta e^{\mu(t-\theta)} \, e^{-\mu t} \, \zeta_t \, dt = \phi_\theta^i e^{-\mu\theta} \, \int_\theta^T \zeta_t \, dt = 0, \tag{3.12}$$

for all $\zeta \in \mathcal{C}(\theta, z_\theta)$. Thus, the left-hand summand on the right-hand side of Equation (3.11) must vanish; i.e.

$$\langle \chi^{i,*}, \zeta \rangle = 0,$$

for any $\zeta \in \mathcal{C}(\theta, z_\theta)$. In particular, this must hold for $\zeta = \chi^{i,*}$ which immediately implies $\|\chi^{i,*}\| = 0$ and thus $\chi^{i,*} \equiv 0$. In summary, we thus obtain that $\chi^{i,*} - \phi \perp \mathcal{C}(\theta, z_\theta)$ if and only if $\chi^{i,*} \equiv 0$. This proves that γ is indeed a strong equilibrium strategy bundle in on $[\theta, T]$. We note that we have by the market clearing condition $\beta_t^I = 0$ for all $t \in [\theta, T]$, that

$$P_t = Q_t - 2\varrho\Delta_t^j = Q_t - 2\varrho(\beta_t^I - \alpha_t^I) = Q_t + 2\varrho\alpha_t^I.$$

This corresponds to the intuitive result that the market clearing price should equal the *aggregate* marginal abatement costs at any time $t \in [\theta, T]$. □

We will devote the following Section 3.5 to the problem of finding a *stochastic* $\mathbb{Q}$-strong equilibrium strategy bundle by means of techniques of stochastic optimal control. To this end, we will first sum up some useful technical preliminaries in Subsection 3.5.1, before we solve each agent i's optimisation problem under $\mathbb{Q}$ in Subsections 3.5.2 and 3.5.3 . In Section 3.7 we will then show that the thusly obtained strategy bundle and price process also constitute a $\mathbb{P}$-weak equilibrium, in line with Proposition 7. In the subsequent Chapter 4, we will apply our equilibrium results to questions on regulatory mechanism design in the context of climate policy.

3.5 Strong Equilibrium under a Risk-Neutral Measure

We will now derive a strong equilibrium with respect to the measure $\mathbb{Q}$. In Subsection 3.7 we will then use the equilibrium dynamics obtained in the following, to present sufficient conditions under which $\mathbb{Q} \sim \mathbb{P}$ and hence the $\mathbb{Q}$-strong equilibrium is also a weak equilibrium under the objective measure $\mathbb{P}$. In order to solve each agent's optimisation problem, we will first introduce some useful technical preliminaries from the theory of optimal stochastic control.

3.5.1 Technical Preliminaries: Optimal Stochastic Control

Our treatise of technical preliminaries in this Subsection 3.5.1 follows the presentation in [Pham, 2009], Chapter 3, the notation in which we amend to fit our purposes and which we augment with explanatory details and economic intuition. We first define the key notions necessary for our treatise. For each agent i, these are the cost function (Definition 18), which simply captures the expected future costs induced by a given strategy; and the associated value function (Definition 19) which captures how the infimum of future expected costs responds to instantaneous changes in the agent's state process.

Agent i's Cost Functional

Definition 18. *We define the agents' cost functional* $J : [0, T] \times \mathbb{R}^3 \times \mathbb{R}^2$ *by*

$$J(t, z, \gamma^i) = \mathbb{E}_t^{\mathbb{Q}} \left[\int_t^T e^{-\mu s} \, v(Z_s^{i,t,z}, \gamma_s^i) \, ds \right],$$

for all $(t, z) \in [0, T] \times \mathbb{R}^3$ *and all* $\gamma^i \in \mathcal{G}^i(t, z)$.

Agent i's Value Function

Definition 19. *We define the value function associated to each agent i's optimisation problem by*

$$w^{i,*}(t, z) = \inf_{\gamma^i \in \mathcal{G}^i(t,z)} J(t, z, \gamma^i).$$

for all $(t, z) \in [0, T] \times \mathbb{R}^3$.

The Dynamic Programming Principle In the following we will lay out ready a useful set of tools that we will use to solve the problem of optimal stochastic control for each agent. Namely, these are the dynamic programming principle (DPP) and the Hamilton-Jacobi-Bellman (HJB) equation. Similar to the discrete-time case, the solutions to continuous-time optimal stochastic control problems can often be found based on the following Theorem 20 (cf. [Pham, 2009], Remark 3.3.3.), a proof of which we will provide further below

The Dynamic Programming Principle (DPP)

Theorem 20. *Let* $t \in [0, T]$. *For all* $\gamma^i \in \mathcal{G}^i(t, z)$ *and all stopping times* $\tau \in [t, T]$ *we have*

$$w^{i,*}(t, z) \leqslant \mathbb{E}_t^{\mathbb{Q}} \left[\int_t^\tau e^{-\mu s} v(Z_s^{i,t,z}, \gamma^i) \, ds + w^{i,*}(\tau, Z_\tau^{i,t,z}) \right]. \tag{3.13}$$

Furthermore, for all $\epsilon > 0$, *there exists a strategy* $\gamma^i \in \mathcal{G}^i(t, z)$ *such that for all stopping times* $\tau \in [t, T]$ *we have*

$$w^{i,*}(t, z) + \epsilon \geqslant \mathbb{E}_t^{\mathbb{Q}} \left[\int_t^\tau e^{-\mu s} v(Z_s^{i,t,z}, \gamma^i) \, ds + w^{i,*}(\tau, Z_\tau^{i,t,z}) \right]. \tag{3.14}$$

The natural interpretation of Equations (3.13) and (3.14) is that the optimisation can be 'split in two' with respect to the time-dimension at any random point $\tau \in [t, T]$. That is, in analogy to dynamic programming in discrete time, where we can start the optimisation at T using some terminal condition, and then work our way back towards $t = 0$, either algorithmically or in closed-form, here we have a similar formulation in continuous time: Once we know the optimal outcome over the remaining timeframe $[\tau, T]$ (i.e. $w^{i,*}(\tau, Z_\tau^{i,t,z})$), we have that, on the one hand, any control implemented over $[t, \tau]$ will at best yield the optimal outcome over the entire period (Equation (3.13)). On the other hand, such strategies can come arbitrarily close to that optimal outcome (Equation (3.14)). Note that we have an equivalent formulation in the form of the following Corollary 21 (cf. [Pham, 2009], Theorem 3.3.1).

Alternative Formulation of the DPP

Corollary 21. *For $(t, z) \in [0, T] \times \mathbb{R}^3$ we have*

$$w^{i,*}(t,z) = \inf_{\gamma^i \in \mathcal{G}^i(t,z)} \sup_{\tau \in [t,T]} \mathbb{E}_t^{\mathbb{Q}} \left[\int_t^\tau e^{-\mu s} v(Z_s^{i,t,z}, \gamma_s^i)\, ds + w^{i,*}(\tau, Z_\tau^{i,t,z}) \right]$$
$$= \inf_{\gamma^i \in \mathcal{G}^i(t,z)} \inf_{\tau \in [t,T]} \mathbb{E}_t^{\mathbb{Q}} \left[\int_t^\tau e^{-\mu s} v(Z_s^{i,t,z}, \gamma_s^i)\, ds + w^{i,*}(\tau, Z_\tau^{i,t,z}) \right]. \tag{3.15}$$

Even though the formulation in Corollary 21 is somewhat more compact, the form presented in Theorem 20 is quite useful for the derivation of the Hamilton-Jacobi-Bellman equation as we shall see in the following. In the interest of completeness, we will first provide a proof of Theorem 20, which, as a by-product yields a proof of Corollary 21, and the formulation provided in Equation (3.15).

Proof of Theorem 20

Proof of Theorem 20. (i) We first prove Equation (3.13). We fix a control $\gamma^i \in \mathcal{G}^i(t, z)$. The solution to the dynamics

$$dZ_t^i = \Gamma^i(t, Z_t^i, \gamma_t^i)dt + \Sigma^i(t, Z_t^i, \gamma_t^i)dW_t^{\mathbb{Q}} \tag{3.16}$$

with fixed initial value Z_0^i is pathwise unique; i.e. any two of its solutions, $\hat{Z}$ and $\tilde{Z}$ satisfy $\mathbb{Q}[\hat{Z}_s = \tilde{Z}_s : s \in [0, T]] = 1$. This implies that for any stopping time $\tau \in [t, T]$ we have

$$Z_s^{i,t,z} = Z_s^{i,\tau,Z_\tau^{i,t,z}} : s \in [\tau, T].$$

That is, if we consider the solution $(Z_s^{i,t,z})_{s \in [t,T]}$ to the SDE (3.16) over $[t, T]$ starting at z, and observe its value at time τ, i.e. $Z_\tau^{i,t,z}$, then the process solving (3.16) over $[\tau, T]$ with fixed initial value $Z_\tau^{i,t,z}$ at time τ, is (a.s.) pathwise identical to the original solution $(Z_s^{i,t,z})_{s \in [t,T]}$. This Markovian structure yields by the law of iterated conditional expectation ('tower property') that

$$J(t, z, \gamma_s^i) = \mathbb{E}_t^{\mathbb{Q}} \left[\int_t^T e^{-\mu s} v(Z_s^{i,t,z}, \gamma_s^i)\, ds \right]$$
$$= \mathbb{E}_t^{\mathbb{Q}} \left[\int_t^\tau e^{-\mu s} v(Z_s^{i,t,z}, \gamma_s^i)\, ds + \int_\tau^T e^{-\mu s} v(Z_s^{i,t,z}, \gamma_s^i)\, ds \right]$$
$$= \mathbb{E}_t^{\mathbb{Q}} \left[\int_t^\tau e^{-\mu s} v(Z_s^{i,t,z}, \gamma_s^i)\, ds + \mathbb{E}_\tau^{\mathbb{Q}} \left[\int_\tau^T e^{-\mu s} v(Z_s^{i,t,z}, \gamma_s^i)\, ds \right] \right]$$
$$= \mathbb{E}_t^{\mathbb{Q}} \left[\int_t^\tau e^{-\mu s} v(Z_s^{i,t,z}, \gamma_s^i)\, ds + J(\tau, Z_\tau^{i,t,z}, \gamma_s^i) \right].$$

Since by definition of $w^{i,*}$ we have $J(t, z, \gamma^i) \geqslant w^{i,*}(t, z)$ for all $t, z \in [0, T] \times \mathbb{R}^3$ and since τ is arbitrary in $[t, T]$ it follows that

$$J(t, z, \gamma^i) \leqslant \sup_{\tau \in [t,T]} \mathbb{E}_t^{\mathbb{Q}} \left[\int_t^\tau e^{-\mu s} v(Z_s^{i,t,z}, \gamma^i)\, ds + w^{i,*}(\tau, Z_\tau^{i,t,z}) \right]$$

Taking the infimum over all admissible controls $\gamma^i \in \mathcal{G}^i(t, z)$ on the left-hand side of the above inequality, we obtain

$$w^{i,*}(t,z) \leqslant \inf_{\gamma^i \in \mathcal{G}^i(t,z)} \mathbb{E}_t^{\mathbb{Q}} \left[\int_t^\tau e^{-\mu s} v(Z_s^{i,t,z}, \gamma^i)\, ds + w^{i,*}(\tau, Z_\tau^{i,t,z}) \right],$$

which is the desired Equation (3.13).

(ii) We now want to prove Equation (3.14). To this end we fix some arbitrary control $\gamma^i \in \mathcal{G}^i(t, z)$ and some arbitrary stopping time $\tau \in [t, T]$. By definition, $w^{i,*}$ is the infimum of the cost functional over the set of admissible controls. Hence for any deviation $\epsilon > 0$ from that infimum and any $\omega \in \Omega$, there must exists a control $\gamma^{i,\epsilon}$ which improves upon that deviation in the sense that

$$J\left(\tau, Z_\tau^{i,t,z}, \gamma^{i,\epsilon}\right) \leqslant w^{i,*}\left(\tau, Z_\tau^{i,t,z}\right) + \epsilon. \tag{3.17}$$

We now define the process $\hat{\gamma}^i$ by

$$\hat{\gamma}_s^i(\omega) = \begin{cases} \gamma_s^i(\omega) : 0 \leqslant s \leqslant \tau(\omega) \\ \gamma_s^{i,\epsilon,\omega}(\omega) : \tau(\omega) \leqslant s \leqslant T \end{cases} \quad \text{for } s \in [t, T],\ \omega \in \Omega.$$

The process $\hat{\gamma}^i$ can be shown to be progressively measurable (see [Bertsekas and Shreve, 1978], Chapter 7) and thus is an element of $\mathcal{G}^i(t,z)$. From Equation (3.17), and by the law of iterated conditional expectation we obtain

$$\begin{aligned}\mathbb{E}_t^Q\left[\int_t^\tau e^{-\mu s}\, v(Z_s^{i,t,z},\gamma_s^i)\, ds + w^{i,*}(\tau, Z_\tau^{i,t,z})\right] \leqslant & \mathbb{E}_t^Q\left[\int_t^\tau e^{-\mu s}\, v(Z_s^{i,t,z},\gamma_s^i)\, ds + J(\tau, Z_\tau^{i,t,z},\gamma^{i,\epsilon})\right] \\ \leqslant & \mathbb{E}_t^Q\left[\int_t^\tau e^{-\mu s}\, v(Z_s^{i,t,z},\gamma_s^i)\, ds + w^{i,*}(\tau, Z_\tau^{i,t,z})\right] + \epsilon\end{aligned}$$

Since $\gamma^i \in \mathcal{G}^i(t,z)$ and $\tau \in [t,T]$ are arbitrary (in particular during $[t,\tau]$), we obtain the desired inequality (3.14), as well as Corollary 21.

□

Note that the statement in Theorem 20 and the equivalent formulation provided by Equation 3.15 is a stronger formulation of the statement that $w^{i,*}(t,z)$ equals the expression

$$\inf_{\gamma\in\mathcal{G}^i(t,z)} \mathbb{E}_t^Q\left[\int_t^\tau e^{-\mu s} v(s, Z_s^{i,t,z},\gamma^i)\, ds + w^{i,*}(\tau, Z_\tau^{i,t,z})\right], \quad (3.18)$$

for all stopping times $\tau \in [t,T]$, which is a common formulation of the Dynamic Programming Principle.

The Hamilton-Jacobi-Bellman Equation The Dynamic Programming Principle can be used to examine the local behaviour of the value function at t^+ for any time $t \in [0,T]$: More precisely, by letting the stopping time τ converge to t, we can examine the increment in $w^{i,*}$ for an infinitesimally small increment in time. We will use the DPP in the form of Equations (3.13) and (3.14) in order to formally derive the so-called Hamilton-Jacobi-Bellman (HJB) partial differential equation, which characterises the local behaviour of the value function at t^+. As is customary, we will later use the HJB equation to find a closed-form representation of the value function (by solving the HJB Equation associated to our equilibrium problem explicitly) and, as a by-product, obtain an optimal stochastic control $\gamma^{i,*}$ for agent i. In the following, will heuristically derive that the value function should satisfy what is called the *Hamilton-Jacobi-Bellman* differential equation, which is a partial differential equation of the form

$$0 = D_t w(t,z) + \mathcal{L}^{(a,b)} w(t,z) + e^{-\mu s} v(s,z,a,b),$$

where w denotes a sufficiently smooth function. The functional $\mathcal{L}^{(a,b)}$ will be specified in the following.

Note that throughout the text, we use the notation D_z, to denote the differential operator with respect to z, and accordingly for other parameters.

Heuristic Derivation of the Hamilton-Jacobi-Bellman Equation

(1) We will first use Equation (3.13) to show that $w^{i,*}$ should satisfy a condition of the form

$$0 \leqslant D_t w^{i,*}(t,z) + \mathcal{L}^{(a,b)} w^{i,*}(t,z) + e^{-\mu s} v(s,z,a,b).$$

To this end, consider some stopping time τ valued in $[t,T]$, and a constant control $\gamma^i \equiv (a,b) \in \mathbb{R}^2$. We then have by Equation (3.13) of the DPP:

$$w^{i,*}(t,z) \leqslant \mathbb{E}_t^Q\left[\int_t^\tau e^{-\mu s} v(Z_s^{i,t,z},\gamma_s^i)\, ds + w^{i,*}(\tau, Z_\tau^{i,t,z})\right]. \quad (3.19)$$

If $w^{i,*}$ is smooth enough, i.e. if $w^{i,*}: [0,T]\times\mathbb{R}^3 \to \mathbb{R}$ is at least once continuously differentiable in the first and at least twice continuously differentiable in the second argument, we can apply Itō's Lemma (in integral form). Recalling that Γ and Σ denote the drift and volatility term in the dynamics of Z^i we obtain

$$w^{i,*}(\tau, Z_\tau^{i,t,z}) = w^{i,*}(t,z) + \int_t^\tau \left(D_t w^{i,*} + \mathcal{L}^{(a,b)}\right)(s, Z_s^{i,t,z})\, ds + \xi_\tau, \quad (3.20)$$

where ξ is a local martingale starting in zero at time t and where $\mathcal{L}^{(a,b)}$ is the infinitesimal generator associated to (a,b) and the dynamics dZ^i of Z^i, defined as

$$\mathcal{L}^{(a,b)} w^{i,*} = \Gamma(t,z,a,b).D_z w^{i,*} + \frac{1}{2}\mathrm{tr}(\Sigma(t,z,a,b)\Sigma'(t,z,a,b) D_z^2 w^{i,*}).$$

Here, the period operator '.' denotes the vector product on $\mathbb{R}^3$, $\mathrm{tr}(\cdot)$ denotes the trace operator on the space of 3×3-matrices and Σ' denotes the transposition of Σ. The infinitesimal generator $\mathcal{L}^{(a,b)}$ associated to (a,b) and dZ^i has a straightforward economic interpretation: First, consider the left-hand component $\Gamma(t,z,a,b).D_z w^{i,*}$. Here, the linear approximation to z at the point (t,z) is scaled in each dimension by the drift component in the dynamics of Z^i. Accordingly, $\Gamma(t,z,a,b).D_z w^{i,*}$ is the linear approximation of the drift in the dynamics $dw^{i,*}$ of $w^{i,*}$. This, together with $D_t w^{i,*}$ in Equation (3.20) forms the component referring to the deterministic derivative in Equation (3.20). The right-hand component of $\mathcal{L}^{(a,b)} w^{i,*}$, on the other hand, captures the effect of random perturbations during $[t,\theta]$ on i's value function: Here, the quadratic approximation is scaled in each dimension with the respective element in the 'covariance

matrix' $\Sigma(t,z,a,b)$.
Equations (3.19) and (3.20) together yield

$$\begin{aligned}
0 &= w^{i,*}(t,z) - \mathbb{E}_t^{\mathbb{Q}}[w^{i,*}(t,z)] \\
&= w^{i,*}(t,z) - \mathbb{E}_t^{\mathbb{Q}}\left[w^{i,*}(\tau, Z_\tau^{i,t,z}) - \int_t^\tau (D_t w^{i,*} + \mathcal{L}^{(a,b)})(s, Z_s^{i,t,z})\, ds + \xi_\tau\right] \\
&\leqslant \mathbb{E}_t^{\mathbb{Q}}\left[\int_t^\tau e^{-\mu s} v(s, Z_s^{i,t,z}, \gamma_s^i)\, ds + w^{i,*}(\tau, Z_\tau^{i,t,z}) - w^{i,*}(\tau, Z_\tau^{i,t,z}) + \int_t^\tau (D_t w^{i,*} + \mathcal{L}^{(a,b)})(s, Z_s^{i,t,z})\, ds + \xi_\tau\right] \\
&= \mathbb{E}_t^{\mathbb{Q}}\left[\int_t^\tau \left(D_t w^{i,*} + \mathcal{L}^{(a,b)} w^{i,*}\right)(s, Z_s^{i,t,z}) + e^{-\mu s} v(Z_s^{i,t,z}, a, b)\, ds + \xi_\tau\right].
\end{aligned}$$

