

QUANTITATIVE ENVIRONMENTAL ECONOMICS
MODELING MARKETABLE PERMITS IN DISCRETE
AND CONTINUOUS TIME

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF
MATHEMATICS, STATISTICS AND INFORMATION TECHNOLOGY
AND THE COMMITTEE ON GRADUATE STUDIES OF
UNIVERSITY OF BERGAMO
IN TOTAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Luca Taschini

October 2008

© Copyright by Luca Taschini 2008

All Rights Reserved

Abstract

Environmental policy instruments, such as marketable permits, exist to help monitor and regulate environmental practices of organizations, i.e. companies and institutions (see [60], [77] and discussion in Chapter 2). Market-based instruments are already employed for the implementation of environmental policies on European scale (European Emission Trading Scheme - EU ETS) and on global scale (Kyoto protocol). In an effort to bridge the gap between the theoretical emission permit price and observed market-price behavior, we investigate the historical time series of the marketable permit price. More precisely, in Chapter 3 we advocate the use of a new GARCH-type structure for the analysis of inherent heteroskedastic dynamics in the returns of SO_2 in the U.S. and of CO_2 emission permits in the EU ETS. In Chapter 4 we show that the presence of asymmetric (or incomplete) information plays a central role. In other words, market-prices of permits are affected by the different information sets based on which market-players found their financial and investment strategies. A CO_2 -option pricing model comparison is developed in Chapter 4.7. The option pricing method can be used for hedging purposes and for pricing CO_2 -linked projects and investments.

Acknowledgments

"We learn more by looking for the answer to a question and not finding it than we do from learning the answer itself."

-Lloyd Alexander

This thesis has benefited from the interaction of three excellent environments.

The first one was the Institute of Economy at the University of Bergamo, where I had the opportunity to start my PhD in computational methods in October 2004 at the department of Mathematics, Statistics and Information Technology. My warmest thanks go to Prof. Rosella Giacometti, Prof. Sergio Ortobelli, Prof. Svetlozar Rachev, Prof. Andrea Resti, Prof. Vittorio Moriggia and Prof. L. Bertoli Barsotti for enabling a simultaneous research and learning work in Bergamo. I want to thank Prof. Marida Bertocchi for her perseverance in convincing me to pursue a one-year visiting at the University of Zürich. I would not be where I stand now without here advices. I would also thank my former classmates Marianna Brunetti, Domenico De Giovanni, Simona Gambaro, Davide Orlandini ed Emilio Russo for the useful discussions and the unforgettable nights spent studying together. The financial support from Credito Bergamasco (Banca Popolare di Verona e Novara) is gratefully acknowledged.

The second important stimulating environment is the Swiss Banking Institute at the University of Zürich. The largest part of this work was carried out here at the Chair of Quantitative Finance. I would like to express my gratitude to the Zürich-based supervising professor Marc Chesney for his excellent guidance and comments during the work. I want to thank Prof. Marc Paoletta and Prof. Pauline Barrieu for their help and contribution to our joint works. Also, I thank Prof. Rajna Gibson Brandon, Prof. Thorsten Hens, Prof. Alexander Wagner, and Prof. Michael Habib for their advice and support during the writing of this thesis. My gratitude goes also to Prof. Walter Farkas, being his assistant in the lecture "*Introduction to Mathematical Finance*" (Master of Advanced Studies in Finance at ETH - Zürich) was both an exiting and threatening experience for me! Discussions with classmates and colleagues of the Institute were always both provocative and stimulating. A special thank goes to Urs Schweri for his enormous patience in listening and transmitting me his outstanding computational skills. The financial support from the University Research Priority Program "Finance and Financial Markets" and by the National Centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK), respectively research instruments of the University of Zürich and of the Swiss National Science Foundation are gratefully acknowledged.

Finally, I want to thank my family and friends for their support during the dissertation process. My parents, my sister, my brother, and Federica's family have always encouraged me in pursuing my way. My deepest gratitude goes to Federica for her love, patience, perseverance, and devotion. She moved to Switzerland renouncing all she had and made my research possible.

Contents

Abstract	V
Acknowledgments	VI
1 Summary	1
2 Environmental Economics	5
2.1 Environmental Economics: A Primer	6
2.2 Marketable Permits Systems and Effectiveness	13
2.3 Modeling Marketable Permits	17
3 Econometric Analysis of the Marketable Permits	21
3.1 An Econometric Approach for Analyzing the SO ₂ Data	23
3.1a Illiquidity	23
3.1b Basic Analysis	24
3.1c Stable-GARCH Model	27
3.1d Mixed Normal GARCH: Model and Numerical Issues	30
3.1e Mixed Normal GARCH: Estimation Results and Diagnostics	35
3.2 Analysis of CO ₂ Returns	40
3.3 Conclusions	46

4	Continuous Time Model: An Equilibrium Approach	48
4.1	Marketable Permits for Air-control	51
4.2	Abatement Opportunities in the Short Term	53
4.3	The Formal Model: "Wait-and-see" for One Company	56
4.4	Two-companies and Multi-periods Trading	61
4.5	Multi-firm and Multi-periods Trading	67
4.6	Numerical Evaluation	68
4.7	CO ₂ Option Pricing	73
4.8	CO ₂ Option Pricing: A Comparison	76
	Appendix A	
	81
A.1	Appendix A.1	81
A.2	Appendix A.2	82
A.3	Appendix A.3	83
A.4	Appendix A.4	86
A.5	Appendix A.5	88
	Bibliography	90

List of Figures

3.1	Daily SO ₂ returns (top), and the SACF of the zeros-removed absolute returns (bottom).	26
3.2	Kernel density (solid) of the SO ₂ return series, with the best-fitting normal density (dashed) and best-fitting symmetric stable density (dash-dot). Right panel is just the magnified view of the right tail.	27
3.3	The p -values from the runs-test performed on segments of the SO ₂ return series. The first segment is the returns in the whole series, the second is from the second return to the end, etc., up to the $(T - 50)$ th observation to the end.	28
3.4	QQ-plot of a simulated time series versus the return series for the SO ₂ price index.	38
3.5	Deviation probability plot; solid is for t -GARCH, dashed is for MixN(3,2). Plotted values are “Deviation” := $100(F_U - \hat{F})$ (vertical axis) versus $100\hat{F}$ (horizontal axis), where F_U is the cdf of a uniform random variable; \hat{F} refers to the empirical cdf formed from evaluating the 1,280 one-step, out-of-sample distribution forecasts at the true, observed, next-day return. . . .	40
3.6	Daily CO ₂ returns	41

3.7	Deviation probability plot for CO ₂ returns based on moving window of size 250; see caption in Figure 3.5 for description. Solid is for t -GARCH, dashed is MixN(3,2), dash-dot is GAt-GARCH, and dotted is GAt-GARCH using a weighted likelihood.	42
4.1	Plot of S_0 , the emission permit price at initial time, as a function of X_0 , the amount of permits sold (negative values) or bought (positive values) at time $t = 0$. We plot the permit price for different $\{\mu : \mu \in \mathbb{R}\}$, left picture, and different $\{\sigma : \sigma \in \mathbb{R}^+\}$, right picture, keeping all other parameters constant. In this example, $N = 170$, $P = 40$ and $Q_0 = 100$	60
4.2	\bar{S}_t permit price evolution (bottom-part) for the pollution parameters $\mu = (0.15; 0.10)$, $\sigma = (0.10; 0.10)$, $Q_0 = (50; 25)$, $N_0 = (52; 25)$. The simulated pollution processes are depicted in the upper ($Q_{1,t}$) and middle-part ($Q_{2,t}$).	69
4.3	\bar{S}_t permit price evolution (bottom-part) for the pollution parameters $\mu = (-0.15; 0.001)$, $\sigma = (0.10; 0.10)$, $Q_0 = (50; 25)$, $N_0 = (52; 25)$. The simulated pollution processes are depicted in the upper ($Q_{1,t}$) and middle-part ($Q_{2,t}$).	70
4.4	\bar{S}_t permit price evolution letting vary the drift and the volatility terms for company one, respectively upper and middle picture, and both the initial permits endowments, lower picture. The common starting pollution parameters are $\mu = [-0.15; 0.001]$, $\sigma = [0.10; 0.10]$, $Q_0 = [50; 25]$, $N_0 = [52; 25]$	80

List of Tables

3.1	Likelihood-based goodness-of-fit for the SO ₂ return series. The best values for each criteria are marked in boldface.	36
3.2	Likelihood-based goodness-of-fit for the CO ₂ return series. The best values for each criteria are marked in boldface.	42
4.1	European Call and Put option prices according to two different option pricing models. The table reports the results for six possible general situations where we allow the pollution drift and volatility terms to vary. Initial pollution levels, $\mathbf{Q}_0 = (50; 25)$, and initial permits endowments, $\mathbf{N}_0 = (55; 25)$, are fixed. The risk-free rate is $r = 0.03$, maturity is $T = 1$; and the penalty $P = 40$. The simulated pollution processes n are 500 for each company and $A = 100$	78

Chapter 1

Summary

"You are using profit motive to achieve a public good, and this is just brilliant. . ."

-Henry Derwent, head of IETA

"Until people consciously realise the situation that the world is in and change their own patterns of behaviour, you can't change anything. . . One of the reasons carbon trading is so acceptable to the powers-that-be is that it doesn't substantially impact on existing operations."