In order to get rid of the expectation $\mathbb{E}_t^{\mathbb{Q}}[\xi_\tau]$, we recall that since ξ is a local martingale, there exists a sequence $(\tilde{\tau}_n)_{n\in\mathbb{N}}$ of stopping times such that $\lim_{n\to\infty} \tilde{\tau}_n = \infty$ a.s. and the stopped process $\xi^{\tilde{\tau}_n}$ is a martingale for all $n \in \mathbb{N}$. Let $\tau_n = (t + \frac{1}{n}) \wedge \tilde{\tau}_n$. Then $\lim_{n\to\infty} \tau_n = t$ and the stopped process ξ^{τ_n} is a martingale for all n and thus has vanishing time-t conditional expectation. Hence we obtain

$$0 \leqslant \mathbb{E}_t^{\mathbb{Q}}\left[\int_t^{\tau_n} \left(D_t w^{i,*} + \mathcal{L}^{(a,b)} w^{i,*}\right)(s, Z_s^{i,t,z}) + e^{-\mu s} v(Z_s^{i,t,z}, a, b)\, ds\right].$$

We want to transfer this inequality to the integrand on the right-hand side. To this end, first notice that the above inequality implies by the smoothness we posited that

$$\begin{aligned}
0 &\leqslant \lim_{n\to\infty} \mathbb{E}_t^{\mathbb{Q}}\left[\frac{1}{\tau_n}\int_t^{\tau_n} \left(D_t w^{i,*} + \mathcal{L}^{(a,b)} w^{i,*}\right)(s, Z_s^{i,t,z}) + e^{-\mu s} v(Z_s^{i,t,z}, a, b)\, ds\right] && (3.21)\\
&= D_t w^{i,*}(t,z) + \mathcal{L}^{(a,b)} w^{i,*}(t,z) + e^{-\mu s} v(s,z,a,b). && (3.22)
\end{aligned}$$

(2) In order to formally derive *equality* in Equation (3.21), let $\gamma^{i,*}$ be an optimal control for agent i. Then by (3.18) we have for all $n \in \mathbb{N}$ that

$$w^{i,*} = \mathbb{E}_t^{\mathbb{Q}}\left[\int_t^{\tau_n} e^{-\mu s}\, v(Z_s^{i,*}, \gamma_s^{i,*})\, ds + w^{i,*}(\tau_n, Z_{\tau_n}^{i,*})\right],$$

where $Z^{i,*}$ is the solution to the SDE governing the system dynamics starting from z at time t with control $\gamma^{i,*}$.
By proceeding similarly as above, we obtain

$$D_t w^{i,*}(t,z) + \mathcal{L}^{\gamma_t^{i,*}} w^{i,*}(t,z) + e^{-\mu s} v(s,z,\gamma_t^{i,*}) = 0.$$

This, together with Equation (3.22) suggests that $w^{i,*}$ should satisfy the partial differential equation

$$0 = D_t w^{i,*}(t,z) + \inf_{a\in G}\left[\mathcal{L}^{(a,b)} w^{i,*}(t,z) + e^{-\mu t} v(t,z,a,b)\right],$$

for all $(t,z) \in [0,T) \times \mathbb{R}^3$, if the infimum is finite. $\triangle$

3.5.2 Explicit Solution of each Agent's Hamilton-Jacobi-Bellman Equation

The objective of the following sections is to give a constructive proof of the existence of a weak equilibrium strategy bundle with respect to $\mathbb{P}$. By Proposition 7, we have that in order to do so, it suffices to show that there exists a $\mathbb{Q}$-strong equilibrium strategy bundle γ^* for some measure $\mathbb{Q} \sim \mathbb{P}$. Proposition 7 will then yield that the same strategy bundle γ^* is a weak equilibrium strategy bundle with respect to $\mathbb{P}$. In the following, we will derive the equilibrium under $\mathbb{Q}$ in closed-form. In particular, we will find a closed-form solution for the equilibrium price process P^* (under $\mathbb{Q}$), induced by an equilibrium strategy bundle γ^*. Based on Girsanov's Theorem, we then present sufficient conditions under which $\mathbb{Q}$ is indeed equivalent to $\mathbb{P}$ and under which, therefore, the conditions in Proposition 7 are satisfied.

Each agent continuously minimises expected abatement and trading costs at each point in time $t \in [0, T]$, where each agent's instantaneous cost function is given by

$$v(Z_t^i, \alpha_t^i, \beta_t^i) = Q_t \alpha_t^i + \varrho \cdot (\alpha_t^i)^2 - P_t \beta_t^i + \nu \cdot (\beta_t^i)^2.$$

We will now heuristically derive an educated guess regarding equilibrium strategies before proving that those strategies actually have the $\mathbb{Q}$-strong equilibrium property. We can develop our educated guess regarding the optimal strategies of each agent by recalling the equilibrium during the 'deterministic period' $[\theta, T]$. By the market clearing condition $\int_I \beta_t^i d\mathbb{I}(i) = 0, \ \forall t \in [0, T]$, we then obtain a market price process based on which we can solve any particular agent's optimisation problem. It will then turn out that the optimisation problem is in fact solved by the educated guess we made in the first place, solving the equilibrium problem altogether.

Developing an Intuition for the Candidate Strategies

Recall that we assumed no uncertainty between times θ and T, and that for $t \in [\theta, T]$ hence we obtained $P_t = e^{\mu(t-\theta)} P_\theta^*$,

$$\alpha_t^{i,*} = \frac{P_t^* - Q_t - 2\nu\Delta_t^{i,*}}{2(\nu + \varrho)}, \quad \text{and} \quad \beta_t^{i,*} = \frac{P_t^* - Q_t + 2\varrho\Delta_t^{i,*}}{2(\nu + \varrho)},$$

for all $i \in I$ and $t \in [\theta, T]$, where

$$\Delta_t^{i,*} = -\frac{\mu e^{\mu t}}{e^{\mu T} - e^{\mu\theta}} R_\theta^i. \tag{3.23}$$

No uncertainty during $[\theta, T]$ furthermore implies that for $t \in [\theta, T]$ we have

$$\begin{aligned} R_t^{i,*} &= R_\theta^i + \int_\theta^t \Delta_s^{i,*}\, ds = R_\theta^i - \frac{R_\theta^i}{e^{\mu T} - e^{\mu\theta}} \int_\theta^t \mu e^{\mu s}\, ds \\ &= R_\theta^i \left(1 - \frac{e^{\mu t} - e^{\mu\theta}}{e^{\mu T} - e^{\mu\theta}}\right) = R_\theta^i \frac{e^{\mu T} - e^{\mu t}}{e^{\mu T} - e^{\mu\theta}}, \end{aligned} \tag{3.24}$$

which in turn implies that $R^{i,*}$ has dynamics

$$dR_t^{i,*} = -R_\theta^i \frac{\mu e^{\mu t}}{e^{\mu T} - e^{\mu\theta}} dt = -h_t R_t^{i,*} dt, \quad \text{for} \quad t \in [\theta, T], \tag{3.25}$$

where for convenience we define

$$h_t = \frac{\mu e^{\mu t}}{e^{\mu T} - e^{\mu t}}.$$

During $[0, \theta]$, when the relevant BAU deployment is subject to uncertainty, the above strategies will serve us as a starting point to obtain the stochastic equilibrium strategies. To begin, notice that Equations (3.23), (3.24), and (3.25) together imply that

$$\Delta_t^{i,*} = -h_t R_t^{i,*} \quad \text{for} \quad t \in [\theta, T].$$

When agents' BAU deployment is subjected to uncertainty, so is R_t^i. However, at time $t \in [0, \theta]$, past changes to R_t^i from stochasticity are obviously incorporated in R_t^i which encourages us to suspect that a risk-neutral agent acts solely on his *expectation* of his future BAU deployment and applies the same *principle* of abatement and trading as in the deterministic case. Accordingly, our educated guess for all but one agent $j \in I \backslash \{i\}$ shall be given by $\tilde{\gamma}^j = (\tilde{\alpha}^j, \tilde{\beta}^j)$, where

$$\tilde{\alpha}_t^j = \frac{\tilde{P}_t - Q_t}{2(\nu + \varrho)} + \frac{\nu}{\nu + \varrho} h_t R_t^j \quad \text{and} \quad \tilde{\beta}_t^j = \frac{\tilde{P}_t - Q_t}{2(\nu + \varrho)} - \frac{\varrho}{\nu + \varrho} h_t R_t^j$$

for $t \in [0, T]$. Since agent i's actions have no impact on the market price, we then obtain by the market clearing condition

$$\int_{I\backslash\{i\}} \beta_t^j d\mathbb{I}(j) = \int_I \beta_t^j d\mathbb{I}(j) = 0$$

that the time-t market price for $t \in [0, T]$ is given by

$$\tilde{P}_t = Q_t + 2\varrho h_t R_t^I. \tag{3.26}$$

Given the above market price, we can consider agent i's problem of optimal stochastic control. This problem's solution will confirm the intuition that agent i's strategy $\gamma^i = (\alpha^i, \beta^i)$ is identical to any agent j's strategy, $j \in I\backslash\{i\}$, when R_t^j is replaced with R_t^i. And accordingly, this will prove that the strategies above form an equilibrium strategy bundle with the induced equilibrium price process given in Equation (3.26).

Substituting for the strategy $\tilde{\alpha}_t^j$, $\tilde{\beta}_t^j$ above, we obtain the dynamics for the process R_t^j:

$$dR_t^j = (\beta_t^j - \alpha_t^j)\, dt + d\mathbb{E}_t^{\mathbb{Q}}[Y^j(0,T)] = -\frac{\mu e^{\mu t}}{e^{\mu T} - e^{\mu t}} R_t^j\, dt + d\mathbb{E}_t^{\mathbb{Q}}[Y^j(0,T)].$$

Solving the above, we obtain:

$$R_t^j = R_0^j\, \frac{e^{\mu T} - e^{\mu t}}{e^{\mu T} - 1} + (e^{\mu T} - e^{\mu t}) \int_0^t \frac{d\mathbb{E}_s^{\mathbb{Q}}[Y^j(0,T)]}{e^{\mu T} - e^{\mu s}}.$$

Integrating over I yields, together with Equation (3.26) that

$$\tilde{P}_t = Q_t + 2\varrho h_t R_t^I = Q_t + 2\varrho \frac{\mu e^{\mu t}}{e^{\mu T} - 1}\, R_0^I + 2\varrho \mu e^{\mu t} \int_0^t \frac{d\mathbb{E}_s^{\mathbb{Q}}[Y^I(0,T)]}{e^{\mu T} - e^{\mu s}}.$$

In particular, we observe that P has the dynamics

$$d\tilde{P}_t = \mu \tilde{P}_t dt + 2\varrho h_t d\mathbb{E}_t^{\mathbb{Q}}[Y^I(0,T)].$$

Agent i's Problem of Optimal Stochastic Control

We now consider the problem of optimal abatement and trading for agent i. Let p denote an observed price and let $P^{t,p}$ denote the price process with time-t value $P_t^{t,p} = p$. Analogously, let $Q_t^{t,q} = q$. At time $s \in [t,T]$, agent i has to bear costs given by

$$v_s(Z_s^{i,t,z}, \alpha_s^i, \beta_s^i) = Q_s^{t,q}\, \alpha_s^i + \varrho \cdot (\alpha_s^i)^2 - P_s^{t,p} \beta_s^i + \nu \cdot (\beta_s^i)^2.$$

Agent i's problem is to find abatement- and trading strategies α^i and β^i respectively, such that, for all $t \in [0,T]$, the cost function J, given by

$$J(t,z,\alpha^i,\beta^i) = J(t,r,p,q,\alpha^i,\beta^i) = \mathbb{E}_t^{\mathbb{Q}}\left[\int_t^T e^{-\mu s} v_s(Z_s^i, \alpha_s^i, \beta_s^i)\, ds\right],$$

is minimised by α^i, β^i for all $z = (r,p,q) \in \mathbb{R}^3$, and such that the constraint $R_T^i = 0$ is satisfied. Let

$$w(t,z) = \inf_{(\alpha^i,\beta^i) \in \mathcal{G}^i(t,z)} J(t,z,\alpha^i,\beta^i)$$

denote the value function for agent i, where we omit the superscript i to improve legibility. The agent observes the state process $Z^i = (R^i, P, Q)$, where

$$\begin{aligned} dR_t^i = (\beta_t^i - \alpha_t^i)\, dt + d\mathbb{E}_t^{\mathbb{Q}}[Y^i(0,T)] &= (\beta_t^i - \alpha_t^i)dt + \sigma_t^i\, dW_t^{\mathbb{Q}}, \\ dP_t = \mu P_t dt + 2\varrho h_t d\mathbb{E}_t^{\mathbb{Q}}[Y^I(0,T)] &= \mu P_t dt + 2\varrho h_t \sigma_t^I\, dW_t^{\mathbb{Q}}, \\ dQ_t &= \mu Q_t dt, \end{aligned}$$

with initial value $Z_0^i = (R_0^i, P_0, Q_0) \in \mathbb{R}^3$.

The HJB Associated to Agent i's Minimisation Problem

We denote by D_z the differential operator with respect to z, and accordingly for other parameters. Accordingly, $D_z^2 = D_z \circ D_z$ denotes the second derivative operator with respect to z. In order to improve legibility, we will in the following simply write w instead of $w(t,z)$ for $(t,z) = (t,r,p,q) \in [0,T] \times \mathbb{R}^3$, where the parameterisation is clear from the context. The Hamilton-Jacobi-Bellman (HJB) Equation associated to the minimisation problem above is given by

$$\begin{aligned} 0 &= D_t w + \inf_{a,b} \left[(b-a) D_r w + \mu p D_p w + \mu q D_q w + \frac{1}{2} \mathrm{tr}(\Sigma\Sigma' D_z^2 w) + e^{-\mu t}(qa + \varrho a^2 - pb + \nu b^2) \right] \\ &= D_t w + \mu p D_p w + \mu q D_q w + \frac{1}{2} \mathrm{tr}(\Sigma\Sigma' D_z^2 w) + \inf_{a,b} \left[(b-a) D_r w + e^{-\mu t}(qa + \varrho a^2 - pb + \nu b^2) \right], \end{aligned}$$

where Σ is the vector

$$\Sigma = \begin{pmatrix} \sigma_t^i \\ 2\varrho h_t \sigma_t^I \\ 0 \end{pmatrix}$$

which implies that

$$\mathrm{tr}(\Sigma\Sigma' D_z^2 w) = (\sigma_t^i)^2 D_r^2 w + 2\varrho h_t \sigma_t^i \sigma_t^I D_p D_r w + 2\varrho h_t \sigma_t^i \sigma_t^I D_r D_p w + 4\varrho^2 h_t^2 (\sigma_t^I)^2 D_p^2 w.$$

Equivalent Formulation of the HJB Equation

Lemma 22. *The HJB Equation can be rewritten as*

$$0 = e^{\mu t} \left(D_t w + \mu p D_p w + \mu q D_q w \right) + \frac{e^{\mu t}}{2} tr(\Sigma\Sigma' D_z^2 w) - \frac{1}{4\varrho} \left(e^{\mu t} D_r w - q \right)^2 - \frac{1}{4\nu} \left(p - e^{\mu t} D_r w \right)^2. \tag{3.27}$$

Proof. We notice that the minimisers a, b in the above equation have to satisfy

$$a = \frac{1}{2\varrho} \left(e^{\mu t} D_r w - q \right) \quad \text{and} \quad b = \frac{1}{2\nu} \left(p - e^{\mu t} D_r w \right). \tag{3.28}$$

Furthermore, we notice that the second-order condition is satisfied for all a, b. The statement then follows immediately from the HJB equation, together with Equation (3.28) $\square$

The Singular Terminal Condition

In order to enforce the constraint $R_T^i = 0$, we impose the singular terminal condition

$$\lim_{t \nearrow T} w(t,r,p,q) = \begin{cases} 0 & : r = 0, \\ \infty & : r \neq 0. \end{cases} \tag{3.29}$$

Exact Solution to Agent i's Hamilton-Jacobi-Bellman Equation

In the following, we will provide an exact solution to agent i's Hamilton-Jacobi-Bellman Equation. To this end, we will first approach the problem with a plausible ansatz leading to a heuristic derivation. As it turns out, our approach does in fact lead to an exact solution, given in Theorem 23.