-Toby Carroll, research fellow, Lee Kuan Yew School of Public Policy

"The point of the market is to find the most efficient way to reduce emissions. . ."

-Bill Hare, Greenpeace

Global warming is on the top of many agendas nowadays. In the last decade several organizations, policy regulators and scholars have been studying and discussing possible strategies to tackle the problem. Most of the proposed solutions include increasing energy

efficiency and conservation, examining the potential for capture, sequestration and storage of carbon, expanding the production of renewable energy, and even reviving nuclear energy production. But most of the recently proposed strategies include a form of cap-and-trade system for carbon dioxide emissions based on marketable permits. In such a system, the regulator allocates a number of emission permits to the responsible installations, and at pre-settled compliance dates, each source must have enough permits to cover all its recorded emissions or be the subject of significant penalties. The rationale for such a system is that the exchange of permits between firms through trading will minimize the overall social costs since companies that can easily reduce emissions will do so, and those for which it is harder will buy permits. Today several countries already implemented cap-and-trade schemes in their efforts to curb greenhouse gases (GHG) and markets for emission permits do exist in the U.S. and in Europe. In Chapter 2 we will survey the environmental economics principles (and assumptions) under which marketable permits are a cost-effective choice to reduce pollution.

One of the aims of emissions trading is the introduction and internalisation of the value of emitting CO₂ in the economy. For the parties participating in the European scheme, optimizing every-day operations while taking into account the value of emission permits seems obvious today. However, knowledge of the statistical distribution of the prices of marketable permits, and their forecastability, were not sufficiently studied by the literature. Such notions are crucial in constructing, among other things, purchasing and risk management strategies. Chapter 3 analyzes the two main markets for permits designed for air-pollution control in Europe and in the U.S. and investigates a model for dealing with the unique stylized facts of this type of data.

The first 4 years of the European cap-and-trade system for carbon dioxide emissions have proven cap-and-trade can work. We address those readers interested in the economic impacts and social results of the EU ETS to the numerous papers cited in Chapter 2. The current thesis concentrates more on the development of a valid dynamic price model. The transaction of permits -market liquidity- has been increasing from 300,000 emission permits per day in January 2005 to 4.9 million permits per day in January 2007. As such, the market does much more than simply transfer permits from companies with a surplus in permits to companies with a deficit. This implies that emission permits are not only considered a compliance tool, but are also transacted as financial contracts, putting them on the same level as other commonly traded securities. Such trading interest is important in establishing the market value of the CO₂ permits and is modeled in Chapter 4. The goal of this Chapter is to develop a mathematical model for the short-run equilibrium price of the emission permit in the presence of partial information, and to analyze the possible economic price equilibria when different initial permits endowments or companies characterizations are introduced.

Today derivative contracts play an important role in the markets for emission permits. By allowing market participants to reduce exposure to price risk, permit buyers and sellers can better plan their businesses. The availability of these markets can eventually provide the means to allow greater risk to be spanned, thus facilitating growth and efficiency in each of the associated industries. In the last part of Chapter 4, we briefly introduce a European-style (financial and real) option pricing model comparison. The benchmark model is that of Black, Merton and Scholes (1973) where one assumes the permit price evolves according to a geometric Brownian motion. The second option approach relies on the equilibrium price dynamics adjusted for the risk, i.e. under the risk neutral probability measure.

The thesis contribution is threefold. First, the econometric analysis provides companies with a comprehensive approach for analyzing the statistical distribution of the prices of emission permits and constructing effective risk management strategies for those portfolios which include emission permits. Second, the equilibrium model is designed to empower policy makers into designing emissions markets capable to meet emissions target while at the same time, accounting for the presence of partial (asymmetric) information and the existence of strategic interactions in the trades of permits. Finally, the equilibrium price dynamic under the risk-neutral measure can be easily employed for pricing CO₂-based (financial and real) contracts such as option derivatives or project investments (for example, those projects developed under the Clean Development Mechanism).

I would mention that some of the technical issues raised by this thesis involve new mathematical models in environmental economics and very general competitive equilibrium problems. They all lead to new mathematical challenges of optimization techniques including conditional Monte Carlo simulations. Finally, we hope the importance of the applications and the relevance of the tools of stochastic analysis and stochastic numerics highlight the role that mathematics has to play in major policy making decisions.

Chapter 2

Environmental Economics

“Private markets are perfectly efficient only if there are no public goods, no externalities, no monopoly buyers or sellers, no increasing returns to scale, no information problems, no transaction costs, no taxes, no common property and no other distortions between the cost paid by buyers and the benefits received by the sellers.”

-Fullerton and Stavins (1998)

Virtually every aspect of economic activity results in greenhouse gas emissions (GHG), so when the environmental revolution arrived in the late 1960s, the (environmental) economists were ready and waiting. Economists had what they saw as a coherent and compelling view of the nature of pollution with a straightforward set of policy implications. The problem of externalities¹ and the associated market failure had long been a part of the micro-economic theory. Economists saw pollution as the consequence of an absence of prices for

¹Externalities refers to situations when the effect of production or consumption of goods and services imposes costs or benefits on others which are not reflected in the prices charged for the goods and services being provided.

certain scarce environmental resources such as clean air and water, and they prescribed the introduction of surrogate prices in the form of unit taxes or “effluent fees” to provide the needed signal to economize on the use of these resources. This explains why the source of the basic economic principles of environmental policy is to be found in the theory of externalities. The literature on this subject is enormous; it encompasses hundreds of books and papers. An attempt to provide a comprehensive and detailed description of the literature on externalities theory reaches beyond the scope of this survey. Instead, we shall attempt in this paper to sketch the central results from this literature, with an emphasis on their implication for the quantitative analysis of the price dynamics of the marketable permits. And so, we begin in section 2.1 with a review of the theory of environmental regulation in which we explore theoretical results regarding the choice among the key policy instruments for the control of externalities: taxes, subsidies and marketable permits. Section 2.2 overviews those factors that affect the effectiveness of marketable permits addressed by several authors in the last decade. Section 2.3 concludes analyzing the recent attempts to develop a valid dynamic price model for emission permits.

2.1 Environmental Economics: A Primer

In a competitive market equilibrium, firms with free access to environmental resources will continue to engage in polluting activities until the marginal return of their production is zero. We thus obtain the familiar result that because of their disregard for the external costs that they impose on others, polluting agents will engage in socially excessive levels of polluting activities. The policy implication of this result is then clear in economics. Polluting agents need to be confronted with a price equal to the marginal external cost of their polluting activities to induce them to internalize at the margin the full social costs of their pursuits. Such a price incentive can take the form of the familiar “Pigouvian tax”, after

[65]. This is a levy on the polluting agent equal to the marginal social damage. However, the Pigouvian solution to the problem of externalities has been the subject of repeated attack along Coasian lines. [21] has elaborated on the externalities question by emphasizing that the root of the problem is that of undefined property rights. The author claims that if the ownership rights to clean air, for instance, were clearly defined and enforced, then self-interested parties would use legal and market mechanisms to bring about a socially acceptable level of externalities. However, such a theory holds in absence of transaction costs and strategic behavior (more on this in the next section).

Beside a tax on polluting activities, two alternative policy instruments have received extensive attention in the literature: subsidies and marketable permits. Early on it was recognized that a subsidy per unit of emissions reduction could establish the same incentive for the abatement activity as a tax of the same magnitude per unit of pollution emitted. Soon it became apparent that there are important asymmetries between these two policy instruments -see [47] for a comprehensive analysis. In particular, they have quite different implications for the profitability of production in a polluting industry: subsidies increase profits, while taxes decrease them. The policy instruments thus have quite different implications for the long-run, entry-exit decisions of firms. The subsidy approach will shift the industry supply curve to the right and result in a larger number of firms and higher industry output, while the Pigouvian tax will shift the supply curve to the left with a consequent contraction in the size of the industry -see [5]. The basic point is that there is a further condition, an entry-exit condition, that long-run equilibrium must satisfy for an efficient outcome. To obtain the correct number of firms in the long run, it is essential that firms pay not only the cost of the marginal damages of their emissions, but also the total cost arising from their emissions. Only if firms bear the total cost of their emissions will the prospective profitability of the enterprise reflect the true social net benefit of entry and exit into the industry.

The second policy instrument is a market-based instrument: marketable permits. Suggested applications for the use of market approach abound in the economics literature, especially in the fields of air and water pollution -see [5]. In a world of perfect knowledge, marketable emission permits are, in principle, a fully equivalent alternative to unit taxes. With a system of marketable permits in place for air-pollution control for instance, instead of setting the proper Pigouvian tax and obtaining the efficient quantity of pollution as a result, the environmental authority could issue (emission) permits equal in the aggregate to the efficient quantity of pollution and allow firms to bid for the permits. One can show that the market-clearing price of the (emission) permits will produce an outcome that satisfies the first-order conditions both for efficiency in pollution abatement activities in the short-run and for entry-exit decisions in the long run. The regulator can, in short, set either “price” (tax) or “quantity” (emission cap) and achieve the desired result. This symmetry between the price and quantity approaches is, however, critically dependent upon the assumption of perfect knowledge. In a setting of imperfect information concerning the marginal benefit and cost functions, the outcomes under the two approaches can differ in important ways.