Heuristic Approach to Solving the Hamilton Jacobi Bellman Equation with Singular Terminal Condition

In the following we will find a solution w to the HJB Equation (3.27), that also satisfies the terminal condition (3.29). To this end, we notice that we can heuristically derive an educated guess on the general form of w from a and b as given in Equation (3.28), together with α^j, β^j for $j \in I \backslash \{i\}$, as observed by firm i: Since each firm is characterised exactly by its state process Z^i, intuition suggests that in equilibrium, all firms' strategies have the same functional form in terms of

$$\gamma^i : [0,T] \times \mathbb{R}^3 \to \mathbb{R}^2, \; (t,z) \mapsto \gamma^i(t,z).$$

Let $(t,z) = (t,r,p,q) \in [0,T] \times \mathbb{R}^3$. By Equation (3.27), we should observe

$$\frac{1}{2\varrho}\left(e^{\mu t} D_r w - q\right) = \frac{p-q}{2(\nu+\varrho)} + \frac{\nu}{\nu+\varrho} h_t r.$$

From this we immediately obtain

$$e^{\mu t} D_r w = \frac{2\nu\varrho}{\nu+\varrho} h_t r + \frac{\varrho}{\nu+\varrho} p + \frac{\nu}{\nu+\varrho} q,$$

i.e.

$$D_r w = \frac{e^{-\mu t}}{\nu+\varrho}\left(2\nu\varrho h_t r + \varrho p + \nu q\right).$$

This suggests that w should have the form

$$w(t,r,p,q) = F(t,r,p,q) + G(t,p,q),$$

where

$$F(t,r,p,q) = \frac{e^{-\mu t}}{\nu+\varrho}\left(\nu\varrho h_t r^2 + (\varrho p + \nu q) r\right),$$

and where the term $G(t,p,q)$ does not contain r. In this case, Equation (3.27) would be equivalent to

$$\begin{aligned} 0 =& e^{\mu t}\left(D_t F + \mu p D_p F + \mu q D_q F\right) + e^{\mu t}\left(D_t G + \mu p D_p G + \mu q D_q G\right) \\ &+ \frac{e^{\mu t}}{2}\operatorname{tr}(\Sigma\Sigma' D_z^2(F+G)) - \frac{1}{4\varrho}\left(e^{\mu t} D_r F - q\right)^2 - \frac{1}{4\nu}\left(p - e^{\mu t} D_r F\right)^2. \end{aligned} \tag{3.30}$$

We are going to plug F and G into the above equation in order to derive requirements on G. To this end, first notice that

$$D_t F(t,r,p,q) = -\mu F_t + \frac{e^{-\mu t}}{\nu+\varrho} \nu\varrho D_t h_t r^2,$$

where the time-t derivative of h with respect to time is given by

$$D_t h_t = D_t\left(\frac{\mu e^{\mu t}}{e^{\mu T} - e^{\mu t}}\right) = \mu h_t + h_t^2.$$

We thereby obtain

$$\begin{aligned} D_t F(t,r,p,q) &= -\mu F_t + \mu F_t - \frac{\mu e^{-\mu t}}{\nu+\varrho}(\varrho p + \nu q) r + \frac{e^{-\mu t}}{\nu+\varrho}\nu\varrho h_t^2 r^2 \\ &= -\frac{\mu e^{-\mu t}}{\nu+\varrho}(\varrho p + \nu q) r + \frac{e^{-\mu t}}{\nu+\varrho}\nu\varrho h_t^2 r^2. \end{aligned}$$

As for $D_p F$ and $D_q F$, we have

$$D_p F(t,r,p,q) = \frac{e^{-\mu t}}{\nu+\varrho}\varrho r \quad \text{and} \quad D_q F(t,r,p,q) = \frac{e^{-\mu t}}{\nu+\varrho}\nu r.$$

We plug these into the first term on the right-hand side of Equation (3.30) and find that

$$\begin{aligned} & e^{\mu t}(D_t F + \mu p D_p F + \mu q D_q F) \\ =& \frac{1}{\nu+\varrho}\left(-\mu(\varrho p + \nu q) r + \nu\varrho h_t^2 r^2 + \mu p \varrho r + \mu q \nu r\right) \\ =& \frac{\nu\varrho h_t^2 r^2}{\nu+\varrho}. \end{aligned} \tag{3.31}$$

Furthermore, we have

$$\frac{1}{4\varrho}\left(e^{\mu t}D_rF-q\right)^2=\varrho\left(\frac{p-q}{2(\nu+\varrho)}+\frac{\nu}{\nu+\varrho}h_tr\right)^2$$
$$=\frac{\varrho}{(\nu+\varrho)^2}\left(\frac{(p-q)^2}{4}+(p-q)\nu h_tr+\nu^2h_t^2r^2\right)$$

and analogously we have

$$\frac{1}{4\nu}\left(p-e^{\mu t}D_rF\right)^2=\nu\left(\frac{p-q}{2(\nu+\varrho)}-\frac{\varrho}{\nu+\varrho}h_tr\right)^2$$
$$=\frac{\nu}{(\nu+\varrho)^2}\left(\frac{(p-q)^2}{4}-(p-q)\varrho h_tr+\varrho^2h_t^2r^2\right).$$

Together, these equations yield

$$\frac{1}{4\varrho}\left(e^{\mu t}D_rF-q\right)^2+\frac{1}{4\nu}\left(p-e^{\mu t}D_rF\right)^2=\frac{1}{\nu+\varrho}\left(\frac{(p-q)^2}{4}+\nu\varrho h_t^2r^2\right). \tag{3.32}$$

We plug Equations (3.31) and (3.32) into the right-hand side of Equation (3.30) and obtain

$$0=e^{\mu t}(D_tG+\mu pD_pG+\mu qD_qG)+\frac{e^{\mu t}}{2}\mathrm{tr}(\Sigma\Sigma'D_z^2w)-\frac{(p-q)^2}{4(\nu+\varrho)}. \tag{3.33}$$

Consider the term $\frac{e^{\mu t}}{2}\mathrm{tr}(\Sigma\Sigma'D_z^2w)$ in Equation (3.30) and recall that

$$\mathrm{tr}(\Sigma\Sigma'D_z^2w)=(\sigma_t^i)^2D_r^2w+2\varrho h_t\sigma_t^i\sigma_t^ID_pD_rw+2\varrho h_t\sigma_t^i\sigma_t^ID_rD_pw+4\varrho^2h_t^2(\sigma_t^I)^2D_p^2w. \tag{3.34}$$

With $w(t,r,p,q)=F(t,r,p,q)+G(t,p,q)$, where G is independent of r, and recalling that

$$F(t,r,p,q)=\frac{e^{-\mu t}}{\nu+\varrho}\left(\nu\varrho h_tr^2+(\varrho p+\nu q)r\right),$$

we obtain

$$D_r^2w=D_r^2F=\frac{2e^{-\mu t}\nu\varrho h_t}{\nu+\varrho},\qquad D_pD_rw=D_rD_pw=D_pD_rF=\frac{e^{-\mu t}\varrho}{\nu+\varrho},$$

as well as

$$D_p^2w=D_p^2G.$$

Plugging this into Equation (3.34) we obtain

$$\mathrm{tr}(\Sigma\Sigma'D_z^2w)=(\sigma_t^i)^2\frac{2e^{-\mu t}\nu\varrho h_t}{\nu+\varrho}+4\varrho h_t\sigma_t^i\sigma_t^I\frac{e^{-\mu t}\varrho}{\nu+\varrho}+4\varrho^2h_t^2(\sigma_t^I)^2D_p^2w$$

and hence

$$\frac{e^{\mu t}}{2}\mathrm{tr}(\Sigma\Sigma'D_z^2w)=\frac{\nu\varrho h_t(\sigma_t^i)^2}{\nu+\varrho}+\frac{2\varrho^2h_t\sigma_t^i\sigma_t^I}{\nu+\varrho}+2\varrho^2h_t^2(\sigma_t^I)^2e^{-\mu t}D_p^2w.$$

The above equation, together with Equation (3.33) suggests that the function G should satisfy

$$-e^{\mu t}(D_tG+\mu pD_pG+\mu qD_qG)=\frac{\nu\varrho h_t(\sigma_t^i)^2}{\nu+\varrho}+\frac{2\varrho^2h_t\sigma_t^i\sigma_t^I}{\nu+\varrho}+2\varrho^2h_t^2(\sigma_t^I)^2e^{-\mu t}D_p^2w-\frac{(p-q)^2}{4(\nu+\varrho)}. \tag{3.35}$$

In order to solve the HJB Equation, it would thus suffice to find a sufficiently smooth function G that satisfies the above equation. Considering the fact that the right-hand side of Equation (3.35) only contains terms that either only depend on time or are quadratic in $p-q$, it is natural to approach this problem with the ansatz

$$G(t,p,q)=H_t\cdot(p-q)^2+\int_t^TC_s\,ds,$$

where H_t is only time-dependent, and where $-e^{\mu t}C_t$ cancels out all terms that only depend on time; i.e.

$$C_t=\frac{\nu\varrho h_t(\sigma_t^i)^2e^{-\mu t}}{\nu+\varrho}+\frac{2\varrho^2h_t\sigma_t^i\sigma_t^Ie^{-\mu t}}{\nu+\varrho}+4\varrho^2h_t^2(\sigma_t^I)^2H_t.$$

In summary, our deliberation suggest that w is of the form

$$w(t,r,p,q)=F(t,r,p,q)+H_t\cdot(p-q)^2+\int_t^TC_s\,ds,$$

with F and C as above. We then have that w as specified above satisfies the HJB Equation if

$$
\begin{aligned}
0 &= e^{\mu t}\left(D_t H_t (p-q)^2 + 2\mu p H_t (p-q) - 2\mu q H_t (p-q)\right) - \frac{(p-q)^2}{4(\nu+\varrho)} \\
&= e^{\mu t}\left(D_t H_t (p^2 - 2pq + q^2) + 2\mu p^2 H_t - 2\mu pq H_t - 2\mu qp H_t + 2\mu q^2 H_t\right) - \frac{(p-q)^2}{4(\nu+\varrho)} \\
&= e^{\mu t}\left((D_t H_t + 2\mu H_t)p^2 - 2(D_t H_t + 2\mu H_t)pq + (D_t H_t + 2\mu H_t)q^2\right) - \frac{(p-q)^2}{4(\nu+\varrho)} \\
&= \left(e^{\mu t}(D_t H_t + 2\mu H_t) - \frac{1}{4(\nu+\varrho)}\right)(p-q)^2.
\end{aligned}
$$

Since the above equation has to hold for any values of p and q, we arrive at the first-order linear ordinary differential equation

$$D_t H_t + 2\mu H_t = \frac{e^{-\mu t}}{4(\nu+\varrho)},$$

which can be solved using standard methods, by which we obtain

$$H_t = \frac{1 - k e^{-\mu t}}{4\mu e^{\mu t}(\nu+\varrho)}, \qquad \text{for some } k \in \mathbb{R}.$$

Notice that the singular terminal condition (3.29) is satisfied for $k = e^{\mu T}$. By employing the above heuristics, we thus do in fact find a solution to the HJB equation with terminal condition (3.29) given by

$$w(t,r,p,q) = F(t,r,p,q) + H_t \cdot (p-q)^2 + \int_t^T C_s \, ds,$$

with the functions F, H and C given as above. We summarise our respective findings in Theorem 23. △

Exact Solution to the Hamilton-Jacobi-Bellman Equation with Singular Terminal Condition

Theorem 23. *The HJB Equation* (3.27)*, together with the terminal condition* (3.29) *is solved by*

$$\tilde{w}(t,r,p,q) = \frac{\mu\nu\varrho r^2}{(e^{\mu T} - e^{\mu t})(\nu+\varrho)} + e^{-\mu t}\left(q + \frac{\varrho(p-q)}{\nu+\varrho}\right) r + \frac{(1 - e^{\mu(T-t)})(p-q)^2}{4\mu e^{\mu t}(\nu+\varrho)} + \int_t^T C_s \, ds$$

where

$$C_s = \frac{\mu\nu\varrho(\sigma_s^i)^2}{(e^{\mu T} - e^{\mu s})(\nu+\varrho)} + \frac{2\varrho^2 h_s \sigma_s^i \sigma_s^I e^{-\mu s}}{\nu+\varrho} + \frac{\varrho^2 h_s^2 (\sigma_s^I)^2 (1 - e^{\mu(T-s)})}{\mu e^{\mu s}(\nu+\varrho)} \quad \text{for} \quad t \leqslant s < T.$$

Proof. The statement follows from differentiation. □

Agent i's candidate strategies

Corollary 24. *The candidate strategy associated to the solution of the HJB Equation obtained in Theorem 23 is given by* $\tilde{\gamma}^i = (\tilde{\alpha}^i, \tilde{\beta}^i)$*, where*

$$\tilde{\alpha}_t^i = \frac{P_t - Q_t}{2(\nu+\varrho)} + \frac{\nu}{\nu+\varrho} h_t R_t^i \quad \text{and} \quad \tilde{\beta}_t^i = \frac{P_t - Q_t}{2(\nu+\varrho)} - \frac{\varrho}{\nu+\varrho} h_t R_t^i.$$

The candidate aggregate abatement is given by

$$\tilde{\alpha}_t = \mu e^{\mu t} \frac{R_0^I}{e^{\mu T} - 1} + \mu e^{\mu t} \int_0^t \frac{d\mathbb{E}_s^{\mathbb{Q}}[Y^I(0,T)]}{e^{\mu T} - e^{\mu s}}$$

and accordingly, the candidate equilibrium price process is given by

$$\tilde{P}_t = Q_t + 2\varrho \frac{\mu e^{\mu t}}{e^{\mu T} - 1} R_0^I + 2\varrho\mu e^{\mu t} \int_0^t \frac{d\mathbb{E}_s^{\mathbb{Q}}[Y^I(0,T)]}{e^{\mu T} - e^{\mu s}}.$$

Proof. Plugging $D_r w$ as obtained from Theorem 23 into Equation (3.28) yields the result. □

3.5.3 Verification Theorem

In this section, we omit the superscript i for legibility. All quantities depending on the agent index i are to be taken as with respect to a single agent $i \in I$. The respective aggregate quantities are still marked with the superscript I.

Note that $\tilde{w}$ is merely a candidate for the value function w^*, just like $(\tilde{\alpha}, \tilde{\beta})$ is merely a candidate for an optimal strategy (α^*, β^*). We therefore conduct a verification in the sense of the following Theorem 25.

In order to prove Theorem 25, we are going to proceed as follows: First we will show that $\tilde{w}(t,z)$ does not exceed the expected costs $J(t,z,\gamma)$ obtained for any arbitrary admissible strategy $\gamma \in \mathcal{G}^i(t,z)$. That is, we will show that $\tilde{w} \leqslant w^*$, with w^* established in Definition 19. Second, we will show that the candidate control $\tilde{\gamma}$ lies in $\mathcal{G}^i(t,z)$ and that the system state SDE $dZ_t = \Gamma(t, Z_t, \gamma_t)dt + \Sigma(t, Z_t, \gamma_t)dW_t^{\mathbb{Q}}$ has a unique solution, given $\gamma = \tilde{\gamma} = (\tilde{\alpha}, \tilde{\beta})$ and any initial value z. Third, we will show that $\tilde{w}$ equals the value function w^* and that $\tilde{\gamma}$ is an optimal control. The proof consists, at its core, of an application of Itō's Lemma and a localisation of the emerging local martingale part to eliminate it in expectation. We then obtain our results by taking the limits with respect to our localising sequences. This is possible due to the growths behaviour of $\tilde{w}$ and of total costs under any admissible strategies, which allows us to make use of the Dominated Convergence Theorem.

The Verification Theorem

Theorem 25. *Let w for any starting values $(t,z) = (t,r,p,q) \in [0,T] \times \mathbb{R}^3$, be given by*

$$\tilde{w}(t,r,p,q) = \frac{\mu\nu\varrho r^2}{(e^{\mu T} - e^{\mu t})(\nu + \varrho)} + e^{-\mu t}\left(q + \frac{\varrho(p-q)}{\nu + \varrho}\right) r + \frac{(1 - e^{\mu(T-t)})(p-q)^2}{4\mu e^{\mu t}(\nu + \varrho)} + \int_t^T C_s \, ds$$

and

$$C_s = \frac{\mu\nu\varrho(\sigma_s)^2}{(e^{\mu T} - e^{\mu s})(\nu + \varrho)} + \frac{2\varrho^2 h_s \sigma_s \sigma_s^I e^{-\mu s}}{\nu + \varrho} + \frac{\varrho^2 h_s^2 (\sigma_s^I)^2 (1 - e^{\mu(T-s)})}{\mu e^{\mu s}(\nu + \varrho)} \quad \text{for} \quad t \leqslant s < T.$$

Then $\tilde{w}$ equals the value function w^* for agent i's optimisation problem. That is

$$\tilde{w}(t,z) = \inf_{\gamma \in \mathcal{G}^i(t,z)} J(t,z,\gamma) = \inf_{\gamma \in \mathcal{G}^i(t,z)} \mathbb{E}_t^{\mathbb{Q}}\left[\int_t^T e^{-\mu s} v(Z_s^{t,z}, \gamma_s) \, ds\right] = w^*(t,z),$$

for all $(t,z) \in [0,T] \times \mathbb{R}^3$. The optimal trading and abatement strategy, which attains the infimum in the above equation, is given by $\gamma^* = (\alpha^*, \beta^*) = (\tilde{\alpha}, \tilde{\beta}) = \tilde{\gamma}$, where

$$\tilde{\alpha}(t,r,p,q) = \frac{1}{2\varrho}\left(e^{\mu t} D_r \tilde{w}(t,r,p,q) - q\right) \quad \text{and} \quad \tilde{\beta}(t,r,p,q) = \frac{1}{2\nu}\left(p - e^{\mu t} D_r \tilde{w}(t,r,p,q)\right)$$

for any $(t,z) = (t,r,p,q) \in [0,T] \times \mathbb{R}^3$. Accordingly, we have $\tilde{w}(t,z) = J(t,z,\tilde{\gamma})$.

First Lemma for the Verification Theorem

Lemma 26. *Let $(t,z) \in [0,T] \times \mathbb{R}^3$. For any admissible strategy $\gamma \in \mathcal{G}^i(t,z)$ we have the inequality*

$$\tilde{w}(t,z) \leqslant J(t,z,\gamma) = \mathbb{E}_t^{\mathbb{Q}}\left[\int_t^T e^{-\mu s} v(Z_s^{t,z}, \gamma_s) \, ds\right],$$

where $(Z_s^{t,z})_{s \in [t,T]}$ denotes the state process obtained for γ with time-t value $Z_t^{t,z} = z$. That is, we have

$$\tilde{w}(t,z) \leqslant w^*(t,z) = \inf_{\gamma \in \mathcal{G}^i(t,z)} J(t,z,\gamma).$$

Proof. Clearly, we have $\tilde{w} \in C^{1,2}([0,T] \times \mathbb{R}^3)$. Consider any initial values $(t,z) \in [0,T] \times \mathbb{R}^3$ and any arbitrary admissible strategy $\gamma \in \mathcal{G}^i(t,z)$. Let $Z^{t,z}$ be the state process obtained for the (arbitrary) strategy γ with time-t value $Z_t^{t,z} = z$. Let $u \in [t,T]$ and recall that $dZ_u^{t,z} = \Gamma(u, Z_u^{t,z}, \gamma_u)du + \Sigma(u, Z_u^{t,z}, \gamma_u)dW_u^{\mathbb{Q}}$ and

$$\mathcal{L}^{\gamma_u}\tilde{w}(u, Z_u^{t,z}) = \Gamma(u, Z_u^{t,z}, \gamma_u).D_z\tilde{w}(u, Z_u^{t,z}) + \frac{1}{2}\mathrm{tr}\left(\Sigma(u, Z_u^{t,z}, \gamma_u)\Sigma'(u, Z_u^{t,z}, \gamma_u)D_z^2\tilde{w}(u, Z_u^{t,z})\right).$$

We then have by Itō's Lemma that for any stopping time τ valued in $[t,T]$ and any $s \in [t,T]$

$$\begin{aligned}\tilde{w}(s \wedge \tau, Z_{s \wedge \tau}^{t,z}) = \tilde{w}(t,z) &+ \int_t^{s \wedge \tau} D_t\tilde{w}(u, Z_u^{t,z}) + \mathcal{L}^{\gamma_u}\tilde{w}(u, Z_u^{t,z}) \, du \\ &+ \int_t^{s \wedge \tau} D_z\tilde{w}(u, Z_u^{t,z})'\Sigma(u, Z_u^{t,u}, \gamma_u) \, dW_u^{\mathbb{Q}},\end{aligned} \tag{3.36}$$

where $s \wedge \tau = \min\{s, \tau\}$. We want to eliminate (in expectation) the local martingale on the rightmost end of the above equation. To this end, we replace the arbitrary stopping time above with the localising sequence $(\tau_n)_{n\in\mathbb{N}}$ of stopping times defined by

$$\tau_n = \inf\left\{s \in [t,\theta] : \int_t^s \left|D_z\tilde{w}(u, Z_u^{t,z})'\Sigma(u, Z_u^{t,z}, \gamma_u)\right|^2 du \geqslant n\right\} \quad \text{for } n \in \mathbb{N},$$

where we use the convention that $\inf \varnothing = \infty$. We notice that $\lim_{n\to\infty} \tau_n = \infty$. Let n be arbitrary in $\mathbb{N}$. We notice that the stopped process

$$\left(\int_t^{s\wedge\tau_n} D_z\tilde{w}(u, Z_u^{t,z})'\Sigma(u, Z_u^{t,z}, \gamma_u)\, dW_u^{\mathbb{Q}}\right)_{s\in[t,T]}$$

is true martingale. We thus obtain from Equation (3.36) that

$$\mathbb{E}_t\left[\tilde{w}(s\wedge\tau_n, Z_{s\wedge\tau_n}^{t,z})\right] = \tilde{w}(t,z) + \mathbb{E}_t\left[\int_t^{s\wedge\tau_n} D_t\tilde{w}(u, Z_u^{t,z}) + \mathcal{L}^{\gamma_u}\tilde{w}(u, Z_u^{t,z})\, du\right]. \tag{3.37}$$