In a seminal paper, [79] investigated this asymmetry between price and quantity instruments and produced a theorem with extremely important policy implications. The theorem establishes the conditions under which the expected welfare gain under a unit tax exceeds, is equal to, or falls short of that under a system of marketable permits (quotas). In short, the theorem states that in the presence of uncertainty concerning the costs of pollution control, the preferred policy instrument depends on the relative steepness of the marginal benefit and cost curves. The intuition of Weitzman is straightforward. Consider, for example, the case where the marginal benefits curve is quite steep but marginal control costs are fairly constant over the relevant range. This could reflect some kind of environmental threshold

effect where, if pollutant concentrations rise only slightly over some range, dire environmental consequences follow. In such a setting, it is clearly important that the environmental authority have a close control over the quantity of emissions. If, instead, a price instrument were employed and the authority were to underestimate the true costs of pollution control, emissions might exceed the critical range with a resulting environmental disaster. In such a case, the Weitzman theorem tells us, quite sensibly, that the regulator should choose the quantity instrument (because the marginal benefits curve has a great absolute slope than the marginal cost curve). Suppose, next, that it is the marginal abatement cost curve that is steep and that the marginal benefits from pollution control are relatively constant over time. The danger here is that because of imperfect information, the regulatory agency might, for example, select an overly stringent standard, thereby imposing large, excessive costs on polluters and society. Under these circumstances, the expected welfare gain is larger under the price instrument. Polluters will not get stuck with inordinately high control costs, since they always have the option of paying the unit tax on emissions rather than reducing their pollution further. The Weitzman theorem thus suggests the conditions under which each of these two policy instruments is to be preferred to the other in the presence of abatement cost uncertainty. Not surprisingly, an even better expected outcome can be obtained by using price and quantity instruments in tandem, see [67].

After two decades, [74] showed that also benefit uncertainty matters. In particular, the instrument-neutrality long identified with equal absolute valued slopes of marginal benefits and marginal costs likewise disappears when there exists a significant correlation between them. Quite remarkably, Stavins' results suggest that the conventional identification - under Weitzman policy instrument recommendations - of a price (tax) instrument can in fact be reversed, to favor instead a quantity (marketable) measure. On the other hand, the results also suggest that it is less likely that the conventional identification of a quantity measure as being more efficient to a price measure will be reversed.

In sum, there is a basic sense in which systems of taxes and marketable emission permits are equivalent: the environmental authority can, in principle, set a price (i.e. a tax) and then adjust it until emissions are reduced sufficiently to achieve the prescribed environmental standard. Alternatively, it can issue the requisite number of permits directly and allow the bidding of polluters to determine the market-clearing price. However, the basic equivalence obscures some crucial differences between the two approaches in a policy setting; they are by no means equivalent policy instruments from the perspective of a regulatory agency. A major advantage of the marketable permit approach is that it gives the environmental authority direct control over the quantity of emissions. Under the tax approach, the regulator must set a tax, and if, for example, the tax turns out to be too low, the pollution will exceed permissible levels. In sum, the regulatory agency might have to enact periodic (and unpopular) increases in taxes. In contrast, a system of marketable permits automatically accommodates itself to growth and inflation. Since there can be no change in the aggregate quantity of emissions without some explicit action on the part of the agency, increased demand will simply translate itself into a higher market-clearing price for permits with no effects on levels of pollution. On the other side, polluters are likely to prefer the permit approach because it can involve lower levels of compliance costs.

As a result, marketable permits have often been identified as *second-best*² approaches to policy design. When we cannot assume the existence of a perfectly competitive equilibrium, market based measures can be seen as effective regulatory instruments for the achievement of predetermined environmental standards, see [4] and next section for further discussions. The chief appeal of economic incentives as the regulatory device for achieving environmental standards is the large potential cost-savings that they promise. There

²The Theory of the Second Best concerns what happens when one or more optimality conditions are not satisfied in an economic model.

is now an extensive body of empirical studies that estimate the cost of achieving standards for environmental quality under existing command-and-control regulatory programs. These are typically programs under which the environmental authority prescribes the treatment procedures that are to be adopted by each source. The studies compare costs under command-and-control programs with those under a more cost effective system of economic incentives. The results have been quite striking: they indicate that control costs under existing programs have often been several times the least-cost levels -see [77] for a survey on cost studies. The source of these large cost savings is the capacity of economic instruments to take advantage of the large differentials abatement costs across polluters. In addition, the information problems confronting regulators under the more traditional command-and-control approaches are enormous and they lead regulators to make only very rough and crude distinctions among sources. In a setting of perfect information, such a problem would, of course, disappear. But in the real world of imperfect information, economic instruments have the important advantage to economize on the need for the environmental agency to acquire information on the abatement costs of individual sources. This is just another example of the more general principles concerning the capacity of markets to deal efficiently with information problems. The estimated cost savings in the studies mentioned above result from a more cost effective allocation of abatement efforts within the context of existing control technologies. From a more dynamic perspective, economic incentives promise additional gains in terms of encouraging the development of more effective and less costly abatement techniques. As reported by [46] and more recently by [10], a market for permits provides a greater incentive to R&D efforts in control technology than will a regulation that specifies some given level of pollution.³

³A U.S. Congressional Budget Office study in 2006 concludes that a strategy combining both research and development to cut GHG emissions and emission permits -in particular carbon dioxide permits- would prove more effective and better balanced economically, than one relying simply on new technologies.

Largely for the reasons mentioned above, policy makers in the U.S. have found marketable permits preferable to taxes as a mechanism for providing economic incentives for pollution control. The major program of this genre is the EPA's Emission Trading Program for sulfur dioxide (SO_2). This market has been created under the 1990 Amendments to the Clean Air Act. It was designed to address the acid rain problem by cutting back annual sulfur emissions by 10 million tons. This market permitted affected power plants to meet their emissions reduction quotas by whatever means they will, including the purchase of extra emissions reductions from other sources. The scheme significantly increased the flexibility with which sources can meet their pollution limitations, and this has been important for it has allowed substantial cost savings -see [32] and [31] for detailed descriptions. In early stage the great majority of the trades have been internal ones.⁴ A real and active market in emissions permits involving different firms has developed under the program only quite recently. For an in deep econometric analysis of the SO_2 market see Chapter 3.

Conversely, in the past, the use of taxes was more prevalent in Europe where they have been extensively employed in systems of water quality management -see [78] for a comprehensive summary. However, taxes have typically been low and have tended to apply to average or expected pollution rather than to provide a clear cost signal at the margin. Moreover, the charges were overlaid on an extensive command-and-control system of regulations that mute somewhat further their effects as economic incentives. Recently, following the example of the Kyoto protocol, European policy makers implemented the largest and most important market for emission permits: the European Emission Trading Scheme (EU ETS). At its launch in 2005, the scheme covered the carbon dioxide (CO_2) emissions from energy-intensive industry sectors in the then 25 member states, responsible for nearly half of the EU's CO_2 emissions. Today, the scheme includes 27 countries and

⁴This fact explains the large presence of zeros in the return time series of SO_2 - see the discussion in Chapter 3.

claims 80 per cent of the value of the world's markets for marketable permits. The scheme has so far worked as it was envisioned: a European-wide price on emission of CO₂ was established, businesses began incorporating this price into their decision-making, and the market infrastructure for a multi-national trading program is now in place.

Since market-based instruments are extensively being used as a tool for pollution control at a regional and international scale, there is an increasing need to develop effective dynamic models for the price of marketable permits. In fact, a valid price model is an essential component for any decision-making process, and for constructing optimal hedging and purchasing strategies in a (carbon) constrained market. Furthermore, firms trade permits not exclusively for compliance purposes but some take also speculative positions, as reported by several analysts in early 2008. The last section surveys the few model attempts which exist in literature.

2.2 Marketable Permits Systems and Effectiveness

One of the first references to marketable permits can be found in the seminal works of [25] and [27]. In these papers the pollution abatement problem is viewed within an economic, cost-benefit framework in conjunction with the concept of property rights introduced in the previous section. Based on such an idea, [60] provides a rigorous theoretical justification of how marketable emission permits leads to the efficient allocation of abatement costs across various "sources of pollution". Necessary and sufficient conditions for market equilibrium and efficiency are derived given the setting of multiple profit-maximizing firms who attempt to minimize total compliance costs. Literature discussing those factors which can adversely affect the performance of marketable permits systems and not addressed directly in Montgomery (1972) has followed. Among the most important, we recall concentration

in the permit market [38] and [55], concentration in the output market [52], the preexisting regulatory environment [8], the degree of monitoring and enforcement [48], and the presence of non-negligible transaction costs in the market [73]. We briefly overview these papers below.

The appeal of using marketable permits as a means of allocating scarce resources stems in large part from the assumption that a market for permits will approximate the competitive ideal. When the competition is not a foregone conclusion, the question arises as to how a firm might manipulate the market to its own advantage. [38] has discussed such issues as market manipulation developing a one-period model where one firm has market power and all transactions of emission permits take place at a single price. The author's principal finding is that the degree of inefficiency observed in the market is systematically related to the distribution of permits. In other words, in the presence of market power, the initial distribution of permits matters, with regard not only to equity considerations but also to cost.⁵ This is to say, it is the demand of the firm with market power which determines the equilibrium price of the emission permits. Building on the theory of cost-minimizing manipulation and the literature on raising rivals' costs, [55] have discussed a different form of market manipulation: exclusionary manipulation. Because permit prices are sensitive to the purchases (or sales) of the firm with market power, exclusionary manipulation can aggravate the inefficiencies which occur in both the market for permits and the product market.⁶

Similarly, the efficiency of marketable permits system depends on the competitiveness

⁵Traditional models of marketable permits view problems of initial permit distribution as being strictly an equity issue. The analysis of [60] is one such example where firms are assumed to be all price takers. For the case of one pollutant and one market, the author finds that the distribution of permits will have no effect on achieving the target in a cost-effective manner.

⁶Anticipating discussions in the next section, in the dynamic model presented in [19] permit prices are sensitive to the strategic trading interactions of all buyers and sellers in the market.

of the output markets in which polluting firms compete. [52] has shown that the introduction of marketable permits increases aggregate “welfare” if the output markets are competitive. In contrast, in the presence of non-competitive output markets, a system of emission permits may reduce social “welfare” even if the market for the emission permits is perfectly competitive.

The strength and effectiveness of the incentives created by a cap-and-trade scheme will depend in large part on the rules that regulators apply to permit transactions. These rules will determine how affected firms will be compensated for investments in emission permits and whether ratepayers and shareholders will share in the benefits of trading emission permits. The influence of uncertainty regarding the regulation policy of public utility commissions in the U.S. market for SO₂ have been discussed by [8]. The authors develop and analyze a model of individual utility decisions that focuses on the choice between purchasing permits or investing in SO₂ abatement measures to comply with the law. The key finding is that policy rules influence the relative cost of investments in emission permits versus switching to low sulfur fuels (a medium-term abatement measure) or investments in emission control equipment (typically long-term abatement measures). Furthermore, such rules may distort the incentives of utilities to adopt the least cost combination of emission permits and other compliance strategies required to satisfy the U.S. regulation.

The degree of monitoring and enforcement has also been subject of several studies. In [48], the author extends the previous research on marketable permits with noncompliant firms. Keeler makes a specific comparison between command-and-control and marketable permits systems when regulatory authorities are unable to achieve full compliance. In particular, the author studies the sensitivity of the shape of the penalty function faced by non-compliant firms. His analysis indicates that under plausible penalty functions marketable permits may allow more pollution or higher fraction of regulated firms out of compliance. These results highlight the importance of implementation in the success of pollution control

strategies relying on marketable permits.

After [2], the presence of transaction costs in the markets for permits was a fact. Though already [39] and [5], among several other authors, have commented on the potential importance of transaction costs in the markets for emission permits, [73] has been the first to include transaction costs into a formal model. Another source of indirect evidence of the prevalence of transaction costs in the U.S. market for SO₂ permits comes from the well documented “internal trading” within firms, as opposed to “external trading” among firms. It has been hypothesized that the crucial difference favoring the internal trades and discouraging the external trades is the existence of significant transaction costs that arise once trades are between one firm and another [39].⁷ Stavins claims that transaction costs reduce the volume of permits trading, regardless of the specific forms that the marginal control cost functions and transaction cost functions take, as long as the marginal control cost functions are nondecreasing.⁸ Not surprisingly, equilibrium permit allocations are sensitive to initial distributions of permits. This result is fully consistent with the Coase Theorem, which states that in the presence of transaction costs, the anticipated outcome from a process of bilateral negotiation is variant with respect to the initial assignment of property rights [21]. In sum, the presence of transaction costs reduce trading levels and the discretion of each environmental agency which, as opposed to [60], “can[not] distribute licenses as it pleases.”

The attempt to find a specific initial permit allocation to overcome some of the problems described so far, can actually result in a scheme that will be far more costly than planned. This may argue for the economist’s favorite permit-allocation mechanism: Auctions. This approach, for instance, becomes even more attractive in the presence of transaction costs. [24] have analyzed the distributional implications of allocating CO₂ emission

⁷This fact explains the large presence of zeros in the return time series of the SO₂ as discussed in [64].

⁸A proof of this statement is found in [72]

permits through auctions rather than through some form of *grandfathering*.⁹ The authors argue that auctioning is superior because it increases efficiency by reducing existing tax distortions; it also offers greater incentives for innovation and gives more flexibility in the distribution of costs; finally, it reduces the need for politically contentious arguments over the allocation of rents. On this latter issue, [75] considers a construction of an allocation scheme in the presence of market power. These authors also point out that auctions may not be chosen due to vested interests bringing on a powerful voice in favor of grandfathering. Finally, [16, 17] has conducted an experimental analysis on auction and rules design.

2.3 Modeling Marketable Permits

As obvious from previous section, literature focusing on the economic and policy aspects of marketable permits is extensive. However, an explicit and formal study of the dynamic price of the emission permits is an almost unexplored area. Most of the present research relies on the key result that, in a competitive market with perfect information, the equilibrium price of the emission permits is equal to the marginal costs of the cheapest pollution abatement solution. This statement underpins the belief that a high price level for emission permits brings about the relevant companies with lower marginal abatement costs in order to exploit consequent price differences. Such companies make profits by lowering the level of pollution more than is necessary to comply with regulations and subsequently sell their unused permits relying on banking opportunity.¹⁰ Instead of limiting intertemporal trading to banking, [69] allows both borrowing and banking and extends the work of [77] and [26]

⁹For a detailed discussion of initial allocation criteria see [3] and references therein. For a comprehensive analysis of the social and economic impact of allocation criteria see [9] and [15].

¹⁰To generate permits, a firm may choose to pollute less than the current standard and sell the “unused” permits to a different firm or deposit them in an “emission bank account” to be used later by the firm or sold at a later time to another firm. The borrowing of permits occurs when a firm pollutes more than its current standard, but the cumulative deficit must be repaid by the end of the planning scheme.

providing a more general treatment of permit trading in continuous time through the use of optimal-control theory. In particular, the author explores the problem of minimizing the cost of intertemporal emission control by N heterogeneous firms in the presence of emission permits that are tradable across firms and through time. In such a setting, firms may directly abate emissions, and they may purchase, sell, bank and borrow emission permits in order to meet applicable standards or to take advantage of any speculative opportunities that may arise. The equilibrium permit price is shown to be constant in time and equal to the marginal cost of pollution abatement when each firm can bank and borrow permits. Conversely, if the firm desires to borrow but this is not permitted, the equilibrium permit price is decreasing. A special case arises when the discount rate is nil. In this situation firms have no incentive to undertake abatement measures until the future. If pollution emission rates becomes more strict through time, firms tend to save more or buy more permits in the beginning time periods for later use. Higher discount rates lower the value of future cost savings and decrease the incentive for firms to bank permits. Perhaps one of the most important findings of Rubin is the ability of firms to shift their emission stream through time as a consequence of banking and borrowing. In particular, when social damages are an increasing function of the level of pollution emitted at any one time and pollution standards are becoming stricter through time, banking is good. It provides cost saving to firms by allowing them to adjust their own internal rates of emission reduction to an externally set regulations. However, when regulations are constant or easing, then allowing firms to borrow will raise social damages while lowering firms' costs.

Though [69] provides a comprehensive treatment of intertemporal emission trading, its analysis has been framed in a world of certainty where strategic interaction was not taken into consideration. [70] has introduced uncertainty in Rubin's model (1996): this reduces the expected permit price growth rate. This paper is one of the first that implicitly analyzes the permit price in a stochastic, continuous-time and infinite-time horizon model. In line

with previous research, in the model of Schennach a level of pollution abatement is chosen such that the current marginal cost of abating equals the current emission permit price. However, this is not true in practice as observed permit market prices are typically far away from their expected theoretical levels.

Recently, in an effort to bridge the gap between theory and observed market-price behavior, an increasing number of empirical studies has been investigating the historical time series of the permit price. In [28] several different diffusion and jump–diffusion processes were fitted to the European CO₂ futures time series. [6] analyze the short-term spot price behavior of CO₂ permits employing a Markov–switching model to capture the heteroskedastic behavior of the return time series. In contrast, in Chapter 3 we advocate the use of a new GARCH-type structure for the analysis of inherent heteroskedastic dynamics in the returns of SO₂ in the U.S. and of CO₂ emission permits in the EU ETS.

With a precise focus on the European emission market and in an attempt to develop a valid dynamic price model, [71] and [34] elaborate a quantitative analysis of the CO₂ permits price founded on the pivotal results from environmental economics literature. In particular, [71] consider one representative agent who decides whether or not to spend money on lowering emission levels. The model is based on the optimal abatement decision of an affected company, therefore it very much depends on its total expected emissions. With a distinction between long-term and short-term abatement measures, [34] concentrate on the energy sector considering n affected utilities which decide their abatement levels by relying on the cheapest possible abatement option in the short-term, i.e. so-called fuel-switching.¹¹ Chapter 4 contributes to this increasing body of literature on quantitative

¹¹It involves the replacement of high–carbon (sulfur) fuels with low–carbon (sulfur) alternatives. The most common form of fuel switching in the U.S. is the replacement of high–sulfur coal with a low–sulfur coal. In Europe, coal is typically replaced by natural gas.

environmental economics. This is a new strand of research which focuses on financial and quantitative issues originating from solutions proposed by environmental economists.

In common with Fehr and Hinz, we differentiate short-term and long-term abatement measures and by means of the dynamic optimization we develop an endogenous model for the emission permit price dynamics. In particular, we assume that companies are characterized by exogenous pollution dynamics and they optimize their cost functions by continuously adjusting their permit portfolio allocations and by choosing the optimal permit amount to purchase (in the shortage permit case) or to sell (in the excess permit case). The result is an equilibrium price for emission permits where the price is sensitive to the trading interaction of all buyers and sellers which found their strategies on their own (different) information sets.