Recall that $\tilde{w}$ satisfies the HJB Equation

$$0 = D_t w(t,z) + \inf_{(a,b)\in\mathbb{R}^2}\left[\mathcal{L}^{(a,b)}w(t,z) + e^{-\mu t}v(z,a,b)\right]$$

for any $(t,z) \in [0,T] \times \mathbb{R}^3$, from which we obtain for the arbitrary strategy $\gamma \in \mathcal{G}^i(t,z)$, and any $u \in [t, s\wedge\tau_n]$:

$$\begin{aligned}
0 &= D_t\tilde{w}(u, Z_u^{t,z}) + \inf_{(a,b)\in\mathbb{R}^2}\left[\mathcal{L}^{(a,b)}\tilde{w}(u, Z_u^{t,z}) + e^{-\mu u}v(Z_u^{t,z}, a, b)\right] \\
&\leqslant D_t\tilde{w}(u, Z_u^{t,z}) + \mathcal{L}^{\gamma_u}\tilde{w}(u, Z_u^{t,z}) + e^{-\mu u}v(Z_u^{t,z}, \gamma_u)
\end{aligned}$$

This, together with Equation (3.37), yields that

$$\mathbb{E}_t\left[\tilde{w}(s\wedge\tau_n, Z_{s\wedge\tau_n}^{t,z})\right] \geqslant \tilde{w}(t,z) - \mathbb{E}_t\left[\int_t^{s\wedge\tau_n} e^{-\mu u}v(Z_u^{t,z}, \gamma_u)\, du\right]. \tag{3.38}$$

We will use the Dominated Convergence Theorem to obtain the limits of the two expected value terms above for $n \to \infty$ and $s \nearrow \theta$. To this end, notice that for any $s \in [t,\theta]$ we have

$$\begin{aligned}
&\left|\tilde{w}((s\wedge\tau_n), R_{(s\wedge\tau_n)}, P_{(s\wedge\tau_n)}, Q_{(s\wedge\tau_n)})\right| \\
\leqslant &\left|\frac{\mu\nu\varrho}{(\nu+\varrho)}\frac{R^2_{(s\wedge\tau_n)}}{(e^{\mu T} - e^{\mu(s\wedge\tau_n)})}\right| + \left|e^{-\mu(s\wedge\tau_n)}\left(Q_{(s\wedge\tau_n)} + \frac{\varrho(P_{(s\wedge\tau_n)} - Q_{(s\wedge\tau_n)})}{\nu+\varrho}\right)R_{(s\wedge\tau_n)}\right| \\
&+ \left|\frac{(e^{\mu(T-(s\wedge\tau_n))} - 1)(P_{(s\wedge\tau_n)} - Q_{(s\wedge\tau_n)})^2}{4\mu e^{\mu(s\wedge\tau_n)}(\nu+\varrho)}\right| + \left|\int_{(s\wedge\tau_n)}^T C_u\, du\right| \\
\leqslant &\left|\frac{\mu\nu\varrho}{(\nu+\varrho)}\frac{R^2_{(s\wedge\tau_n)}}{(e^{\mu T} - e^{\mu(s\wedge\tau_n)})}\right| + \left|e^{-\mu(s\wedge\tau_n)}Q_{(s\wedge\tau_n)}R_{(s\wedge\tau_n)}\right| + \left|\frac{\varrho}{\nu+\varrho}e^{-\mu(s\wedge\tau_n)}(P_{(s\wedge\tau_n)} - Q_{(s\wedge\tau_n)})R_{(s\wedge\tau_n)}\right| \\
&+ \left|\frac{(e^{\mu(T-(s\wedge\tau_n))} - 1)(P_{(s\wedge\tau_n)} - Q_{(s\wedge\tau_n)})^2}{4\mu e^{\mu(s\wedge\tau_n)}(\nu+\varrho)}\right| + \left|\int_{(s\wedge\tau_n)}^T C_u\, du\right|.
\end{aligned} \tag{3.39}$$

We are going to show that each of the terms on the right-hand side of the above inequality is dominated by a function of finite expectation. For the first term, we have that

$$\left|\frac{\mu\nu\varrho}{(\nu+\varrho)}\frac{R^2_{(s\wedge\tau_n)}}{(e^{\mu T} - e^{\mu(s\wedge\tau_n)})}\right| = \frac{\mu\nu\varrho}{(\nu+\varrho)}\left|\frac{R^2_{(s\wedge\tau_n)}}{(e^{\mu T} - e^{\mu(s\wedge\tau_n)})}\right| \leqslant \frac{\mu\nu\varrho}{(\nu+\varrho)}\sup_{n\in\mathbb{N}}\left|\frac{R^2_{(s\wedge\tau_n)}}{(e^{\mu T} - e^{\mu(s\wedge\tau_n)})}\right|,$$

where the supremum on the right-hand side is finite since γ is an element of $\mathcal{G}^i(t,z)$ and thus has the properties set out in Proposition 11. As for the second term on the right-hand side of Equation (3.39) we have

$$\left|e^{-\mu(s\wedge\tau_n)}Q_{(s\wedge\tau_n)}R_{(s\wedge\tau_n)}\right| = \left|Q_0 R_{(s\wedge\tau_n)}\right| \leqslant Q_0 \sup_{n\in\mathbb{N}}\left|R_{(s\wedge\tau_n)}\right|,$$

where the finiteness of the supremum on the right-hand side again results from $\gamma \in \mathcal{G}^i(t,z)$. We now turn to the third term in Equation (3.39), for which we find

$$\begin{aligned}
&\left|\frac{\varrho}{\nu+\varrho}e^{-\mu(s\wedge\tau_n)}(P_{(s\wedge\tau_n)}-Q_{(s\wedge\tau_n)})R_{(s\wedge\tau_n)}\right|\\
=&\frac{\varrho}{\nu+\varrho}\left|e^{-\mu(s\wedge\tau_n)}\left(Q_{(s\wedge\tau_n)}+2\varrho\mu\frac{e^{\mu(s\wedge\tau_n)}}{e^{\mu T}-e^{\mu(s\wedge\tau_n)}}R^I_{(s\wedge\tau_n)}-Q_{(s\wedge\tau_n)}\right)R_{(s\wedge\tau_n)}\right|\\
=&\frac{\varrho}{\nu+\varrho}\left|\left(2\varrho\mu\frac{R^I_{(s\wedge\tau_n)}}{e^{\mu T}-e^{\mu(s\wedge\tau_n)}}\right)R_{(s\wedge\tau_n)}\right|\\
=&\frac{\varrho}{\nu+\varrho}\left|\left(2\varrho\mu\frac{R^I_{(s\wedge\tau_n)}}{e^{\mu T}-e^{\mu(s\wedge\tau_n)}}\right)R_{(s\wedge\tau_n)}\right|\\
\leqslant&\frac{\varrho}{\nu+\varrho}\left(2\varrho\mu\sup_{n\in\mathbb{N}}\left|\frac{R^I_{(s\wedge\tau_n)}}{e^{\mu T}-e^{\mu(s\wedge\tau_n)}}\right|\right)\sup_{n\in\mathbb{N}}|R_{(s\wedge\tau_n)}|,
\end{aligned}$$

where the finiteness of $\sup_{n\in\mathbb{N}}\left|\frac{R^I_{(s\wedge\tau_n)}}{e^{\mu T}-e^{\mu(s\wedge\tau_n)}}\right|$ and $\sup_{n\in\mathbb{N}}|R_{(s\wedge\tau_n)}|$ follow from $\gamma^j\in\mathcal{G}^j(t,z)$ for all $j\in I$: More precisely, we find that

$$\sup_{n\in\mathbb{N}}\left|\frac{R^I_{(s\wedge\tau_n)}}{e^{\mu T}-e^{\mu(s\wedge\tau_n)}}\right|=\sup_{n\in\mathbb{N}}\left|\frac{\int_I R^j_{(s\wedge\tau_n)}\,d\mathbb{I}(j)}{e^{\mu T}-e^{\mu(s\wedge\tau_n)}}\right|\leqslant\sup_{n\in\mathbb{N}}\left|\frac{\mathbb{I}(I)\sup_{j\in I}(R^j_{(s\wedge\tau_n)})}{e^{\mu T}-e^{\mu(s\wedge\tau_n)}}\right|.$$

Regarding the fourth term on the right-hand side of Equation (3.39), we find by Jensen's inequality that

$$\begin{aligned}
&\left|\frac{(e^{\mu(T-(s\wedge\tau_n))}-1)(P_{(s\wedge\tau_n)}-Q_{(s\wedge\tau_n)})^2}{4\mu e^{\mu(s\wedge\tau_n)}(\nu+\varrho)}\right|\\
=&\left|\frac{(e^{\mu T}-e^{\mu(s\wedge\tau_n)})}{4\mu e^{2\mu(s\wedge\tau_n)}(\nu+\varrho)}\left(Q_{(s\wedge\tau_n)}+2\varrho\mu\frac{e^{\mu(s\wedge\tau_n)}}{e^{\mu T}-e^{\mu(s\wedge\tau_n)}}R^I_{(s\wedge\tau_n)}-Q_{(s\wedge\tau_n)}\right)^2\right|\\
\leqslant&\left|\frac{(e^{\mu T}-e^{\mu(s\wedge\tau_n)})}{4\mu e^{2\mu(s\wedge\tau_n)}(\nu+\varrho)}4\varrho^2\mu^2\frac{e^{2\mu(s\wedge\tau_n)}(R^I_{(s\wedge\tau_n)})^2}{(e^{\mu T}-e^{\mu(s\wedge\tau_n)})^2}\right|\\
=&\left|\frac{\varrho^2\mu}{(\nu+\varrho)}\frac{(R^I_{(s\wedge\tau_n)})^2}{(e^{\mu T}-e^{\mu(s\wedge\tau_n)})}\right|\\
\leqslant&\frac{\varrho^2\mu}{(\nu+\varrho)}\sup_{n\in\mathbb{N}}\left(\frac{(R^I_{(s\wedge\tau_n)})^2}{(e^{\mu T}-e^{\mu(s\wedge\tau_n)})}\right),
\end{aligned}$$

where, again, the finiteness of the supremum on the right-hand side of the above inequality is due to $\gamma^j\in\mathcal{G}^j(t,z)$ for all $j\in I$. As for the fifth term in Equation (3.39), we have by our assumptions on σ and σ^I that all of the terms

$$\frac{\sigma_u^2}{e^{\mu T}-e^{\mu u}},\quad\frac{\sigma_u\sigma_u^I}{e^{\mu T}-e^{\mu u}},\quad\text{and}\quad\frac{(\sigma_u^I)^2}{e^{\mu T}-e^{\mu u}}$$

are bounded on $[0,T]$ and hence the right-hand side of

$$\left|\int_{(s\wedge\tau_n)}^T C_u\,du\right|\leqslant\sup_{n\in\mathbb{N}}\left|\int_{(s\wedge\tau_n)}^T C_u\,du\right|$$

is also finite. We thus find that by the Dominated Convergence Theorem, along with the continuity of $\tilde{w}$ and the pathwise continuity of $Z^{t,z}$, that

$$\lim_{n\to\infty}\mathbb{E}_t^{\mathbb{Q}}\left[\tilde{w}(s\wedge\tau_n,Z^{t,z}_{s\wedge\tau_n})\right]=\mathbb{E}_t^{\mathbb{Q}}\left[\lim_{n\to\infty}\tilde{w}(s\wedge\tau_n,Z^{t,z}_{s\wedge\tau_n})\right]=\mathbb{E}_t^{\mathbb{Q}}\left[\tilde{w}(s,Z^{t,z}_s)\right]$$

as well as: $\lim_{s\nearrow\theta}\mathbb{E}_t^{\mathbb{Q}}\left[\tilde{w}(s,Z^{t,z}_s)\right]=\mathbb{E}_t^{\mathbb{Q}}\left[\lim_{s\nearrow\theta}\tilde{w}(s,Z^{t,z}_s)\right]=\mathbb{E}_t^{\mathbb{Q}}\left[\tilde{w}(\theta,Z^{t,z}_\theta)\right]$. Furthermore, we have

$$\left|\int_t^{s\wedge\tau_n}e^{-\mu u}v(Z^{t,z}_u,\gamma_u)\,du\right|\leqslant\int_t^{s\wedge\tau_n}\left|e^{-\mu u}v(Z^{t,z}_u,\gamma_u)\right|\,du$$

and since $\gamma\in\mathcal{G}^i(t,z)$, the integral on the right-hand side is finite. Therefore, we can again apply the Dominated Convergence Theorem which implies

$$\lim_{n\to\infty}\mathbb{E}_t^{\mathbb{Q}}\left[\int_t^{s\wedge\tau_n}e^{-\mu u}v(Z^{t,z}_u,\gamma_u)\,du\right]=\mathbb{E}_t^{\mathbb{Q}}\left[\int_t^{s}e^{-\mu u}v(Z^{t,z}_u,\gamma_u)\,du\right].$$

as well as $\lim_{s\nearrow\theta}\mathbb{E}_t^{\mathbb{Q}}\left[\int_t^s e^{-\mu u}v(Z_u^{t,z},\gamma_u)\,du\right]=\mathbb{E}_t^{\mathbb{Q}}\left[\int_t^\theta e^{-\mu u}v(Z_u^{t,z},\gamma_u)\,du\right]$. In summary, we obtain with Equation (3.38) that

$$\tilde{w}(t,z)\leqslant\mathbb{E}_t^{\mathbb{Q}}\left[\int_t^\theta e^{-\mu u}v(Z_u^{t,z},\gamma_u)\,du+\tilde{w}(\theta,Z_\theta^{t,z})\right]\leqslant\mathbb{E}_t^{\mathbb{Q}}\left[\int_t^T e^{-\mu u}v(Z_u^{t,z},\gamma_u)\,du\right],$$

where the last inequality follows from the optimality of $\tilde{\gamma}$ on $[\theta,T]$, as established in Theorem 17. □

Second Lemma for the Verification Theorem

Lemma 27. *Let $z=(r,p,q)\in\mathbb{R}^3$. The candidate control $\tilde{\gamma}=(\tilde{\alpha},\tilde{\beta})$ given by*

$$\tilde{\alpha}(t,r,p,q)=\frac{1}{2\varrho}\left(e^{\mu t}D_r\tilde{w}(t,r,p,q)-q\right)\quad\text{and}\quad\tilde{\beta}(t,r,p,q)=\frac{1}{2\nu}\left(p-e^{\mu t}D_r\tilde{w}(t,r,p,q)\right)\tag{3.40}$$

is an element of $\mathcal{G}^i(t,z)$ and the system state SDE,

$$dZ_t=(dR_t,dP_t,dQ_t)'=\Gamma(t,Z_t,\tilde{\alpha}_t,\tilde{\beta}_t)dt+\Sigma(t,Z_t,\tilde{\alpha}_t,\tilde{\beta}_t)dW_t^{\mathbb{Q}},$$

where

$$\Gamma(t,Z_t,\tilde{\alpha}_t,\tilde{\beta}_t)=\begin{pmatrix}(\tilde{\beta}_t-\tilde{\alpha}_t)\\ \mu P_t\\ \mu Q_t\end{pmatrix}\quad\text{and}\quad\Sigma(t,Z_t,\tilde{\alpha}_t,\tilde{\beta}_t)=\begin{pmatrix}\sigma_t\\ 2\varrho h_t\sigma_t^I\\ 0\end{pmatrix}$$

has a unique strong solution $Z^{t,z}$ for any time-t initial value $z=(r,p,q)$ and any $t\in[0,T]$.

Proof. We have $\sigma\equiv 0$ on $[\theta,T]$. Thus it suffices to show existence and uniqueness of a strong solution to the above SDE only on $[0,\theta]$, since it is clearly given on $[\theta,T]$. To this end, we will treat each component in dZ_t separately. Resolving $D_r\tilde{w}$ in Equation (3.40) yields

$$\alpha_t=\frac{P_t-Q_t}{2(\nu+\varrho)}+\frac{\nu}{\nu+\varrho}\,h_tR_t\quad\text{and}\quad\beta_t=\frac{P_t-Q_t}{2(\nu+\varrho)}-\frac{\varrho}{\nu+\varrho}\,h_tR_t.$$

From this we obtain the dynamics for the process R,

$$dR_t=(\beta_t-\alpha_t)\,dt+\sigma_t dW_t^{\mathbb{Q}}=-\frac{\mu e^{\mu t}}{e^{\mu T}-e^{\mu t}}R_t\,dt+\sigma_t dW_t^{\mathbb{Q}}.\tag{3.41}$$

We find that for any $\bar{R},\hat{R}\in\mathbb{R}$

$$\left|\frac{\mu e^{\mu t}}{e^{\mu T}-e^{\mu t}}\bar{R}-\frac{\mu e^{\mu t}}{e^{\mu T}-e^{\mu t}}\hat{R}\right|\leqslant\frac{\mu e^{\mu t}}{e^{\mu T}-e^{\mu \theta}}\,|\bar{R}-\hat{R}|\leqslant\frac{\mu e^{\mu\theta}}{e^{\mu T}-e^{\mu\theta}}\,|\bar{R}-\hat{R}|$$

and thus the SDE (3.41) satisfies a Lipschitz-condition with Lipschitz-constant $\mu e^{\mu\theta}/(e^{\mu T}-e^{\mu\theta})$. Furthermore, the SDE satisfies the linear growth condition

$$\left|\frac{\mu e^{\mu t}}{e^{\mu T}-e^{\mu t}}\hat{R}\right|+|\sigma_t|\leqslant\frac{\mu e^{\mu\theta}}{e^{\mu T}-e^{\mu\theta}}\,|\hat{R}|+|\sigma_t|,$$

where $\mathbb{E}^{\mathbb{Q}}\left[\int_0^T|\sigma_s|^2\,ds\right]<\infty$ follows from our assumptions on σ. We thus find by [Pham, 2009], Section 1.3 in Chapter 1, that the SDE (3.41) has a unique strong solution $R^{t,\hat{R}}$ with $R_t^{t,\hat{R}}=\hat{R}$. Furthermore, we find that since the above finiteness applies for all agents, $\sigma_t h_t$ is bounded by assumption and Q trivially has a unique strong solution, the price process with $P_t=Q_t+2\varrho h_t\sigma_t dW_t^{\mathbb{Q}}$ for $t\in[0,\theta]$ exists and is unique. We thereby find that there is a unique strong solution $Z^{t,z}$ to the original SDE in question with $Z_t^{t,z}=z$. We also find that for all $s\in[t,T]$

$$\int_t^s\left|\Gamma(u,Z_u^{t,z},\tilde{\gamma}_u)\right|^2+\left|\Sigma(u,Z_u^{t,z},\tilde{\gamma}_u)\right|^2\,du<\infty,$$

which means that $\tilde{\gamma}$ satisfies the boundedness condition we imposed on the state process for any $\gamma\in\tilde{\mathcal{G}}^i$. Furthermore we find that agent i does in fact incur finite costs:

$$\mathbb{E}_t^{\mathbb{Q}}\left[\int_t^T e^{-\mu s}\left|v(s,Z_s^{t,z},\tilde{\gamma}_s)\right|\,ds\right]<\infty.$$