Chapter 3

Econometric Analysis of the Marketable Permits

"Lack of clarity. . . post-2012 is countering growth of markets such as the EU ETS. . . The market is truly at a crossroads as participants appreciate the complexity and risks of carbon trading."

-Andrew Ertel, chief executive of Evolution Markets (May - 2008)

Title IV of the Clean Air Act Amendment (CAAA) in the U.S., and the Emission Trading Scheme in Europe (EU ETS), created *de facto* property rights for pollution, referred to as *emission permits* or *allowances* in the programs, that can be freely traded.¹ The right gives relevant subjects² complete flexibility in determining how they will comply with their obligations under the programs, see Chapter 2 for a comprehensive introduction to the basic principles of environmental economics. Permits can be traded nationally in the case

¹The original version of this chapter was co-written with Marc Paoletta and is to appear in *Journal of Banking & Finance*, under the title "An Econometric Analysis of the Emission Allowance Prices" .

²Title IV in the U.S. affects mainly electric utilities and EU ETS affects different sectors like iron, steel, cement, glass and ceramics, pulp and paper producers and the energy sector as well.

of the U.S., and internationally under the EU ETS, with no necessity of prior approval. The purchase and holding of permits is not restricted to the companies affected by the programs—which means that all sources, as well as third parties such as brokers, are free to buy and sell permits with any other party.³

Knowledge of the statistical distribution of the prices of marketable permits, and their forecastability, are crucial in constructing, among other things, purchasing and risk management strategies in those markets recently affected by tighter environmental regulations. Therefore, this chapter analyzes the two main markets for permits designed for air-pollution control and described in the introduction of Chapter 2: the CO₂ market in Europe, and the SO₂ market in the U.S. This chapter investigates a model for dealing with the unique stylized facts of this type of data. Its effectiveness in terms of model fit and out-of-sample value-at-risk forecasting, as compared to models commonly used in risk-forecasting contexts, is demonstrated.

Along with the working papers of [28] and [6], this paper provides one of the first econometric investigations of the behavior of the new emission permits. Our approach is completely different than that used in both the aforementioned papers. [28] use a jump-diffusion model to approximate the random behavior of the CO₂ emission spot price, while [6] analyze the short-term spot price behavior of CO₂ emission permits and employ a Markov-switching model to capture the heteroskedastic behavior of the return series. In contrast, we build upon a recently developed GARCH-type structure particularly suited to the stylized facts of the data. Our approach differs also from those alternative pricing models driven by economic considerations. According to this strand of literature, energy

³As mentioned in a footnote in Chapter 2, the option of full transferability of emission permits leads to the relevant question of which type of *agent* act in the emission market with respect to the different priorities and aims. We leave this question for future research.

prices and climatic conditions are the most important drivers of the emission permits (see for instance [20], [13], and [23] among others). Though the empirical findings of these models differ notably due to diverging input variables and different modeling approaches, the common result is a loose identification of what are the factors and with which intensity they affect the price of the emission permits.

The remainder of the chapter proceeds as follows: Sections 3.1 and 3.2 provide the empirical analysis of the SO₂ and CO₂ returns, respectively. Section 3.3 provides concluding remarks and ideas for future research.

3.1 An Econometric Approach for Analyzing the SO₂ Data

We consider a time-series model applied to the return series, $r_t = 100(\ln p_t - \ln p_{t-1})$, generated by the price sequence p_t . This section discusses in detail the analysis on the SO₂ spot price data set, using the 1,780 returns from January 4, 1999 to May 16, 2006. The spot closing prices SO₂ have been collected by the Chicago Climate Exchange on the OTC market.

3.1a Illiquidity

The emission permits market is nonstandard due to the fact that the traded asset is itself a so-called *nonstandard commodity*. Electric power companies, for example, do not physically need the emission-right to produce and, therefore, to pollute; in most cases it is feasible to delay action and wait for new information before purchasing permits. This also applies the other way around: A firm that holds more permits than it expects to need may still hold onto the surplus because they have some option value, given that it may be costly to get them back once they are sold. So, *illiquidity* arises endogenously from the fact that firms can

emit without having permits and thus fear that they may face a market squeeze at the end of the year. The relevance of the (incomplete) information-flow in the market for permits and the impact of the strategic interaction of relevant companies is subject of research in Chapter 4 for the specific case of the CO₂ market.

Historically, markets for permits have never been completely liquid. However, volume and liquidity on the emission permit markets have increased over time, particularly from the end of 2005. Nevertheless, the data set exhibits a relatively large (29%) number of zero returns, due in part because of the relatively small number of agents interacting on the market. This is a common finding of exchange-traded assets which, on a daily scale, possesses a low floating capital and an even lower traded volume. For example, SO₂ is a regional problem in the U.S., where covered utilities include only a few hundred large energy producers (a few thousand facilities). As anticipated in Chapter 2, another plausible explanation is that within-firm trading at the same price could be taking place. For example, AES, the largest electricity producer in the U.S., has numerous facilities covered by the CAAA Title IV, and it is not unreasonable to assume that emission permits are financially transferred from one balance sheet to another, at market price.⁴

3.1b Basic Analysis

Below we will present a statistical analysis of the returns, emphasizing the interplay between the standard features of the data (fat tails and volatility clustering) and the less-standard fact that the data exhibit a much greater percentage of zero-returns than the more commonly analyzed financial markets.

⁴At the time of writing, data to confirm this are not available, although from the USEPA website, the intra-company transaction volume is available, thus confirming at least that within-firm trading does take place on a large scale. A similar argument is described in [29]. Furthermore, as discussed in Chapter 2, within-firm trading reduced transaction costs.

The top panel in Figure 3.1 plots the return data, from which we clearly see the presence of volatility clustering. The sample autocorrelation function (SACF) for the returns is typical in the sense that very little correlation structure is present in the data, and is not shown. Unsurprisingly, there is much stronger correlation involving the absolute returns. The bottom panel of Figure 3.1 shows the SACF of the absolute returns, but having first stripped the data set of the zeros: By removing the zeros, the revealed correlation structure is stronger, though the SACF with the zeros still in place is similar - graph available from the author. Thus, other than the larger-than-usual number of zeros, the returns exhibit the usual stylized facts of asset returns, including a very low predictive component for the mean, strong volatility clustering, and tails which are far fatter than the normal.

To emphasize the tail behavior, Figure 3.2 shows a kernel density estimate of the returns data, but having removed the zeros (explaining the small dip in the curve near zero), which does not affect the tail of the distribution, but would otherwise jeopardize the quality of a fitted distribution. The kernel density of the data with the zeros looks similar, but with a higher peak near zero. Fitting a normal or stable distribution⁵ to the returns, including the zeros, leads to a very misleading fit in the tails. A mixture of normals could be fit to such data, and this is done in the more general context of the conditionally heteroskedastic model below. We see that the nonzero returns are virtually symmetric, obviating the need for distributions which support asymmetry. The graph also shows an overlaid normal with matching mean and variance, and a location-scale symmetric stable distribution (fit via maximum likelihood). While the normal fit is disastrous, the stable distribution fits the data rather well, with estimated tail index (and estimated standard error) $\alpha = 1.453 (0.045)$, location term $\mu = 0.092 (0.047)$ and scale term $c = 0.980 (0.037)$.

In light of the excess number of zeros, a conditional time series model for the returns

⁵See [66], Paoletta (2007, Ch. 8) and the references therein for the history, the theoretical and practical value, and computational issues of using the stable Paretian distribution in financial modelling.

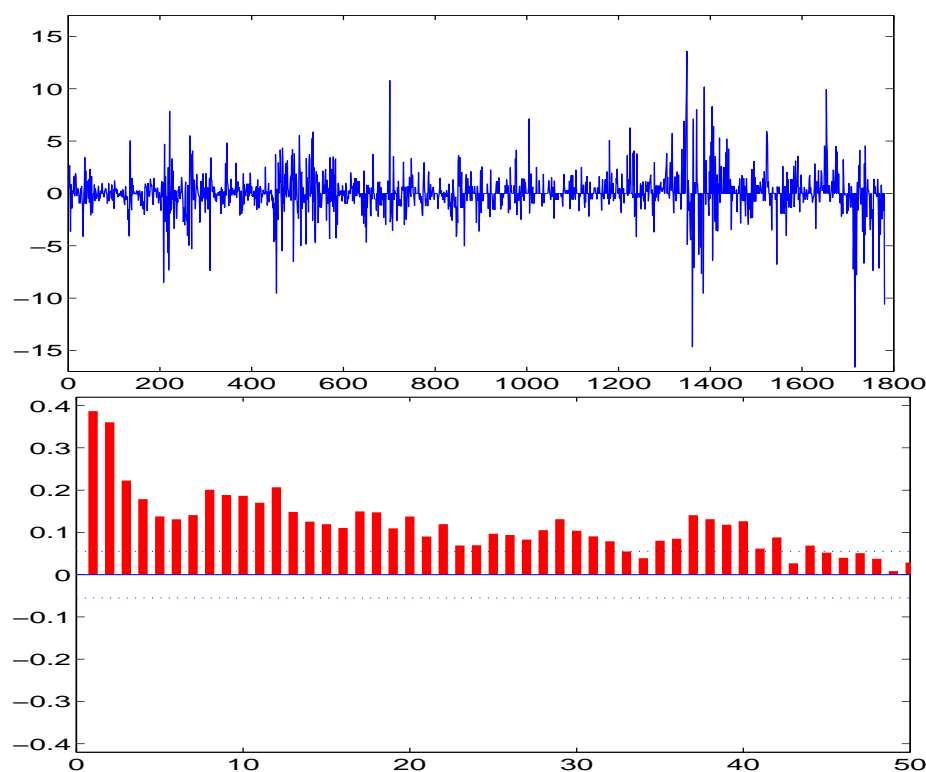


Figure 3.1: Daily SO₂ returns (top), and the SACF of the zeros-removed absolute returns (bottom).