Therefore, $\tilde{\gamma}\in\mathcal{G}^i(t,z)$. □

Proof of the Verification Theorem 25

Proof. Let $Z^{t,z}$ denote the state process obtained for the candidate strategy $\tilde{\gamma}$. For any stopping time τ valued in $[t,\infty)$ and any $s \in [t,T]$, we have by Itō's Lemma:

$$\tilde{w}(s\wedge\tau, Z^{t,z}_{s\wedge\tau}) = \tilde{w}(t,z) + \int_t^{s\wedge\tau} D_t\tilde{w}(u,Z^{t,z}_u) + \mathcal{L}^{\tilde{\gamma}(t,z)}\tilde{w}(u,Z^{t,z}_u)\,du \\ + \int_t^{s\wedge\tau} D_z\tilde{w}(u,Z^{t,z}_u)'\Sigma(Z^{t,z}_u,\tilde{\gamma}_u)\,dW^{\mathbb{Q}}_u. \tag{3.42}$$

In particular we have for the localising sequence

$$\tau_n = \inf\left\{s \geqslant t : \int_t^s \left|D_z\tilde{w}(u,Z^{t,z}_u)'\Sigma(Z^{t,z}_u,\tilde{\gamma}_u)\right|^2\,du \geqslant n\right\} \quad \text{for } n\in\mathbb{N}$$

that the stopped local martingale part in Equation (3.42) is a true martingale and hence vanishes in expectation, i.e.

$$\mathbb{E}^{\mathbb{Q}}_t\left[\tilde{w}(s\wedge\tau_n, Z^{t,z}_{s\wedge\tau_n})\right] = \tilde{w}(t,z) + \mathbb{E}^{\mathbb{Q}}_t\left[\int_t^{s\wedge\tau_n} D_t\tilde{w}(u,Z^{t,z}_u) + \mathcal{L}^{\tilde{\gamma}(t,z)}\tilde{w}(u,Z^{t,z}_u)\,du\right]. \tag{3.43}$$

Since $\tilde{\gamma}$ is chosen such that $\tilde{\gamma}_u$ is the unique minimiser on the right-hand side of the HJB equation

$$0 = D_t w(u,Z_u) + \inf_{(a,b)\in G}\left[\mathcal{L}^{(a,b)}w(u,Z_u) + e^{-\mu u}v(u,Z_u,a,b)\right]$$

for any (u,Z_u), we have

$$0 = D_t w(u,Z_u) + \mathcal{L}^{\tilde{\gamma}(t,z)}w(u,Z_u) + e^{-\mu u}v(u,Z_u,\tilde{\gamma}_u).$$

From this obtain for Equation (3.43) that

$$\mathbb{E}^{\mathbb{Q}}_t\left[\tilde{w}(s\wedge\tau_n, Z^{t,z}_{s\wedge\tau_n})\right] = \tilde{w}(t,z) - \mathbb{E}^{\mathbb{Q}}_t\left[\int_t^{s\wedge\tau_n} e^{-\mu u}v(u,Z^{t,z}_u,\tilde{\gamma}_u)\,du\right].$$

Since $\tilde{\gamma}\in\mathcal{G}^i(t,z)$ we can repeat the argument in the proof of Lemma 26 and use the Dominated Convergence Theorem to obtain

$$\mathbb{E}^{\mathbb{Q}}_t\left[\tilde{w}(s,Z^{t,z}_s)\right] = \tilde{w}(t,z) - \mathbb{E}^{\mathbb{Q}}_t\left[\int_t^s e^{-\mu u}v(u,Z^{t,z}_u,\tilde{\gamma}_u)\,du\right]$$

and let $s \nearrow T$ to deduce

$$\tilde{w}(t,z) = \mathbb{E}^{\mathbb{Q}}_t\left[\int_t^T e^{-\mu u}v(u,Z^{t,z}_u,\tilde{\gamma}_u)\,du\right].$$

That is, $\tilde{w}(t,z) = J(t,\tilde{\gamma},z) \geqslant \inf_{\gamma\in\mathcal{G}^i(t,z)} J(t,\gamma,z) = w^*(t,z)$, where w^* is the value function associated to i's optimisation problem. Recalling that from Lemma 26 we have $w(t,z) \leqslant w^*(t,z)$ we finally have $\tilde{w} = w^*$, which is what we wanted to prove. ☐

3.6 Uniqueness of the Equilibrium

Notice that our assumptions on $\mathcal{G}^i$ for all $i \in I$ yield that any equilibrium price process must be a diffusion process, since it must always be a component of each agent's state process. This, together with a simple arbitrage-argument yields that any equilibrium price process induced by a strategy bundle in $(\mathcal{G}^i)_{i\in I}$ must satisfy an SDE of the form

$$dP_t = \mu P_t dt + \hat{\sigma}_t(Z_t)dW_t^{\mathbb{Q}},$$

where $\hat{\sigma}$ must be compatible with the assumptions on admissible strategies. However, we also notice from the Verification Theorem above, that each agent's optimal strategies do not depend on σ^I and thus do not depend on $(h_t\sigma_t^I)_{t\in[0,T]}$ and would thus be valid for all $\hat{\sigma}$ that are compatible with the growth conditions on Σ^i for all $i \in I$. Therefore, we obtain that our $\mathbb{Q}$-strong equilibrium strategy bundle γ^* is in fact unique in $(\mathcal{G}^i)_{i\in I}$.

3.7 Weak Equilibrium Property under the Objective Measure

We solved the model above under some measure $\mathbb{Q}$, a sufficient condition for the equivalence to $\mathbb{P}$ of which we are now going to provide. Under these conditions, our $\mathbb{Q}$-strong equilibrium also has the $\mathbb{P}$-weak equilibrium property. As not entirely unconventional, we will to this end, use Girsanov's famous theorem, a one-dimensional version of which is given in Theorem 28.

Girsanov's Theorem

Theorem 28 (Girsanov, 1-dimensional)**.** *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$ be a filtered probability space where the filtration $(\mathcal{F}_t)_{t\geqslant 0}$ is generated by a $\mathbb{P}$-Brownian motion $W^{\mathbb{P}} = (W_t^{\mathbb{P}})_{t\geqslant 0}$. Let a process $\lambda_t = (\lambda_t)_{t\geqslant 0}$ be an appropriately measurable stochastic process and define L_t as*

$$L_t = \exp\left(-\int_0^t \lambda_s \, dW_s^{\mathbb{P}} - \frac{1}{2}\int_0^t \lambda_s^2 \, ds\right).$$

If $L = (L_t)_{t\geqslant 0}$ is a strictly positive martingale under $\mathbb{P}$, then there exists a probability measure $\mathbb{Q}$ with $\mathbb{P} \sim \mathbb{Q}$ such that the Radon-Nikodým derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ on $(\Omega, \mathcal{F}_t)$ is given by

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = L_t$$

for all $t \in [0,T]$ and the process $W^{\mathbb{Q}} = (W_t^{\mathbb{Q}})_{t\geqslant 0}$ defined by

$$W_t^{\mathbb{Q}} = W_t^{\mathbb{P}} + \int_0^t \lambda_s \, ds, \text{ for all } t \in [0,T],$$

is a Brownian motion under $\mathbb{Q}$.

First, we appreciate that under the objective measure $\mathbb{P}$, the commodity price return is not necessarily equal to the risk-free rate μ. On the contrary, agents internalise the risk associated to investments in low-carbon technologies (or any investment into reducing the consumption of the commodity in question), and the commodity becomes a risky asset. Accordingly, investors request a risk-premium ζ_t for investing into the commodity at price P_t at time t. Given the equilibrium dynamics above, it is thus natural to require that under $\mathbb{P}$, the price dynamics should be given by

$$dP_t = (\mu + \zeta_t)P_t dt + 2\varrho h_t \sigma_t^I dW_t^{\mathbb{P}}, \tag{3.44}$$

where $W^{\mathbb{P}}$ is a standard Brownian motion under $\mathbb{P}$. Recall that under $\mathbb{Q}$, the price process has dynamics

$$dP_t = \mu P_t dt + 2\varrho h_t \sigma_t^I dW_t^{\mathbb{Q}}. \tag{3.45}$$

We require that no strong arbitrage opportunities shall exist in equilibrium, when we evaluate it under the objective measure $\mathbb{P}$. By Lemma 4 we know that such arbitrage is indeed not possible if $\mathbb{P} \sim \mathbb{Q}$.

Naturally, we require the two price dynamics under $\mathbb{P}$ and $\mathbb{Q}$ in Equations (3.44) and (3.45) to be consistent. If that is the case, then we observe that the Brownian motions $W^{\mathbb{P}}$ and $W^{\mathbb{Q}}$ have to satisfy

$$dW_t^{\mathbb{Q}} = dW_t^{\mathbb{P}} + \frac{\zeta_t P_t}{2\varrho h_t \sigma_t^I} \, dt \quad \text{for } t \in [0,\theta)$$

and $dW_t^{\mathbb{Q}} = dW_t^{\mathbb{P}}$ for $t \in [\theta, T]$. We see that in our case, sufficient conditions for $\mathbb{Q} \sim \mathbb{P}$ can be reduced to a growth condition on the relation between the time-t risk-premium ζ_t and the volatility coefficient in the rate of return of allowances price under $\mathbb{P}$, commonly denoted formally by dP_t/P_t. We can interpret the risk-premium as a quantification of agents' risk-aversion, and it is intuitive to model this as a function of the volatility parameter in dP_t/P_t.

In order to derive our equilibrium dynamics under $\mathbb{P}$, we will make a simplifying assumption on how agents' risk-aversion, quantified by a risk-premium on the allowance price returns, depends on these returns' volatility coefficient.

It would be counterintuitive to assume that the relation in Equation (3.46) is directly time-dependent but should at least have martingale character in the sense that, unless unpredicted economy-wide events (such as large-scale economic crises) or unpredicted asset-specific events (such as technological progress) influence agents' risk-aversion, market sentiment and its impact on risk-aversion should remain unchanged throughout $[0,\theta)$. For simplicity, we are going to refrain from modelling the dependency of risk-aversion on such unforeseen developments and assume that ζ_t is simply proportional to the volatility coefficient in dP_t/P_t; i.e. that there exists a constant $k \geqslant 0$ such that that

$$\zeta_t = k \cdot \frac{2\varrho h_t \sigma_t^I}{P_t} \text{ for } t \in [0,\theta) \tag{3.46}$$

and $\zeta_t = 0$ for $t \in [\theta, T]$. In line with Girsanov's Theorem we then obtain that the process L_t given by

$$L_t = \exp\left(-\int_0^t \frac{\zeta_s P_s}{2\varrho h_s \sigma_s^I} \, dW_s^{\mathbb{P}} - \frac{1}{2}\int_0^t \left(\frac{\zeta_s P_s}{2\varrho h_s \sigma_s^I}\right)^2 \, ds\right),$$

is a $\mathbb{P}$-martingale by Novikov's condition and hence we obtain the desired change of measure by the Radon-Nikodým derivative $d\mathbb{Q}/d\mathbb{P}|_{\mathcal{F}_t} = L_t$.

Consider again Equation (3.46). As we see, ζ_t depends on both σ_t^I and P_t. In particular, any dynamic changes to the supply of commodity units could potentially alter the volatility coefficient on the right-hand side of Equation (3.46) and hence have an impact on the risk-premium. While the impact on supply volatility may be straightforward and simply given by design, the impact on price volatility and, ultimately, on the risk-premium is subject to the endogenous market reaction of the model. In Chapter 4 we will explore this impact in greater detail and examine the relationship between the risk-premium and total costs under a generic supply management policy.

Chapter 4

Policy Relevance and Application

In the following, we will focus our attention mainly on the entirety of regulated agents. In order to improve legibility, we thus omit the superscript I for aggregate quantities.

This chapter is devoted to extending our findings from Chapter 3 to a problem of optimal regulatory control that is particularly relevant in the context of climate relevant resources. Inspired by the recent reform of the EU ETS to include a Market Stability Reserve (MSR) and the preceding political and academic debate, we will consider a generic mechanism that adjusts the supply of a resource in response to changes in agents' respective bank. We let $f = (f_t)_{t\in[0,T]}$ denote the regulator's ex-ante allocation schedule. At each point in time t, the resource allocation will be adjusted based on the bank B_t, which itself is subject to exogenous shocks via the equilibrium dynamics. Our mechanism will provide an abstraction of the rules implemented by the EC MSR[1].

The effects of *any* such and similar mechanism will depend on how *responsive* it is in its operation. That is to say, how strong are the changes made to the regulatory environment when changes to exogenous factors occur? We want to concretise the notion of responsiveness in order to gain an understanding of how a system can benefit in terms of cost-efficiency from a regulatory framework that can react to exogenous changes. In particular, we note that the effects of any mechanism similar to the MSR depends on the degree by which the regulator is willing to deviate from its ex-ante schedule f and, possibly, the ex-ante cap. We will find that the readiness to accept ex-post changes to the (expected) cap is intimately linked to the responsiveness of the system – a phenomenon that we will expound in some detail in the following section.

4.1 Cap Stringency and Responsiveness

Recall that in Section 3.5 we obtained a market equilibrium strategy bundle $(\gamma^{i,*})_{i\in I}$ where $\gamma^{i,*} = (\alpha^{i,*}, \beta^{i,*})$, with

$$\alpha_t^{i,*} = \frac{P_t - Q_t}{2(\nu+\varrho)} + \frac{\nu}{\nu+\varrho}\, h_t R_t^i$$

$$\text{and} \quad \beta_t^{i,*} = \frac{P_t - Q_t}{2(\nu+\varrho)} - \frac{\varrho}{\nu+\varrho}\, h_t R_t^i,$$

for all $t \in [0,T]$. Also recall that agent i's time-t expected residual abatement requirement is given by

$$R_t^i = \mathbb{E}_t^{\mathbb{Q}}[E^i(t,T) - A^i(t,T)] - B_t^i + c.$$

Furthermore, recall that the induced equilibrium price process is given by

$$P_t = Q_t + R_0 \frac{2\varrho\mu e^{\mu t}}{e^{\mu T}-1} + 2\varrho\mu e^{\mu t}\int_0^t \frac{d\xi_s}{e^{\mu T} - e^{\mu s}},$$

where

$$d\xi_s = d\mathbb{E}_s^{\mathbb{Q}}\left[E(0,T) - A(0,T)\right].$$

The process ξ reflects changes in the agents' expectations about their required abatement as a result of changes in the expected total BAU deployment $\mathbb{E}_t^{\mathbb{Q}}[E(0,T)]$ and expected total allocations $\mathbb{E}_t^{\mathbb{Q}}[A(0,T)]$. With a supply adjustment mechanism in place, changes to BAU deployment may partially or fully be offset by changes to the remaining expected allocation: In case of an economic crisis such as most recently observed in 2007/08, we may witness (and indeed witnesses in the example) a mounting bank of unused commodity units due to a massive decline in expected production. Accordingly, the commodity's market price dropped significantly and remained on a low level, causing doubt in the efficacy of the policy instrument altogether. In light of the long time horizons and substantial uncertainties on behalf of market participants, these claims should not necessarily be rejected offhand. A responsive mechanism may compensate part of such shocks by reducing or increasing the resource allocation, such that shocks are not transferred one-to-one onto the residual abatement requirement R_t and, as a side-effect, the price response to unanticipated events may be mitigated.

Below, we will consider such a responsive mechanism, where the stringency by which the regulator adheres to the ex-ante cap $\int_0^T f_t dt$ may be loosened in favour of regulatory responsiveness toward unanticipated events.[2] In order to structure our analysis and interpretation, we will first establish some useful notions and terminology. Firstly, consider the general concept of *cap stringency*: This term shall henceforth carry the concept of how far the regulator is willing to deviate from the ex-ante cap. In the example of emissions trading systems, the regulator may, depending on the bank of unused allowances, remove allowances from the allocation or introduce more allowances than previously planned. The higher the stringency, however, the less the regulator is willing to do so – a higher cap stringency implies a lower propensity to deviate from the ex-ante cap.

[1] Therein, less allowances will be made available next year than prescribed by the ex-ante schedule f, if the current year's bank exceeds a certain threshold. Conversely, if the current bank has dropped below another (different) threshold, more allowances will be released to the market than scheduled ex-ante. In our responsive mechanism, we do not assume any threshold levels but continuously adjust the allocation schedule for any values of B_t^i, $i \in I$. However, we provide some insights into threshold levels later in the text.

[2] Note that in climate relevant applications, the cap should typically only be altered within an acceptable range; i.e. up to a GHG limit that prevents catastrophic damages to the environment. In the following we neglect this constraint for tractability reasons but note that within this range, the optimal instrument is *not* necessarily a border solution on the policy spectrum.

Note that we will use the term *cap stringency* only in a qualitative sense. In the interest of useful quantification, we introduce another notion, complementary to cap stringency: Each time the resource allocation is adjusted, we assume that the regulator does so based on a fixed set of rules and on the *bank* of unused commodity units. The farther the bank deviates from the reference point $c \in \mathbb{R}$, the stronger the response will be in absolute terms. However, we will let the regulator decide ex-ante on the *relative* strength of his response; that is to say: the regulator may decide ex-ante how strong the response is in relation to the difference $|B_t - c|$. A relatively 'weak' response will let the bank deviate more freely from the point c, which preserves the ex-ante cap to a large extend. In this case, the mechanism is only slightly responsive, while the *cap stringency* is rather high. In Section 4.2, we will quantify this relative responsiveness in terms of a parameter $\delta \geqslant 0$, labelled *adjustment rate*. With a high adjustment rate, the strength of the response, relative to the difference $|B_t - c|$ will be high, which may require the regulator to deviate farther from the ex-ante cap. In that sense, the terms *cap stringency* and *responsiveness*, the latter being quantified by the *adjustment rate*, represent complementary notions, which will be quite useful in the following.

4.2 A Regulatory Mechanism with Adjustable Responsiveness

In this section, we construct a responsive mechanism that covers a spectrum of policies between a pure price instrument and a pure quantity instrument by means of parameterisation with an adjustment rate δ. This will lay the ground for Section 4.3, where we will show how the regulator can identify an adjustment rate that minimises expected aggregate total abatement costs. In particular, we will identify a trade-off between a high responsiveness, which, as we shall see, lowers the agents' costs of adjusting to shocks on the one hand; and a low responsiveness, which yields potential for agents to benefit from inter-temporal cost-saving opportunities on the other hand.