would have to account for any dependency structure in the occurrence of zeros. To test this, we use the standard runs-test, reducing the data to a sequence of zeros (zero return) and ones (nonzero return), and using the asymptotic normality of the test statistic. Inspection of the return series shows that relatively more zeros occur in the first third of the data set (from January, 1999 up to March, 2002) than later, so that the runs-test applied to the whole series leads to blatant rejection of the null hypothesis. To account for this, we could test just, say, the last two thirds of the data, but more informative is to perform the test on the return series starting from the s th return to the end of the series, $s = 1, \dots, T - 50$, where $T = 1,780$ is the total number of returns. Figure 3.3 plots the resulting p -values, and

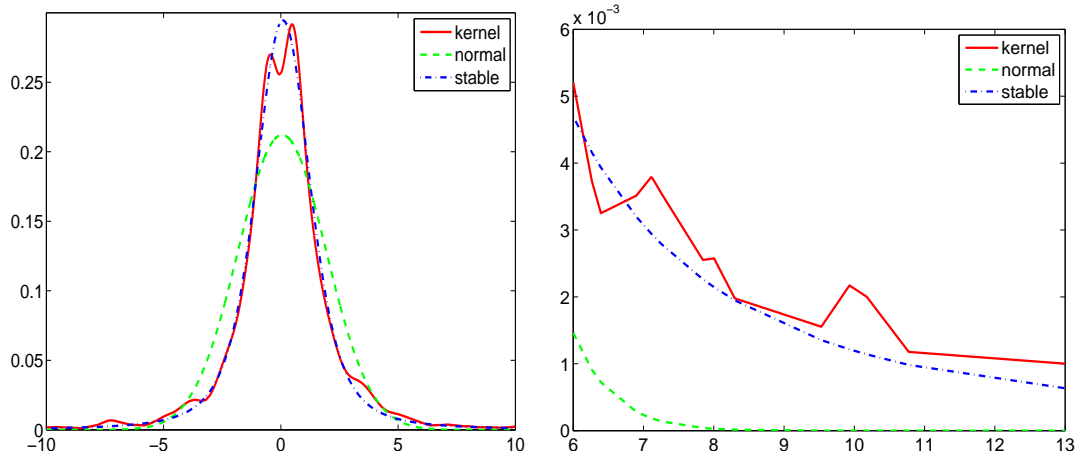


Figure 3.2: Kernel density (solid) of the SO₂ return series, with the best-fitting normal density (dashed) and best-fitting symmetric stable density (dash-dot). Right panel is just the magnified view of the right tail.

shows that, for the latter half of the data set, the null hypothesis of no correlation structure in the occurrence of zeros cannot be rejected. This fact is important for the conditional model employed below; if the assumption of randomness of the zeros were not tenable, then a more complicated model involving Markov-switching structures would have been necessary.

3.1c Stable-GARCH Model

Perhaps the most common effective conditional model used in both academic and financial institutions contexts for the analysis of asset returns data is a variant of the power-GARCH(r, s) model given by

$$r_t = \mu_t + \sigma_t z_t, \quad \sigma_t^d = \theta_0 + \sum_{i=1}^r \theta_i |r_{t-i} - \mu_{t-i}|^d + \sum_{j=1}^s \phi_j \sigma_{t-j}^d, \quad d > 0. \quad (3.1)$$

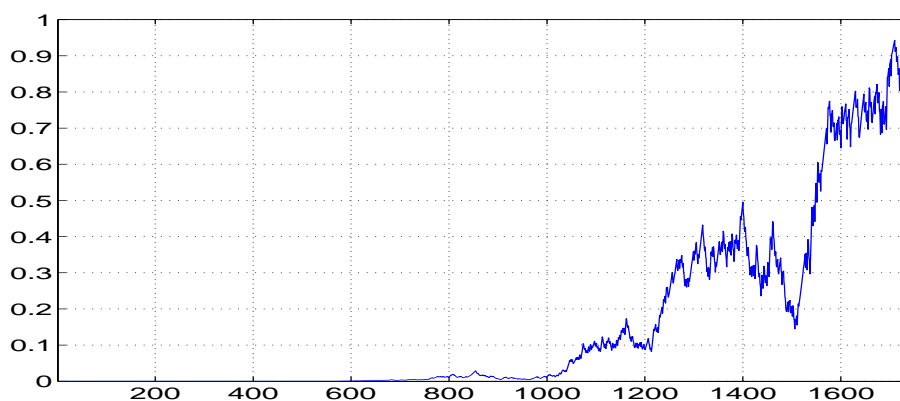


Figure 3.3: The p -values from the runs-test performed on segments of the SO_2 return series. The first segment is the returns in the whole series, the second is from the second return to the end, etc., up to the $(T - 50)$ th observation to the end.

In the ubiquitous $r = s = 1$ case, we require $\theta_0 > 0$, $\theta_1 \geq 0$, $\phi_1 \geq 0$. Here, $z_t \stackrel{\text{iid}}{\sim} f_Z(\cdot)$ with f_Z a zero-location, unit-scale continuous probability density function (pdf). In the standard GARCH model, f_Z is the Gaussian density and $d = 2$; the t -GARCH takes f_Z to be the Student's t pdf and $d = 2$; the stable-GARCH model, denoted $S_{\alpha,\beta}$ -GARCH, takes f_Z to be the $S_{\alpha,\beta}$ density and $d = 1$; see [58] for details.⁶

The stable-GARCH model possesses two important advantages over use of the t -GARCH. Firstly, the $S_{\alpha,\beta}$ -GARCH in-sample fit and out-of-sample Value-at-Risk (VaR) forecasting ability are generally superior to t -GARCH [57]. Secondly, the use of stable Paretian innovations is theoretically more appealing because of its relation to the generalized central limit theorem (GCLT) and which, via the stability property of the distribution, can be tested and confirmed to be applicable in some financial asset return series [61].

⁶In general, d can be estimated. For the stable-GARCH model, we require $0 < d < \alpha$. In practice, $\alpha > 1$ and, for numerous financial return series, the out-of-sample forecasting ability is barely affected by the choice of $d \in [1, \alpha)$.

For the SO₂ data, because of the aforementioned issue with a preponderance of zeros in the return series, the data generating process (DGP) is not consistent with any typical distributional assumption (such as Student's t or stable Paretian), in either the unconditional or conditional (GARCH) case (though a mixture model suggests itself, and is the method used below). Because fatter-than-normal tails of a unimodal distribution imply a more peaked center, the excess amount of zeros will have the effect of causing the tail index (the thickness of the tail) to be biased downwards (thicker), thus overestimating the risk of extreme tail events.

To illustrate, we estimate the $S_{\alpha,\beta}$ -GARCH model for the SO₂ return series. The estimated asymmetry term is $\hat{\beta} = -0.003$, which is practically and statistically insignificant. Note that the resulting estimate of the tail index α pertains to the stable Paretian innovations of the GARCH process describing the returns, i.e., the GARCH effects (which also give rise to the fat-tails of the returns) are taken into account, so that the resulting index will be greater (i.e., correspond to thinner tails) than the unconditional counterpart. The resulting estimate of the tail index (and approximate standard error) are $\hat{\alpha} = 1.0278$ (0.015), suggesting the plausibility of Cauchy ($\alpha = 1$) innovations, which do not even possess a finite mean and is therefore not a tenable assumption. If we numerically restrict $\hat{\alpha}$ to be above 1.3, we obtain $\hat{\alpha} = 1.498$ (0.018), showing that a local maximum of the likelihood does exist in a "plausible" region of the sample space. This occurs because the innovations in the conditional model are not stable Paretian (there are too many zeros, yielding the near-Cauchy fit), but the tails are thinner than Cauchy, which resulted in the trade-off value for $\hat{\alpha}$ of about 1.5.

If, for illustrative purposes, we strip all the zeros from the return series and then fit the $S_{\alpha,0}$ -GARCH model, we obtain $\hat{\alpha} = 1.640$ (0.022). This value is in agreement with the range of estimated tail indexes of numerous other financial data sets and is a far better reflection of the true thickness of the (conditional) tail. However, given the ad-hoc nature

of the removal of zeros, this is still an unsatisfying approach for building a realistic model of the return series. Similar results are obtained when using the t -GARCH model.