We begin by indexing the regulator's allocation adjustments to the aggregate bank of commodity units: At each time $t \in [0, T]$, if the current aggregate bank is above the fixed level $c \in \mathbb{R}$, we let a fraction δdt (where $\delta \geqslant 0$) of the difference $|B_t^i - c|$ be instantaneously removed from the allocation, for each $i \in I$. Conversely, if the aggregate bank drops below the level c, the allocation is increased by $\delta \cdot |B_t^i - c| dt$, compared to the ex-ante schedule.[3] Here, we abstract the workings of the MSR, in which allowances are removed from (or added to) *next* year's allocation and set aside into (or taken from) a so-called 'reserve'. For the purpose of an analysis of the policy spectrum in our continuous-time setting, however, we opt for an *instantaneous* allocation adjustment.

We point out that for each agent i's optimisation problem in the previous chapter to be applicable, any assumptions on the BAU deployment process, together with the aggregate system dynamics, must admit a representation of ξ^i as $d\xi_t^i = \sigma_t^i dW_t^{\mathbb{Q}}$, where the process σ^i is deterministic and bounded with $\sigma^i \equiv 0$ on $[\theta, T]$. Regarding the random perturbations of BAU deployment, we here make a fairly simple assumption for the sake of tractability: Let each agent i's cumulated BAU deployment process be given by

$$E^i(0,t) = \int_0^t g_s^i \, ds \; + \; \int_0^t \kappa_s^i \, dW_s^{\mathbb{Q}},$$

where g^i and κ^i shall be some deterministic and bounded processes and $\kappa_t^i = 0$ for $t \in [\theta, T]$. As we shall see as a by-product of the following Proposition 29, our responsive mechanism and the above assumption on BAU deployment are in fact compatible with the assumption that $d\xi_t^i = \sigma_t^i dW_t^{\mathbb{Q}}$ and σ^i will indeed turn out to be deterministic and bounded with $\sigma^i \equiv 0$ on $[\theta, T]$.

We can express the instantaneous change in the aggregate bank of commodity units by

$$dB_t = f_t dt + \delta(c - B_t)dt - g_t dt - \kappa_t dW_t^{\mathbb{Q}} + \alpha_t dt.$$

Note that in contrast to the EC MSR, there is neither an explicit nor an implicit announcement that those units removed from the allocation will be re-introduced to the market later-on, or (in case of additional allocation) vice-versa. With our mechanism, each adjustment to the allocation schedule carries with it an instantaneous deviation from the *ex-ante* cap at the same amount, although the *expected ex-post* cap may be altered by a different amount, subject to the market's endogenous reaction and future allocation adjustments. For example, an instantaneous allocation adjustment may remove some commodity units from the allocation, but future allocation adjustments are expected to re-introduce some of those units later-on. Accordingly, the time-t instantaneous change $d\mathbb{E}_t^{\mathbb{Q}}[A(0,T)]$ to expected total allocations may differ from $\delta \cdot |B_t - c| \; dt$ and is subject to the equilibrium dynamics. We will solve this interdependency in Proposition 29, where we derive the aggregate bank's dynamics in equilibrium, under the adjustment mechanism. This step in our analysis will prove fundamental to the quantification of aggregate expected compliance costs and hence for the design of an optimal policy, which we will discuss in Section 4.3.

For the purpose of demonstration, first consider the case of a rather large adjustment rate δ, close to 100% per time-unit. In this case, almost the entire difference $|B_t - c|$ is removed from or added to the allocation, which tightens the bank closely around the level c. In this case, the cap may have to be adjusted heavily, depending on the level of shocks to BAU deployment that the regulator has to compensate. Accordingly, we say that this case requires an *almost fully floating cap*. And since the price is effectively fixed, we are close to the pure price extreme of the policy spectrum.

Towards the other end of the policy spectrum lies the case where the adjustment rate δ is close to zero. Here, the allocation adjustments only partially compensate shocks to BAU deployment, and only to a small extend. The bank thus fluctuates relatively freely from the level c, and agent's abatement responds heavily to changes in BAU deployment. Accordingly, those shocks are almost fully reflected in the market price. The small volume of allocation adjustments requires the regulator to alter the ex-ante cap only slightly and thus the policy instrument is close to a *pure quantity instrument*, such as a 'pure' emissions trading system (e.g. the EU ETS before the introduction of additional measures like the MSR). When

[3] When discretising the model, the time increment dt may be replaced by some fixed timespan (e.g. a year) such that the following exemplary intuition holds: δ would correspond to the percentage of $|B_t - c|$ that is added to or removed from next year's allocation. In our continuous-time setting, δ is simply the adjustment rate of the regulator's allocation.

$\delta = 0$, there are no adjustments to the allocation and, accordingly, no adjustments to the cap, which simply results in a pure quantity instrument.

For a given adjustment rate level $\delta \geqslant 0$, consider the aggregate abatement in equilibrium:

$$\alpha_t = \mu e^{\mu t} \frac{R_0(\delta)}{e^{\mu T} - 1} + \mu e^{\mu t} \int_0^t \frac{d\xi_s(\delta)}{e^{\mu T} - e^{\mu s}} \tag{4.1}$$

Recall that the process ξ represents changes to the agents' expectations of BAU deployment versus allocations. With the responsive mechanism in place, ξ is subject to the adjustment rate δ, and the abatement curve responds to the mechanism, the allocation adjustments of which, in turn, respond to the aggregate abatement curve. (We will solve this inter-dependency in Proposition 29.) Consider the first term on the right-hand side of Equation (4.1), where $R_0(\delta)$ is the expected total required abatement, given only the information present at time 0. If no unanticipated events were to occur (that is to say, if $\xi \equiv 0$) the second term in Equation (4.1) would vanish and agents would simply spread their abatement effort over time, taking into account the cost differences imposed by the interest rate μ.

When unanticipated changes to future required abatement occur, the respective information is continuously incorporated into agents' abatement behaviour, as represented by the second term on the right-hand side of Equation (4.1). At each time $s \in [0, t]$, any changes $d\xi_s$ to expected required abatement are spread over time, taking into account the interest rate μ.

Recall that unanticipated changes to the required abatement, given by $d\xi_s = d\mathbb{E}_s^{\mathbb{Q}}[E(0,T) - A(0,T)]$, incorporate both changes to the expected total BAU deployment $\mathbb{E}_s^{\mathbb{Q}}[E(0,T)]$, as well as changes to the expected total allocations $\mathbb{E}_s^{\mathbb{Q}}[A(0,T)]$. Obviously, when changes in expected total BAU deployment were perfectly compensated by matching changes in total expected allocations (i.e. in the hypothetical[4] case $\delta \to \infty$), there would be no changes to the expected required abatement. This case would correspond to a perfectly responsive mechanism: The term $d\xi_s$ would vanish and the abatement would simply increase at the rate μ, just like in the deterministic case. And since, in equilibrium, the market price equals marginal abatement costs, the price would remain constant in discounted terms. If such a perfect compensation of $d\mathbb{E}_s^{\mathbb{Q}}[E(0,T)]$ by $d\mathbb{E}_s^{\mathbb{Q}}[A(0,T)]$ would be present throughout the entire regulated period $[0, T]$, the mechanism would be fully responsive and would, in fact, realise a pure price instrument, since the market would effectively fix the price without any deviations from external shocks. In terms of responsiveness, the mechanism would thus lie on what we refer to as the *price extreme* of a policy spectrum, along which different degrees of responsiveness may be realised. On the other end of this spectrum lies a pure quantity instrument, where $\delta = 0$ and the cap is never relaxed or tightened; and where , accordingly, there is no adjustment to expected total allocations $\mathbb{E}_s^{\mathbb{Q}}[A(0,T)]$. In this case, the price is effectively allowed to float freely, as the adjustment to shocks is fully borne by the agents, through their adjustment in abatement.

Notice that a responsive mechanism can compensate some of the shocks to BAU deployment and thus may reduce costs of adjustment on behalf of the agents. However, it also limits the temporal flexibility provided to agents and may therefore reduce the cost-efficiency of a quantity instrument such as emissions trading systems. In the extreme case of a fully responsive mechanism, the market priced is effectively fixed, turning the system into a de-facto tax. In what follows, we will examine this trade-off more closely and show how it can be solved to the end of inferring a mechanism that is optimal in terms of agents' expected aggregate total abatement costs.

The Bank under the Responsive Mechanism

We will now derive a closed-form expression for the time-t (aggregate) bank of unused commodity units.

The Bank under the Responsive Mechanism in Closed-Form

Proposition 29. *The time-t bank under the responsive mechanism is given in closed-form by*

$$B_t = B_0 e^{-\delta t} + \frac{\mu(e^{\mu t} - e^{-\delta t})}{(\delta + \mu)(e^{\mu T} - 1)} R_0 - \frac{e^{\mu t}}{V_t(\delta, r)} \int_0^t e^{-\mu s}\, V_s(\delta, r) \kappa_s \, dW_s^{\mathbb{Q}} + \int_0^t e^{\delta(s-t)} (f_s - g_s + \delta c)\, ds,$$

where

$$R_0 = \mathbb{E}_0^{\mathbb{Q}}[Y(0,T)] = -\frac{(\delta + \mu)(e^{\mu T} - 1)}{\mu(e^{\mu T} - e^{-\delta T})} \left(B_0 e^{-\delta T} + \int_0^T e^{\delta(s-T)} (f_s - g_s + \delta c)\, ds - c \right).$$

Proof. For convenience, we define the process $\varepsilon = (\varepsilon_t)_{t \in [0,T]}$ by $d\varepsilon_t = \kappa_t dW_t^{\mathbb{Q}}$. We then have the dynamics of the aggregate bank given by

$$dB_t = f_t dt + \delta(c - B_t)dt - g_t dt - d\varepsilon_t + \alpha_t dt, \tag{4.2}$$

which is solved by the following expression for the time-t bank:

$$B_t = B_0 e^{-\delta t} + \int_0^t e^{\delta(s-t)} (\alpha_s + f_s + \delta c - g_s)\, ds - \int_0^t e^{\delta(s-t)}\, d\varepsilon_s \tag{4.3}$$

Notice that the above equation expresses B_t in terms of the (past and present) abatement $(\alpha_s)_{s \in [0,t]}$, which itself depends on the dynamics of the (expected) bank through $\mathbb{E}_s^{\mathbb{Q}}[Y(0,T)]$, $0 \leqslant s \leqslant t$. More precisely we have that

$$\alpha_t = \mu e^{\mu t} \frac{R_0}{e^{\mu T} - 1} + \mu e^{\mu t} \int_0^t \frac{d\mathbb{E}_s^{\mathbb{Q}}[Y(0,T)]}{e^{\mu T} - e^{\mu s}},$$

[4]We forego a rigorous treatment of the limit case in mathematical terms, since our interest lies on the spectrum of policies rather than the extreme case of a pure tax.

where, in turn, the dynamics of $d\mathbb{E}_s^{\mathbb{Q}}[Y(0,T)]$ are given by

$$d\mathbb{E}_t^{\mathbb{Q}}[Y(0,T)] = d\mathbb{E}_t^{\mathbb{Q}}\left[\int_0^T d\varepsilon_s - \int_0^T \delta B_s\, ds\right] = d\varepsilon_t + \delta d\mathbb{E}_t^{\mathbb{Q}}\left[\int_0^T B_s\, ds\right].$$

In order to obtain α_t and B_t in closed-form, we thus need to solve the above interdependencies between the processes α and B. In order to simplify notation, we first recall that

$$h_t = \frac{\mu e^{\mu t}}{e^{\mu T} - e^{\mu t}}.$$

Using this notation, the dynamics of aggregate abatement can be expressed as

$$d\alpha_t = \mu\alpha_t dt + h_t d\mathbb{E}_t^{\mathbb{Q}}[Y(0,T)]. \tag{4.4}$$

We also recall that by the terminal constraint, we have that $d\mathbb{E}_t^{\mathbb{Q}}[B_T] = 0$ is satisfied for all $t \in [0,T]$, which we apply to Equation (4.3). Isolating the term depending on α then yields

$$d\mathbb{E}_t^{\mathbb{Q}}\left[\int_0^T e^{\delta s}\alpha_s\, ds\right] = d\mathbb{E}_t^{\mathbb{Q}}\left[\int_0^T e^{\delta s}(g_s - f_s - \delta c)\, ds - B_0 + \int_0^T e^{\delta s}\, d\varepsilon_s\right] = e^{\delta t} d\varepsilon_t. \tag{4.5}$$

We can make use of the dynamics of α in equation (4.4) in order to establish:

$$\int_0^t \alpha_s e^{\delta s}\, ds = \frac{e^{\delta t}}{\delta + \mu}\left(\alpha_t - \alpha_0 e^{-\delta t} - \int_0^t e^{\delta(s-t)} h_s\, d\mathbb{E}_s^{\mathbb{Q}}[Y(0,T)]\right).$$

The above equation, together with Equation (4.5), then yields:

$$d\mathbb{E}_t^{\mathbb{Q}}[Y(0,T)] = \frac{V_t(\delta,\mu)}{h_t}\, d\varepsilon_t, \tag{4.6}$$

where $V_t(\delta,\mu) = (\delta+\mu)/(e^{(\delta+\mu)(T-t)} - 1)$. From this we finally obtain

$$B_t = B_0 e^{-\delta t} + \frac{\mu(e^{\mu t} - e^{-\delta t})}{(\delta+\mu)(e^{\mu T} - 1)} R_0 - \frac{e^{\mu t}}{V_t(\delta,\mu)} \int_0^t e^{-\mu s}\, V_s(\delta,\mu)\, d\varepsilon_s + \int_0^t e^{\delta(s-t)}(f_s - g_s + \delta c)\, ds. \tag{4.7}$$

Note that $R_0 = \mathbb{E}_0^{\mathbb{Q}}[Y(0,T)]$. Using the compliance condition $B_T = c$, we can now derive the expression for $\mathbb{E}_0^{\mathbb{Q}}[Y(0,T)]$:

$$\mathbb{E}_0^{\mathbb{Q}}[Y(0,T)] = -\frac{(\delta+\mu)(e^{\mu T} - 1)}{\mu(e^{\mu T} - e^{-\delta T})}\left(B_0 e^{-\delta T} + \int_0^T e^{\delta(s-T)}(f_s - g_s + \delta c)\, ds - c\right).$$

Thus, the time-t bank B_t is indeed determined in closed-form by Equation (4.7). □

Notice that by Equation (4.6), the proof of Proposition 29 yields as a by-product that $d\xi_t = \sigma_t dW_t^{\mathbb{Q}}$, where σ is deterministic and bounded with $\sigma_t = 0$ for $t \in [\theta, T]$. Therefore, under the responsive mechanism, the optimisation problem for each agent in the previous chapter is indeed valid and the equilibrium thus persists under this regulation.

In Figure 4.1 we showcase the relation between the adjustment rate and the aggregate bank under an exemplary parameterisation. The figure exemplifies how the aggregate bank is dispersed, according to its quantiles at 90% confidence level, with and without a positive adjustment rate. In both cases, we notice that by spreading their abatement effort over time, agents start to accumulate a bank of commodity units at the beginning of the regulated timeframe, and draw this bank down as time moves forward and the terminal constraint approaches. However, agents' banking behaviour is noticeably influenced by whether or not the adjustment mechanism is in place: In the left-hand diagram, the adjustment rate is set to zero. In this case, agents adjust their yearly abatement in response to uncompensated shocks, bearing the full costs of adjusting their strategy each time. Accordingly, the full extend of exogenous shocks is reflected in the realised banking curves, illustrated in the respective quantiles.

The right-hand diagram illustrates the case where the adjustment rate is set to 5% per time-unit. In this case, the mechanism adjusts the allocation up or down, based on the aggregate bank. The realised banking curves in each scenario are thus dispersed more tightly, illustrated by their quantiles. In terms of the confidence level, we can thus say that the bank is kept within a 'tighter' band. Furthermore, the agents bear lower costs of adjusting to exogenous shocks, since their effect is partially compensated through adjustments to the allocation (and the cap). Notice, however, that agents' banking behaviour is also affected by anticipated allocation adjustments, which influences the banking behaviour - even in the absence of shocks. The regulator thus interferes with agents' inter-temporal optimisation, potentially reducing the value of their inter-temporal flexibility. We will examine this phenomenon more closely in the next section, where we calculate total aggregate abatement costs as a function of the adjustment rate, in order to pick the most cost-efficient policy solution.

The Bank under the Responsive Mechanism

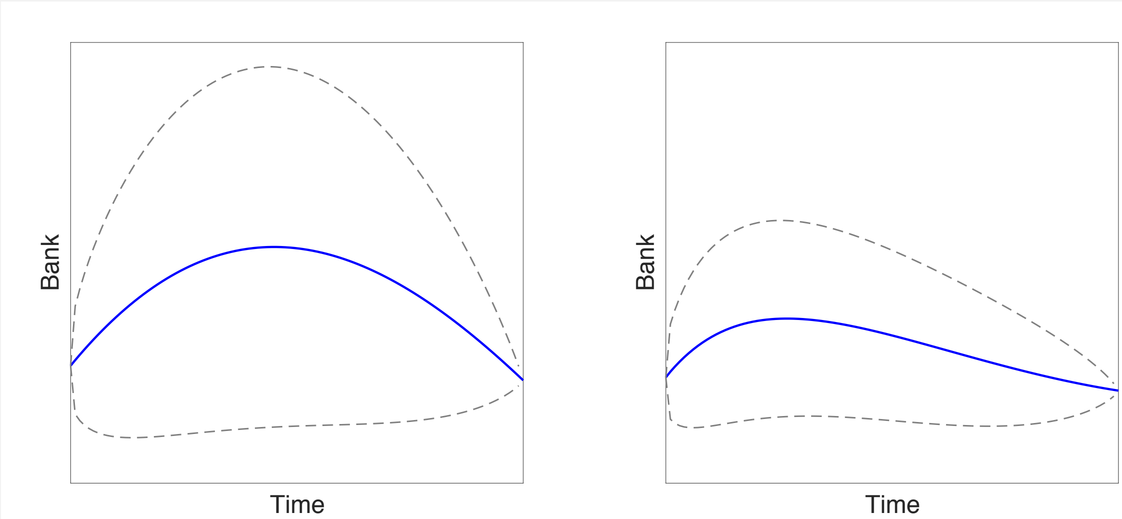

Figure 4.1: The figure shows the aggregate bank in a discretised setting. The left diagram pertains to the case where the responsive mechanism is deactivated, $\delta = 0$, whereas the right diagram shows the bank when $\delta = 5\%$ per time-unit. In each plot, the solid line shows the expected bank throughout $[0, T]$, where T is set to 30 time units. The figure is obtained for an initial bank of $B_0 = 2$Bn units, expected BAU deployment of 2Bn units per time-unit, and a constant allocation of 250Mn units per time-unit. The 'target' level c is set to 500Mn banked units. The dashed line in each plot shows the 5% and 95% quantiles for the aggregate bank. We assumed for both diagrams a standard deviation in BAU deployment of 200Mn units per time-unit. The abatement costs are parameterised by $Q_0 = 5$ Euros/unit and $\varrho = 0.25$ Euros per squared unit. The risk-free interest rate is set to $\mu = 3\%$ per time-unit.

Side-Note: Thresholds

As a side-note, we point out that in a real-world application of our stylised mechanism, the regulator may choose only to adjust the allocation schedule when the deviations $|B_t - c|$ exceed certain boundaries. For example, the regulator may feel inclined to avoid micro-managing the market, which may cause over-steering the bank or evoke a perceived increase in regulatory risk by over-regulation. Furthermore, as evidenced in the EU ETS, there may be boundaries for the bank in which the market is believed to "function orderly"[a] on behalf of the regulator or other stakeholders. Notwithstanding political motives or special interests in the level of such boundaries, we can make some observations on their general concept in the context of our model. In fact, using our framework, the implementation for such boundaries can be approximated with a confidence interval at some predetermined confidence level. More precisely, instead of the adjustment δ, the regulatory lever of control may be a confidence level such that some boundaries are respected by the bank with some 'acceptable' probability. Given a stochastic model for the process $B = (B_t)_{t\in[0,T]}$, and a supply adjustment mechanism such as above, this would then induce an adjustment rate δ. Implementing the generic mechanism above will then result in the predetermined boundaries being respected by the bank with an acceptable probability. In fact, in a real-world application, the regulator may choose to cease allocation adjustment when the number of banked commodity units lies within the desired interval. The regulator could formulate his mechanism using the interval bounds as threshold levels for the aggregate bank, which is congruent to the EC MSR's mode of action.

In our example, the stochastic part in the BAU deployment process is driven simply by a Brownian motion. From Proposition 29, we thus obtain that the Bank is normally distributed, $B_t \sim \mathcal{N}(a_t, b_t^2)$, where

$$a_t = B_0 e^{-\delta t} + \frac{\mu(e^{\mu t} - e^{-\delta t})}{(\delta + \mu)(e^{\mu T} - 1)} R_0 + \int_0^t e^{\delta(s-t)} (f_s - g_s + \delta c)\, ds$$

is the mean, and

$$b_t^2 = \frac{e^{2\mu t}}{V_t^2(\delta, \mu)} \int_0^t e^{-2\mu s}\, V_s^2(\delta, \mu)\, \kappa_s^2\, ds$$

is the variance. In the context of the 'acceptable boundaries' described above, let π_t denote the probability that the bank stays within the band $[l_t, u_t]$. We then have

$$\pi_t = \Phi\left(d_t^{(1)}\right) - \Phi\left(d_t^{(2)}\right),$$

where $\Phi(\cdot)$ represents the cumulative distribution function of the standard normal distribution and the parameters $d_t^{(1)}$ and $d_t^{(2)}$ are given by

$$d_t^{(1)} = \frac{u_t - a_t}{b_t} \quad \text{and} \quad d_t^{(2)} = \frac{l_t - a_t}{b_t}.$$

As anticipated, the probability π_t can now be expressed as a function $C_t(\delta) = \pi_t$. The respective bank interval can now be maintained with a confidence level π_t if the adjustment rate δ is set to $C_t^{-1}(\pi_t)$ and vice-versa.

[a] See [The European Commission, 2012].

4.3 Expected Total Abatement Costs and Policy Implications

The objective of this section is to show how the regulator can select an adjustment rate that minimises total aggregate abatement costs in expectation. Note that in the sense of regulatory optimisation, our analysis ignores the societal benefit of reducing the total aggregate deployment under the new regulation. That is, the benefit of implementing a specific expected ex-post cap is explicitly omitted in the optimisation problem and total economic costs are isolated. Our reason for doing so is based on our main application: Recall that CO_2, as well as comparable greenhouse gases are stock pollutants. As pointed out by [Newell and Pizer, 2008], the marginal benefit curve (of reducing emissions) is flat compared to the marginal (aggregate) abatement cost curve (as long as total emissions do not exceed the boundary of catastrophic implications): The benefit of reducing e.g. 1 billion tonnes of GHG *more* than necessary for the avoidance of catastrophic damages is negligible, compared to the costs of implementing that additional restriction.[5] Hence, maximising societal benefits minus economic costs can be well approximated by minimising only economic costs (given the above boundary). In the interest of tractability, we ignore the 'catastrophic' boundary condition, but point out that in a real-world application, some total cap should not be exceeded by allocation adjustments. The EC MSR, for example, implements this condition by ceasing additional allocations, once a previously filled reserve runs empty.

In order to illustrate how different adjustment rates affect expected total aggregate abatement costs, we first express those costs explicitly in Corollary 30. Note that we ignore total trading costs for simplicity since this would not add any qualitative value to our discussion later-on. Furthermore, we calculate the expectation of costs under the risk-neutral measure $\mathbb{Q}$, in order to first illustrate the effects of our mechanism irrespective of the level or functional representation of risk-aversion. In Section 4.3.2, we will then examine the effects of risk-aversion more closely by considering realised costs and risk-premia under different levels for δ.

Expected Total Aggregate Abatement Costs

Corollary 30. *Expected total aggregate abatement costs are given by*

$$\mathbb{E}^{\mathbb{Q}}\left[\int_0^T e^{-\mu t}\left(Q_t\alpha_t + \varrho\alpha_t^2\right)dt\right] = Q_0R_0 + \varrho\mu\frac{R_0^2}{e^{\mu T}-1} + \varrho\mu\int_0^T \frac{d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t}{\left(e^{\mu T}-e^{\mu t}\right)},$$

where $\langle\cdot\rangle$ refers to quadratic variation.

Proof. We first notice that we have

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[\int_0^T e^{-\mu t}\left(Q_t\alpha_t + \varrho\alpha_t^2\right)dt\right] &= \mathbb{E}^{\mathbb{Q}}\left[\int_0^T e^{-\mu t}Q_t\,\alpha_t\,dt\right] + \varrho\,\mathbb{E}^{\mathbb{Q}}\left[\int_0^T e^{-\mu t}\alpha_t^2\,dt\right]\\ &= Q_0\,\mathbb{E}^{\mathbb{Q}}\left[\int_0^T \alpha_t\,dt\right] + \varrho\,\mathbb{E}^{\mathbb{Q}}\left[\int_0^T e^{-\mu t}\alpha_t^2\,dt\right].\end{aligned} \tag{4.8}$$

Also, recall that we have for all δ

$$\int_0^t e^{\delta s}\alpha_s ds = \frac{1}{\delta+\mu}\left(e^{\delta t}\alpha_t - \alpha_0 - \int_0^t e^{\delta s}h_s d\mathbb{E}_s^{\mathbb{Q}}[Y(0,T)]\right).$$

We can use the above equation with $\delta = 0$ and $t = T$ to resolve the first term on the right-hand side of Equation (4.8):

$$\begin{aligned}Q_0\,\mathbb{E}^{\mathbb{Q}}\left[\int_0^T \alpha_t\,dt\right] &= \frac{Q_0}{\mu}\left(\mathbb{E}^{\mathbb{Q}}[\alpha_T] - \alpha_0 - \mathbb{E}^{\mathbb{Q}}\left[\int_0^T h_t d\mathbb{E}_t^{\mathbb{Q}}[Y(0,T)]\right]\right)\\ &= \frac{Q_0}{\mu}\left(\mathbb{E}^{\mathbb{Q}}[\alpha_T] - \alpha_0\right).\end{aligned}$$

[5] Note, however, that an adjustment rate that benefits agents by making the system more responsive *may*, lead to a lower ex-post cap. In contrast to societal benefits, the *economic* benefits in terms of cost reduction through responsiveness *are* included in our cost analysis.

By the expression for aggregate abatement we obtain

$$\mathbb{E}^{\mathbb{Q}}[\alpha_T] - \alpha_0 = \mu e^{\mu T}\frac{R_0}{e^{\mu T}-1} - \mu\frac{R_0}{e^{\mu T}-1} = \mu R_0.$$

We thus arrive at

$$Q_0\,\mathbb{E}^{\mathbb{Q}}\left[\int_0^T \alpha_t\,dt\right] = Q_0 R_0. \tag{4.9}$$

Regarding the second term on the right-hand side of Equation (4.8), we have

$$\int_0^T e^{-\mu t}\alpha_t^2\,dt = -\frac{1}{\mu}\int_0^T \alpha_t^2\,de^{-\mu t} = \frac{1}{\mu}\left(\alpha_0^2 - e^{-\mu T}\alpha_T^2 + \int_0^T e^{-\mu t}d\alpha_t^2\right).$$

We can use the fact that

$$\begin{aligned} d\alpha_t^2 &= 2\alpha_t d\alpha_t + d\langle\alpha\rangle_t = 2\alpha_t(\mu\alpha_t dt + h_t d\mathbb{E}_t^{\mathbb{Q}}[Y(0,T)]) + h_t^2 d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t \\ &= 2\mu\alpha_t^2 dt + 2\alpha_t h_t d\mathbb{E}_t^{\mathbb{Q}}[Y(0,T)] + h_t^2 d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t \end{aligned}$$

to deduce that

$$\begin{aligned} &\int_0^T e^{-\mu t}\alpha_t^2\,dt \\ =&\frac{1}{\mu}\left(\alpha_0^2 - e^{-\mu T}\alpha_T^2 + \int_0^T e^{-\mu t}2\mu\alpha_t^2 dt + \int_0^T e^{-\mu t}2\alpha_t h_t d\mathbb{E}_t^{\mathbb{Q}}[Y(0,T)] + \int_0^T e^{-\mu t}h_t^2 d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t\right) \\ =&\frac{1}{\mu}\left(\alpha_0^2 - e^{-\mu T}\alpha_T^2 + \int_0^T e^{-\mu t}2\alpha_t h_t d\mathbb{E}_t^{\mathbb{Q}}[Y(0,T)] + \int_0^T e^{-\mu t}h_t^2 d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t\right) + 2\int_0^T e^{-\mu t}\alpha_t^2 dt. \end{aligned}$$

This implies that

$$\int_0^T e^{-\mu t}\alpha_t^2\,dt = \frac{1}{\mu}\left(e^{-\mu T}\alpha_T^2 - \alpha_0^2 - \int_0^T e^{-\mu t}2\alpha_t h_t d\mathbb{E}_t^{\mathbb{Q}}[Y(0,T)] - \int_0^T e^{-\mu t}h_t^2 d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t\right)$$

and consequently

$$\varrho\,\mathbb{E}^{\mathbb{Q}}\left[\int_0^T e^{-\mu t}\alpha_t^2\,dt\right] = \frac{\varrho}{\mu}\left(e^{-\mu T}\left(\mathrm{Var}^{\mathbb{Q}}[\alpha_T] + \left(\mathbb{E}^{\mathbb{Q}}[\alpha_T]\right)^2\right) - \alpha_0^2 - \int_0^T e^{-\mu t}h_t^2 d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t\right),$$

where we used, in particular, that $d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t/dt$ is deterministic. Since $d\mathbb{E}_t^{\mathbb{Q}}[Y(0,T)]/dW_t$ is deterministic and bounded in $[0,T]$ we obtain

$$e^{-\mu T}\mathrm{Var}^{\mathbb{Q}}[\alpha_T] = e^{-\mu T}\mu^2 e^{2\mu T}\int_0^T \frac{d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t}{\left(e^{\mu T}-e^{\mu t}\right)^2} = \mu^2\int_0^T \frac{e^{\mu T}d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t}{\left(e^{\mu T}-e^{\mu t}\right)^2}.$$

Also notice that

$$\int_0^T e^{-\mu t}h_t^2 d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t = \mu^2\int_0^T \frac{e^{\mu t}d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t}{\left(e^{\mu T}-e^{\mu t}\right)^2}$$

and hence

$$e^{-\mu T}\mathrm{Var}^{\mathbb{Q}}[\alpha_T] - \int_0^T e^{-\mu t}h_t^2 d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t = \mu^2\int_0^T \frac{d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t}{\left(e^{\mu T}-e^{\mu t}\right)}.$$

Furthermore,

$$e^{-\mu T}\left(\mathbb{E}^{\mathbb{Q}}[\alpha_T]\right)^2 - \alpha_0^2 = e^{-\mu T}\mu^2 e^{2\mu T}\frac{R_0^2}{(e^{\mu T}-1)^2} - \mu^2\frac{R_0^2}{(e^{\mu T}-1)^2} = \mu^2\frac{R_0^2}{e^{\mu T}-1}.$$

Together, we find that

$$\varrho\,\mathbb{E}^{\mathbb{Q}}\left[\int_0^T e^{-\mu t}\alpha_t^2\,dt\right] = \varrho\mu\frac{R_0^2}{e^{\mu T}-1} + \varrho\mu\int_0^T \frac{d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t}{\left(e^{\mu T}-e^{\mu t}\right)}. \tag{4.10}$$

Equations (4.9) and (4.10) together yield

$$\mathbb{E}^{\mathbb{Q}}\left[\int_0^T e^{-\mu t}\left(Q_t\alpha_t + \varrho\alpha_t^2\right)dt\right] = Q_0R_0 + \varrho\mu\frac{R_0^2}{e^{\mu T}-1} + \varrho\mu\int_0^T \frac{d\langle\mathbb{E}^{\mathbb{Q}}[Y(0,T)]\rangle_t}{\left(e^{\mu T}-e^{\mu t}\right)},$$

which is what we wanted to prove. □

4.3.1 The Optimal Adjustment Rate

We will now use our closed-form expression for expected total aggregate abatement costs to discuss the trade-off the regulator faces when selecting a cost-minimising adjustment rate δ.

On the one hand, recall that due to exogenous shocks to BAU deployment, the agents continuously adapt their strategies, which increases their total costs. The total of necessary adjustments can be reduced by the implementation of a positive adjustment rate. And, with a very large adjustment rate, close to 100% per time-unit, the regulator can almost fully eradicate any variation in agents' expected abatement requirement, leading to a de-facto fixed price policy. It is therefore intuitive that the higher the adjustment rate, the lower the *cost of adjustments* on behalf of agents should be.

On the other hand, agents use their ability to bank commodity units throughout the regulated timeframe to minimise their expected total costs. The profitability or viability of inter-temporal cost-saving opportunities may be affected by a positive adjustment rate: The higher the adjustment rate, the 'tighter' the bank is forced to fluctuate around the level c, and inter-temporal cost-saving opportunities are more and more reduced as δ increases.

Consider the total costs derived in Corollary 30; i.e.

$$Q_0 R_0 + \varrho\mu \frac{R_0^2}{e^{\mu T} - 1} + \varrho\mu \int_0^T \frac{d\langle\xi\rangle_t}{\left(e^{\mu T} - e^{\mu t}\right)}, \tag{4.11}$$
$$\text{where} \quad \xi_t = \mathbb{E}_t^{\mathbb{Q}}[Y(0,T)].$$

Analytically, we can identify the above-mentioned trade-off in the interaction of the three summands on the right-hand side of Equation (4.11): The first two terms $Q_0 R_0$ and $\varrho\mu R_0^2/(e^{\mu T} - 1)$ quantify the costs in the absence of uncertainty. Herein we can, in particular, isolate any effects the adjustment rate has on total costs that are irrespective of responsiveness and its benefits. To that end, consider the initial expected abatement requirement $R_0 = \mathbb{E}_0\left[E(0,T) - A(0,T)\right] - B_0 + c$. First, consider the case where the initial bank B_0 is higher than c (e.g. $B_0 > 0$ and $c = 0$). In this case, if $\delta > 0$, the mechanism will, in expectation, decrease the total cap, compared to the case where $\delta = 0$ – and this effect will intensify with an increasing adjustment rate. Accordingly, the first term in Equation (4.11), $Q_0 R_0$, will increase due to an increasing initial abatement requirement R_0. Conversely, if $B_0 < c$, the cap will increase with an increasing $\delta > 0$, and the term $Q_0 R_0$ will decrease accordingly.

As for the second term on the right-hand side of Equation (4.11), i.e. $\varrho\mu R_0^2/(e^{\mu T} - 1)$, recall that the higher the adjustment rate, the tighter the bank is kept around c. Accordingly, inter-temporal optimisation opportunities will be affected since strategies that entail a large bank are no longer as profitable or viable. Since ϱ parameterises the costs associated to variations in abatement, agents have an incentive to spread their abatement across the regulated timeframe. This leads to the typical banking behaviour illustrated in Figure 4.1: Agents first accumulate commodity units and then steadily draw them down (cf. [Rubin, 1996] and [Schennach, 2000]). The associated cost-saving opportunities, however, are affected by a positive adjustment rate, since the benefits of such banking behaviour is partially offset by the mechanisms propensity to diminish a larger bank. Accordingly, the banking curve is skewed, compared to the case where $\delta = 0$, increasing costs quadratically whenever R_0 is increased by $\delta > 0$.

As for the third term on the right-hand side of Equation (4.11), recall that $d\langle\xi\rangle_t$ is the time-t quadratic variation of the expected total abatement requirement. Accordingly, $\varrho\mu \int_0^T \frac{d\langle\xi\rangle_t}{(e^{\mu T} - e^{\mu t})}$ captures costs due to agents' strategy adjustments. The regulator can reduce those costs by increasing the adjustment rate δ: The more responsive the system is, the smaller the total adjustments become on behalf of the agents. The variability in expected required abatement is reduced by an increasing δ, decreasing the above term in Equation (4.11).

As we can see, there is an inherent trade-off regarding expected aggregate total abatement costs when selecting an adjustment rate. We illustrate this trade-off in Figure 4.2, which shows expected aggregate total abatement costs as a function of the adjustment rate, computed for a set of illustrative parameters.

Consider the results shown in Figure 4.2. For an adjustment rate of zero (on the left end of the horizontal axis), when the cap is unaffected, the policy instrument lies at the fixed-cap end of the spectrum, as implemented in the former EU ETS setup, before recent reforms took place. At this extreme, agents spread their abatement effort over time, taking into account the risk-free rate μ. The presence of exogenous shocks, however, inflict costs through the adjustments that agents make in their response. When the adjustment rate is gradually increased, these adjustments decrease, which in turn reduces costs through strategy adjustments. Towards the fixed-price end of the spectrum (towards the right end of the horizontal axis), however, costs start to increase again: The costs of foregone cost savings from inter-temporal cost-saving opportunities surpass the benefits from the mechanism's responsiveness. The loss of benefits from exploiting differences in marginal abatement costs across time exceed the agents' cost savings caused from shock-mitigation by the mechanism.

Expected Total Aggregate Abatement Costs versus the Adjustment Rate

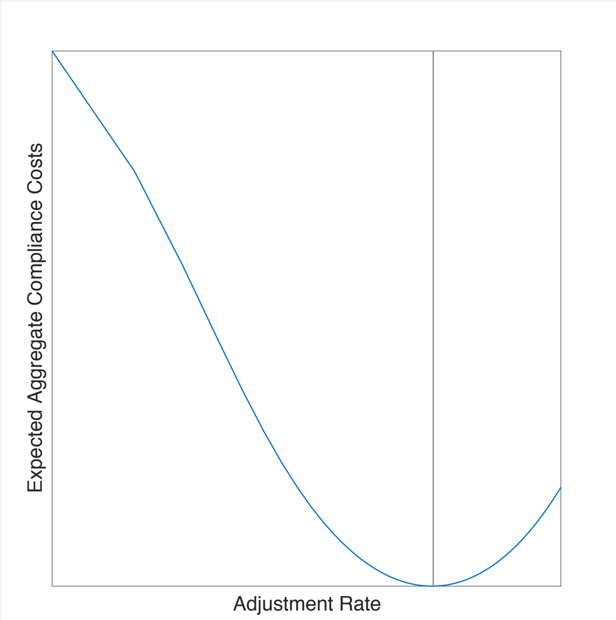

Figure 4.2: Expected aggregate total abatement costs as a function of the adjustment rate between 0 (left end of the spectrum) and 100% per time-unit (right end of the spectrum), obtained for an initial bank of $B_0 = 2$Bn units, an expected BAU deployment of 2Bn tonnes per time-unit and a constant allocation of 250Mn units per time-unit; the standard deviation in BAU deployment is 200Mn units per time-unit, and $c = 0$. The time horizon is set to 30 time-units. Abatement costs are parametrised by $Q_0 = 5$ Euros/unit and $\varrho = 0.25$ Euros per squared units. The risk-free interest rate is $\mu = 3\%$. **The figure shows the curve on a logarithmic scale on both axes and is normalised to values between** 0 **and** 1.