In fact, the abundance of zeros also precludes effective use of other GARCH-type models which otherwise tend to perform excellent in terms of VaR forecasting. In particular, one might think that the GARCH-EVT model, which focuses on tail estimation of the residuals of GARCH-filtered returns via methods of extreme value theory, would be particularly suited for VaR prediction of the SO₂ data, given its unique behavior in the center (but not the tails) of the distribution. The problem, however, is that the choice of innovations assumption used with the GARCH filter in the first step of the GARCH-EVT model is decisive for its forecasting performance, as detailed in [50]. Thus, the same problem arises as with the use of conventional fat-tailed-GARCH models. Similar findings apply to the use of the—otherwise highly effective—method of filtered historical simulation (FHS). More encouragingly, Kuester et al. (2005) show that the mixed-normal GARCH model (which is our proposed solution to the zeros problem) delivers highly competitive VaR forecasts on par with the quality of GARCH-EVT and FHS.

3.1d Mixed Normal GARCH: Model and Numerical Issues

As discussed above, the problem with GARCH formulation (3.1) for the data under study is the excess number of zero-returns, which precludes use of the usual array of distributions useful in this context, such as Student's t , skewed t extensions, hyperbolic, and $S_{\alpha,\beta}$, to name a few.

One candidate distribution which is perfectly suited for capturing this phenomenon is to use a mixture-model, taking f_Z to be a weighted sum of two or more pdfs. It might seem natural to have one component be degenerate at zero, and the other(s) continuous, but it suffices, and is operationally easier to implement, to choose all components of f_Z to

be continuous pdfs, in this case, each Gaussian, with the first one possessing a very small variance and a mean at zero (more on this below).

A further advantage of the mixture model is that it lends itself to economic interpretation. For example, [51] attribute volatility clustering and the emergence of fat-tailed returns mainly to agents' switching between fundamentalist and chartist strategies. A mixture of two or more normals could arise from different groups of actors, with one group acting, for example, more volatile than the others, or, possibly, processing market information differently. This idea is related to recent research with experimental data by [49], who show that heterogeneous fundamental information is a major source for the emergence of fat tails and volatility clustering.⁷ This could apply to the SO₂ market: With the approaching of the more stringent Phase II in Title IV and with the SO₂ emission level taking shape for the different utilities, companies obtained a better indication of their Phase I net positions and some appeared to refrain from speculation covering their short positions on a forward basis and saving the remaining permits.

A GARCH-type model with mixed normal innovations, denoted MixN-GARCH, has already been proposed and studied independently by [37] and [1]. The model was not designed with the zeros-problem in mind, but rather motivated by the aforementioned economic interpretation of different groups of market participants, and the fact that the mixture of normals distribution is extremely flexible, fat-tailed and asymmetric, thus easily able to capture the distributional regularities of financial returns data. A third benefit of the model is that it automatically induces time-varying skewness and kurtosis, which have been advocated in this context by [41], [42], [68], and [11]. Finally, and of great practical importance, [37] and [50] have demonstrated that the model delivers highly competitive out-of-sample

⁷Similar results are also obtained in the equilibrium model developed in Chapter 4, where large drops in the simulated equilibrium price are a result of the presence of incomplete information among the relevant market players.

VaR forecasts.

We say time series $\{\epsilon_t\}$ is generated by an n -component MixN-GARCH(r, s) process if the conditional distribution of ϵ_t is an n -component mixed normal with zero mean, i.e.,

$$\epsilon_t \mid \mathcal{F}_{t-1} \sim \text{MN}(\boldsymbol{\omega}, \boldsymbol{\mu}, \boldsymbol{\sigma}_t^2), \quad (3.2)$$

where the mixed normal pdf is $\sum_{j=1}^n \omega_j \phi(z; \mu_j, \sigma_j^2)$, ϕ is the normal pdf, $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)'$ is the set of component weights such that $\omega_j \in (0, 1)$ and $\sum_{j=1}^n \omega_j = 1$, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)'$ is the set of component means, such that, to ensure $E[\epsilon_t] = 0$, $\mu_n = -\sum_{j=1}^{n-1} (\omega_j/\omega_n) \mu_j$, and $\boldsymbol{\sigma}_t^{(2)} = (\sigma_{1t}^2, \dots, \sigma_{nt}^2)' \in \mathbb{R}_+^n$ are the positive component variances at time t .

In order to model the dynamics in the second (and higher) moments of the returns, the $n \times 1$ component variances $\boldsymbol{\sigma}_t^{(2)}$ are allowed to evolve according to the GARCH-like structure

$$\boldsymbol{\sigma}_t^{(2)} = \boldsymbol{\gamma}_0 + \sum_{i=1}^r \boldsymbol{\gamma}_i \epsilon_{t-i}^2 + \sum_{j=1}^s \boldsymbol{\Psi}_j \boldsymbol{\sigma}_{t-j}^{(2)}, \quad (3.3)$$

where $\boldsymbol{\gamma}_i = (\gamma_{i1}, \gamma_{i2}, \dots, \gamma_{in})'$, $i = 0, \dots, r$, are $n \times 1$ vectors, and $\boldsymbol{\Psi}_j$, $j = 1, \dots, s$, are $n \times n$ matrices with typical entry $\psi_{j hk} = [\boldsymbol{\Psi}_j]_{hk}$ (which we write as ψ_{hk} when, as in most applications, $s = 1$). We restrict $\boldsymbol{\Psi}_j$ to be diagonal, which, as discussed in [37], yields a much more parsimonious model with little loss in goodness-of-fit. In this case, and with $r = s = 1$, the parameter constraints $\gamma_{0i} > 0$, $\gamma_{1i} \geq 0$, and $\psi_{ii} \geq 0$, $i = 1, \dots, n$, are necessary and sufficient to ensure the nonnegativity of the variance terms. With one component ($n = 1$), the model reduces to the standard GARCH model. With two or more components, the model is able to capture the asymmetry and most of the excess kurtosis common in normal-GARCH residuals. Moreover, with $n \geq 2$, the structure of (3.3) also gives rise to rich conditional dynamics in the 2nd, 3rd and 4th moments which cannot be modeled by the classic GARCH model of the form (3.1) with any distributional assumption,

but do appear in real financial returns data [37].

It has been found that the component of the mixture assigned to the most volatile observations often consists of randomly and infrequently occurring jumps in the volatility, so that a GARCH structure is not required. We denote by $\text{MixN}(n, g)$ the model given by (3.2) and (3.3), with n component densities, but such that only g , $1 \leq g \leq n$, follow a GARCH process (and $n - g$ components restricted to be constant). In the context of the SO₂ returns, the component which picks up the zeros will also not require a GARCH component, so that we only entertain models of the form $\text{MixN}(n, g)$, $1 \leq g < n$, for $n \geq 2$, and, for comparison purposes, the $\text{MixN}(1, 1)$, which is just the standard GARCH model.

In line with the vast majority of studies involving the standard GARCH model and those involving the $\text{MixN-GARCH}(r, s)$, the choice $r = s = 1$ has been found to be adequate for the SO₂ returns, and we subsequently suppress reference to r and s . As is common in GARCH applications, the AR(1) structure $r_t = a_0 + a_1 r_{t-1} + \epsilon_t$ is included in the model to pick up the extremely mild autocorrelation structure in the mean. Thus, all future reference to a $\text{MixN}(n, g)$ model implies an AR(1)- $\text{MixN-GARCH}(1, 1)$ structure with diagonal Ψ_1 matrix. So, for example, in the $\text{MixN}(3, 2)$ case (which, anticipating our results below, is the preferred model), (3.3) takes the form

$$\begin{bmatrix} \sigma_{1t}^2 \\ \sigma_{2t}^2 \\ \sigma_{3t}^2 \end{bmatrix} = \begin{bmatrix} \gamma_{01} \\ \gamma_{02} \\ \gamma_{03} \end{bmatrix} + \begin{bmatrix} \gamma_{11} \\ \gamma_{12} \\ 0 \end{bmatrix} \epsilon_{t-1}^2 + \begin{bmatrix} \psi_{11} & 0 & 0 \\ 0 & \psi_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma_{1,t-1}^2 \\ \sigma_{2,t-1}^2 \\ \sigma_{3,t-1}^2 \end{bmatrix}. \quad (3.4)$$

In this case, there are 13 parameters to estimate, noting that μ_3 and ω_3 are constrained, as discussed above, and γ_{13} and ψ_{33} are held at zero.

In general, the number of components, n , needs to be determined empirically. As discussed in [37], standard model likelihood-based selection criteria can be successfully

employed to compare models with different numbers of components. For a model with K parameters, T observations and log-likelihood L , evaluated at the maximum likelihood estimator, we report the $AIC = -2L + 2K$ and $BIC = -2L + K \log T$; see [14] for original references and a detailed textbook presentation of the use of these measures.

The likelihood of the general $AR(p)$ -MixN-GARCH(r, s) model is straightforward to program and evaluate, and its numeric maximization has proven to be unproblematic using standard quasi-Newton-type optimization routines (as implemented in Matlab). One non-standard issue which arises in the context of the data in our study involves the point masses at zero, which are picked up as one of the n components in the MixN(n, g) model. Because these form a degenerate distribution, one variance component, namely γ_{03} in (3.4), is zero, and the likelihood is not defined. One way around this is to set γ_{03} and μ_3 to zero and ω_3 (the weight of this component) to the percentage of zeros in the data set. This turns out to be problematic because the other normal components (which are close to centered around zero) have a “gap” at zero, which (given the discrete nature of the returns data) renders the normal distribution inappropriate. Instead, we propose to replace the zero returns with realizations of i.i.d. normally distributed random variables with mean zero and standard deviation σ_k , where σ_k is chosen to be a small number relative to the unconditional variance of the returns.