Note that although the mechanism will tend to decrease the total number of commodity units allocated (since $B_0 > 0$ and $c = 0$), and this effect intensifies with δ increasing from zero, there is at first a steep decline in expected aggregate total abatement costs towards the left-hand side of the figure.

4.3.2 The Effects of the Adjustment Rate on the Risk Premium

In the previous section, we demonstrated how the adjustment rate δ can be used to stipulate the degree by which the burden of adjusting to shocks is borne by agents as opposed to the regulator. With a low adjustment rate on the one end of the policy spectrum, agents may need to adjust their banking behaviour heavily in response to changes in their BAU deployment, which drives up their costs based on their convex cost structure. In this case, the ex-post cap will be relatively close to its ex-ante plan (high cap stringency). This case corresponds to a policy that is relatively near (or identical, if $\delta = 0$) to a pure quantity instrument. With a high adjustment rate on the other end of the spectrum, a large portion of the shocks to BAU deployment are offset by allocation adjustments, which may require the regulator to heavily adjust the cap ex-post (low cap stringency). This policy is relatively near to a pure price instrument, since the price variability reflects the variability in aggregate abatement. By setting an adjustment rate $\delta > 0$, the regulator can thus curb the price variability – and hence reduce the risk associated with investing in the commodity.

When investments in the commodity are uncertain, investors are only willing to purchase units of the commodity if the rate of return exceeds the risk-free rate μ: Otherwise, investing in the risk-free asset will always be preferred by any risk-averse agent. By the natural no-arbitrage assumption, we should therefore see a positive risk-premium ζ_t on top of the risk-free rate μ. Note that future abatement cost reduction constitutes a payoff from investing in the commodity, the same way speculative investment yields a (potential) payoff. Hence, future revenue streams from re-selling the commodity and future compliance costs should both be discounted at the same (risk-adjusted) discount rate $\mu + \zeta_t$. Alternative investments may promise higher (discounted) returns which would lead agents to postpone their abatement and instead deploy part of their current bank of commodity units. Notice in particular, that a high risk-premium should thus lead to postponed abatement, which would skew the banking curve away from the most cost-efficient equilibrium pattern. This phenomenon has in fact been observed in the literature before, e.g. by [Fell, 2015] and [Ellerman et al., 2015b].

As set out in Section 3.7 of the previous chapter, it is natural to assume that the risk-premium ζ_t is a monotonically increasing function of the volatility of commodity prices. And as we have set out in previous paragraphs, the variability in abatement, and hence in commodity prices decreases in the adjustment rate δ. Hence we should observe that the higher the adjustment rate, the lower the risk premium becomes. In order to quantify the impact of δ on ζ_t, first formally consider the time-t commodity price return in terms of $W_t^{\mathbb{Q}}$:

$$\frac{dP_t}{P_t} = \mu dt + \frac{2\varrho V_t(\delta,\mu)\kappa_t}{P_t} dW_t^{\mathbb{Q}} \tag{4.12}$$

As we can see, the return (per time-unit) is comprised of the risk-free rate μ, plus a stochastic component which reflects the uncertainty of future required abatement. Under the risk-neutral measure $\mathbb{Q}$, we can thus (heuristically) express the expected rate of return by $\mathbb{E}^{\mathbb{Q}}[dP_t/P_t] = \mu dt$, which reflects the $\mathbb{Q}$-strong no-arbitrage condition (i.e. not even $\mathbb{Q}$-weak arbitrage opportunities in the form of outperforming strategies exist). However, under the objective measure $\mathbb{P}$, the expected rate of return should include a risk-premium ζ_t and should thus have a heuristic representation of the form $\mathbb{E}^{\mathbb{P}}[dP_t/P_t] = (\mu + \zeta_t)dt$.

Recall that in order to keep our analysis simple and transparent, we assumed in Section 3.7 that the risk-premium increases linearly in the volatility coefficient of commodity price returns. That is, we assume that ζ_t is proportional to the volatility coefficient in Equation (4.12):

$$\zeta_t = k \cdot \frac{2\varrho V_t(\delta, \mu)\kappa_t}{P_t} \tag{4.13}$$

Here, the constant k represents the (overall) level of the agents' risk-aversion. We recall that the function $V_t(\delta, \mu)$ is given by

$$V_t(\delta, \mu) = \frac{\delta + \mu}{e^{(\delta+\mu)(T-t)} - 1}. \tag{4.14}$$

Note that in the context of Figure 4.3, the parameter k in Equation (4.13) is simply used to conveniently normalise ζ_t to values between 0 and 1.

By Section 3.7 of the previous chapter, we thus formally obtain the following commodity price return (per time-unit) under the objective measure:

$$\frac{dP_t}{P_t} = (\mu + \zeta_t)dt + \frac{2\varrho V_t(\delta, \mu)\kappa_t}{P_t} dW_t^{\mathbb{P}} \tag{4.15}$$

And we indeed have the heuristics $\mathbb{E}^{\mathbb{P}}[dP_t / P_t] = (\mu + \zeta_t)dt$.

From Equations (4.13) and (4.14) we can infer the impact of the adjustment rate δ on the risk-premium ζ_t. To this end, notice that the denominator on the right-hand side of Equation (4.14) increases exponentially in δ, while the numerator increases only linearly. Accordingly, a high adjustment rate entails a small value for $V_t(\delta, \mu)$.

Thus, a high adjustment rate, corresponding to a very responsive mechanism, induces a small risk-premium in Equation (4.13). In this case, the commodity price increases at a rate slightly above the risk-free rate μ. Notice that this is in line with the intuition that an extremely responsive mechanism is close to equivalent to a fixed-price policy – a de-facto tax that increases at the risk-free rate (and thus is, in discounted terms, constant).

Conversely, a small adjustment rate δ induces a high value for $V_t(\delta, \mu)$, which in turn entails a high risk-premium. When the policy is only slightly or not at all responsive, all shocks to BAU deployment are fully borne by the agents through their abatement adjustments. These adjustments then drive price volatility, rendering investments in the commodity relatively risky.

In Figure 4.3, we illustrate the relation between the adjustment rate δ, the risk-premium ζ_t, and the aggregate total abatement costs by plotting the respective results for a number of simulated BAU deployment scenarios under different adjustment rates.

Realised Total Aggregate Abatement Costs versus Risk-Premia for Different Adjustment Rates

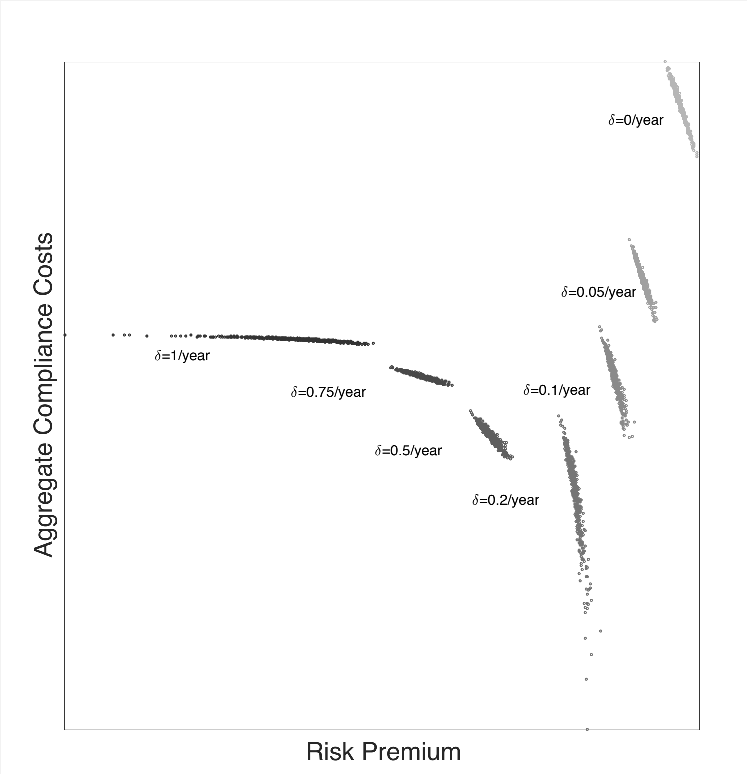

Figure 4.3: Realised total aggregate abatement costs and risk-premia for 500 simulated BAU deployment paths under different adjustment rates in a discretised setting; obtained for an initial bank of $B_0 = 2$Bn units, an expected BAU deployment of 2Bn units per time-unit and a constant allocation of 250Mn units per time-unit; the standard deviation in BAU deployment is 200Mn units per time-unit, and $c = 0$. The time horizon is set to 30 time-units. Abatement costs are parameterised by $Q_0 = 5$ Euros/unit and $\varrho = 0.25$ Euros per squared unit. The risk-free interest rate is $\mu = 3\%$. **The figure shows a scatterplot on a logarithmic scale on both axes, where both dimensions are normalised to values between 0 and 1.**

We simulate 500 BAU deployment paths on a discretised time-grid, in order to illustrate the impact of δ on risk-premia and costs. Each dot in Figure 4.3 represents the outcome of one simulation, where we selected seven different adjustment rates to illustrate the effect of the mechanism across the policy spectrum. The adjustment rate increases from $\delta = 0$ (with respective results appearing in light grey in the upper right-corner) to 100% per time-unit[6] (with results appearing in dark grey towards the left-hand side of the Figure). The realised aggregate abatement costs (vertical axis) and risk-premia (horizontal axis) for each adjustment rate can be identified as distinct clouds of scattered points across the figure, highlighted in different shadings. Note that the figure shows a plot with logarithmic scales on both axes. Accordingly, the results for lower adjustment rates are scattered farther apart than for higher adjustment rates, which corresponds to the fact that the variability in risk-premia and compliance costs diminishes with increasing responsiveness.

Figure 4.3 confirms our previous intuition behind the decomposition of total aggregate abatement costs: The riskiness of investing in the commodity decreases monotonically with an increasing responsiveness. Accordingly, the cost of adjusting to shocks on behalf of the agents is reduced when increasing δ. However, this increase in responsiveness comes at the cost of thinning out the profitability of inter-temporal cost-saving opportunities, and, in this example, an overall decreased cap. In this example we see that, beyond an adjustment rate of about 20% per time-unit, these effects override the benefits associated to the reduced uncertainty regarding the expected abatement requirement.

4.4 Conclusions

By means of our generic responsive mechanism, we examined a spectrum of policies between a pure quantity instrument (such as emissions trading systems with a fixed cap) on the one hand, and pure price instruments (such as fixed taxes or fixed tax trajectories) on the other hand. Each point on the policy spectrum is characterised by a level of cap stringency, i.e. the degree of liberty by which the regulator allows the cap to respond to changes in the system due to shocks in the commodity demand. We made this characterisation possible by employing an adjustment rate δ, which parameterises the strength of the regulatory response. In that sense, this parameter δ represents the converse notion to cap stringency, since the

[6] Note that in the discretised setting, allocation adjustments at each point in time are effective only at the respective next point in time. Hence, even for an adjustment rate of 100% per time-unit, some uncertainty remains.

determination to carry out strong responses may require substantial deviations from the ex-ante cap, whereas responding only weakly tends to require only minor deviations.

The system dynamics, which are subject to both uncertainty and the responsive mechanism, depend upon the reciprocal responses from the market and the regulator through time. Since the agents' optimisation problem depends on the expected future development of the system state, the problem of finding the equilibrium dynamics is subject to the complexities of solving the above reciprocity. In particular, when agents are risk-averse, the stochastic behaviour of the entire system becomes an intriguing topic. In the present work, we solved this inter-dependency and obtained the system's $\mathbb{Q}$-strong equilibrium dynamics in closed form. With the price dynamics readily available, we then show that the obtained equilibrium is also $\mathbb{P}$-weak, and the risk-premia obtained under an exemplary linearity assumption (with respect to the functional representation of risk-aversion) are obtained as a by-product. Finally, we demonstrated that by choosing an optimal adjustment rate, the regulator can maximise the cost-efficiency of the system in a stochastic setup.

Bibliography

[Bain, 1948] Bain, J. S. (1948). Output quotas in imperfect cartels. *The Quarterly Journal of Economics*, 62(4):617–622.

[Bertsekas and Shreve, 1978] Bertsekas, D. P. and Shreve, S. E. (1978). *Stochastic optimal control: The discrete time case*, volume 23. Academic Press New York.

[Doda, 2016] Doda, B. (2016). How to price carbon in good times... and bad! *WIREs Climate Change*, 7(7):135–144.

[Ellerman et al., 2015a] Ellerman, D., Valero, V., and Zaklan, A. (2015a). An analysis of allowance banking in the EU ETS. EUI Working Paper RSCAS 2015/29.

[Ellerman et al., 2015b] Ellerman, D., Valero, V., and Zaklan, A. (2015b). An Analysis of Allowance Banking in the EU ETS. EUI Working Paper RSCAS 2015/29.

[Ellerman and Wing, 2003] Ellerman, D. and Wing, I. (2003). Absolute Versus Intensity-Based Emission Caps. *Climate Policy*, 3(1):7–20.

[Fell, 2015] Fell, H. (2015). Comparing Policies to Confront Permit Over-allocation. Unpublished working paper.

[Fell and Morgenstern, 2010] Fell, H. and Morgenstern, R. (2010). Alternative Approaches to Cost Containment in a Cap-and-Trade System. *Environmental and Resource Economics*, 47(2):275–297.

[Fell et al., 2012] Fell, H., Morgenstern, R., and Palmer, K. (2012). Soft and hard price collars in a cap-and-trade system: A comparative analysis. *Journal of Environmental Economics and Management*, 64:183–198.

[Frino et al., 2010] Frino, A., Kruk, J., and Lepone, A. (2010). Liquidity and Transaction Costs in the European Carbon Futures Market. *Journal of Derivatives and Hedge Funds*, 16(2):100–115.

[Golosov et al., 2014] Golosov, M., Hasslet, J., Krussell, P., and Tsyvinski, A. (2014). Optimal Taxes on Fossil Fuel in General Equilibrium. *Econometrica*, 82:41–88.

[Grafton, 1996] Grafton, R. Q. (1996). Individual transferable quotas: theory and practice. *Reviews in Fish Biology and Fisheries*, 6(1):5–20.

[Grosjean et al., 2014] Grosjean, G., Acworth, W., Flachsland, C., and Marschinski, R. (2014). After monetary policy, climate policy: is delegation the key to EU ETS reform? *Climate Policy*, 1:1–25.

[Grüll and Taschini, 2011] Grüll, G. and Taschini, L. (2011). Cap-and-Trade Properties under Different Hybrid Scheme designs. *Journal of Environmental Economics and Management*, 61(1):107–118.

[Hartley and Hartley, 2002] Hartley, J. E. and Hartley, J. E. (2002). *The representative agent in macroeconomics*. Routledge.

[Hepburn, 2006] Hepburn, C. (2006). Regulation by Prices, Quantities, or Both: A Review of Instrument Choice. *Oxford Review of Economic Policy*, 22(2):226–247.

[Heutel, 2012] Heutel, G. (2012). How should environmental policy respond to business cycles? Optimal policy under persistent productivity shocks. *Review of Economic Dynamics*, 15:244–264.

[Hotelling, 1931] Hotelling, H. (1931). The economics of exhaustible resources. *Journal of political Economy*, 39(2):137–175.

[Jotzo and Pezzey, 2007] Jotzo, F. and Pezzey, J. (2007). Optimal intensity targets for greenhouse gas emissions trading under uncertainty. *Environmental and Resource Economics*, 38(2):259–284.

[Kollenberg and Taschini, 2016a] Kollenberg, S. and Taschini, L. (2016a). Dynamic supply adjustment and banking under uncertainty: The market stability reserve. *Available at SSRN*.

[Kollenberg and Taschini, 2016b] Kollenberg, S. and Taschini, L. (2016b). Emissions trading systems with cap adjustments. *Journal of Environmental Economics and Management*, 80:20–36.

[Krylov, 2008] Krylov, N. V. (2008). *Controlled diffusion processes*, volume 14. Springer Science & Business Media.

[Landis, 2015] Landis, F. (2015). Interpolation of Abatement Cost According to PACE. Unpublished working paper.

[Medina et al., 2014] Medina, V., Pardo, A., and Pascual, R. (2014). The Timeline of Trading Frictions in the European Carbon Market. *Energy Economics*, 42(3):378–394.

[Newell and Pizer, 2008] Newell, R. G. and Pizer, W. A. (2008). Regulating Stock Externalities under Uncertainty. *Journal of Environmental Economics and Management*, 56:221–233.

[Perino and Willner, 2015] Perino, G. and Willner, M. (2015). The Price and Emission Effects of a Market Stability Reserve in a Competitive Allowance Market. WiSo-HH Working Paper Series, No. 28.

[Pham, 2009] Pham, H. (2009). *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science & Business Media.

[Pizer, 2002] Pizer, W. A. (2002). Combining Price And Quantity Controls To Mitigate Global Climate Change. *Journal of Public Economics*, 85:409–434.

[Poundstone, 1992] Poundstone, W. (1992). Prisoner's dilemma: John von neuman, game theory, and the puzzle of the bomb.

[Roberts and Spence, 1976] Roberts, M. J. and Spence, M. (1976). Effluent charges and licenses under uncertainty. *Journal of Public Economics*, 5(3):193–208.

[Rubin, 1996] Rubin, J. D. (1996). A model of intertemporal emission trading, banking, and borrowing. *Journal of Environmental Economics and Management*, 31:269–286.

[Schennach, 2000] Schennach, S. M. (2000). The economics of pollution permit banking in the context of title iv of the 1990 clean air act amendments. *Journal of Environmental Economics and Management*, 40:189–210.

[Teeguarden, 1969] Teeguarden, D. E. (1969). Allocating public timber by transferable purchasing quotas: A policy analysis. *Natural Resources Journal*, 9(4):576–589.

[The European Commission, 2012] The European Commission (2012). Report from the Commission to the European Parliament and the Council: The state of the European carbon market in 2012 . Technical report, The European Commission, Brussels.

[The European Parliament and Council, 2015] The European Parliament and Council (2015). Decision (EU) 2015/1814 of the European Parliament and of the Council of 6 October 2015 concerning the establishment and operation of a market stability reserve for the Union greenhouse gas emission trading scheme and amending Directive 2003/87/EC.

[Unold and Requate, 2001] Unold, W. and Requate, T. (2001). Pollution control by options trading. *Economics Letters*, 73:353–358.

[Weitzman, 1974] Weitzman, M. L. (1974). Prices vs. Quantities. *Review of Economic Studies*, 41(4):683–691.