At first blush, this appears to be an uncomfortable solution, because adding random noise to the data implies both a loss of “objectivity” as well as non-reproducibility of our estimation results. With respect to the first issue of objectivity, we note the relationship between this approach and that of [40], which is a quasi-Bayesian approach to estimating the (unconditional) mixed normal distribution involving (algebraically) adding observations to the data which reflect prior information and then maximizing a quasi-likelihood. His method not only results in greater numeric stability of estimation, but also (with a nonzero

and very small amount of prior information) leads to better small-sample estimation performance than pure maximum likelihood. In our model context, we observe the same results in terms of improved reliability of the numerical maximization of the likelihood and, as verified by simulations, more accurate parameter estimates. For the second issue of non-reproducibility, we remark that the parameter estimates are not sensitive (with respect to their approximate standard errors) to a range of σ_k values from 0.01 to 0.2. In what follows, we use this method with $\sigma_k = 0.1$.

3.1e Mixed Normal GARCH: Estimation Results and Diagnostics

Table 3.1 reports the likelihood-based goodness-of-fit measures for the various MixN(n, g) fitted models, as well as the t -GARCH. As expected, the worst performer is MixN(1,1), the standard (one-component) normal-GARCH model. Much better than the standard GARCH is the t -GARCH model, although because of the zero-returns issue discussed above, it performs, as expected, disastrously compared to the mixture models which, by design, can pick up the zeros.

To help confirm that there is no structure in the pattern of zeros throughout the return series, we also estimated the MixN(2,2) model. Observe that its likelihood is virtually the same as the MixN(2,1), showing that there is no GARCH dynamic in the component which picks up the zero-returns. From the table, we see that the best model according to both the AIC and BIC is the MixN(3,2). With respect to the AIC, the MixN(4,2) and MixN(4,3) models are close, while the BIC strongly favors the MixN(3,2). We subsequently restrict

Model	K	L	AIC	BIC
<i>t</i> -GARCH	6	-2968.6	5949.3	5982.16
MixN(1,1)	5	-4072.0	8134.0	8181.42
MixN(2,1)	8	-2919.6	5855.2	5899.07
MixN(2,2)	10	-2919.3	5858.6	5913.44
MixN(3,1)	11	-2873.8	5769.6	5829.93
MixN(3,2)	13	-2835.7	5697.4	5768.70
MixN(4,2)	16	-2834.5	5701.0	5781.75
MixN(4,3)	18	-2831.6	5699.2	5797.92

Table 3.1: Likelihood-based goodness-of-fit for the SO₂ return series. The best values for each criteria are marked in boldface.

attention to the MixN(3,2). The estimated volatility equation (3.4) for the MixN(3,2) is⁸

$$\begin{bmatrix} \sigma_{1t}^2 \\ \sigma_{2t}^2 \\ \sigma_{3t}^2 \end{bmatrix} = \begin{bmatrix} 0.621 \\ 0.121 \\ 0.013 \end{bmatrix} + \begin{bmatrix} 0.427 \\ 0.212 \\ 0 \end{bmatrix} \epsilon_{t-1}^2 + \begin{bmatrix} 0.846 & 0 & 0 \\ 0 & 0.649 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma_{1,t-1}^2 \\ \sigma_{2,t-1}^2 \\ \sigma_{3,t-1}^2 \end{bmatrix}, \quad (3.5)$$

with the remaining parameters in (3.2) given by $\hat{\omega}_1 = 0.165$, $\hat{\omega}_2 = 0.595$ (implying $\hat{\omega}_3 = 0.239$), $\hat{\mu}_1 = -0.001$, $\hat{\mu}_2 = 0.001$ (implying $\hat{\mu}_3 = -0.045$). Observe that the weight of the component associated with the zero-returns (the third component) is $\hat{\omega}_3 = 23.9\%$, which is, as expected, somewhat less than the unconditional (model free) estimate of the percentage of zeros (29%) in the data set, because the other two components are normal distributions with modes near zero, and thus account for some of the zero-returns.

As shown in Haas et al. (2004, Section 2.2), a measure of volatility persistence of the MixN-GARCH(1,1) model is the largest eigenvalue of $\Psi_1 + \gamma_1 \omega'$, which indeed reduces

⁸Parameter estimates and standard errors for the MixN(2,1), MixN(3,1) and MixN(3,2) models are tabulated in a working version of this chapter - available upon request.

for $n = 1$ component to the well-known persistence measure of $\theta_1 + \phi_1$ in the normal-GARCH(1,1) model with notation in (3.1). For our fitted MixN(3,2) model, this measure is 0.964, which, being less than unity, implies the model is covariance stationary. Looking separately at the two components driving the nonzero observations, we see that the first component, which accounts for an estimated $16.5/(16.5+59.5)=21.7\%$ of the volatility of the nonzero returns, has a persistence parameter of $\hat{\gamma}_{11} + \hat{\psi}_{11} = 1.273$. Thus, this GARCH component accounts for the extreme observations in the sample and, taken alone, is not covariance stationary. However, the fitted weight is sufficiently small so that the overall model is stationary. The second component has a persistence parameter of $\hat{\gamma}_{12} + \hat{\psi}_{22} = 0.861$, which is clearly picking up the milder movements in the returns. Typically, an estimated univariate GARCH model has a persistence parameter $\theta_1 + \phi_1$ very close to one (or equal to one; the so-called integrated, or IGARCH model), whereas the MixN-GARCH model can disentangle the volatility components into “mild” and “wild” ones, but still yielding an overall stationary process.

To help confirm that the fitted MixN(3,2) model successfully reflects the distributional properties of the SO₂ return series, Figure 3.4 shows a QQ-plot of the actual data and a (same length) simulated time series generated from the fitted model. The graph is typical of numerous generated ones, and shows that the entire distribution, most notably the tails, is well-captured by the proposed model.

Besides demonstrating that the MixN(3,2) model is a plausible approximation to the true (and undoubtedly far more complicated) DGP, the excellent fit in the tails shown in the QQ-plot has obvious implications for calculations of risk measures such as VaR. To substantiate this latter claim, we conduct the following out-of-sample forecasting exercise. Starting with the first 500 observations, we estimate the MixN(3,2) model and calculate the forecasted distribution function $\hat{F}_{501|500}$, and evaluate it at the observed return r_{501} . This is repeated for moving windows of length 500 until the end of the sample, with parameter

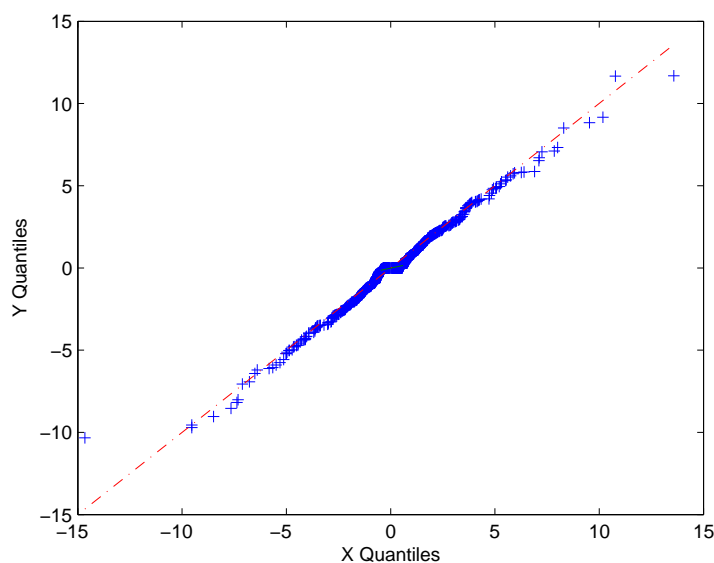


Figure 3.4: QQ-plot of a simulated time series versus the return series for the SO_2 price index.

re-estimation done every 20 steps. The resulting values indicate the quality of the one-step-ahead predicted VaR. In particular, if the model is accurate, then 1% of the 1,280 tail values should be less than 0.01. For the MixN(3,2) model, we obtain 0.94%, which is indeed extremely close to the nominal value of 1%.

To compare, we repeat the same exercise with the t -GARCH model. We argued above that, because of the zeros-problem, this model is highly misspecified and will not be expected to perform well. In this case, the t -GARCH model gave an average of 1.17% violations. The MixN(3,2) value is preferable both in terms of being closer to the nominal value of 1%, as well as being less than the target value instead of above it, as underprediction of risk can be costlier than overprediction. Nevertheless, the value of 1.17 is still reasonably

accurate; it is certainly not as bad as one might have expected from a demonstrably misspecified model, and this begs explanation. Figure 3.5 shows what is going on. The plot shows, for each of the two models, a graphical depiction of the deviations from their nominal values of *all* the empirical VaR values from under 1% to 10%. Thus, an ideal model would yield a straight line at zero, and in the plot, the above results for the 1% nominal VaR value are shown at the vertical dashed line.

We see that, indeed, the *t*-GARCH model performs well at and below the 1% nominal VaR value, but worsens in a steep linear fashion as the nominal value increases, indicating that the model fit is poor. This is not the case for the MixN(3,2) model, which is consistently accurate throughout the whole tail. Essentially, the *t*-GARCH model “got lucky” at 1% (which, perhaps conveniently, is among the most widely-used values), similar to many studies which have shown that the standard GARCH model happens to perform reasonably well at the 5% value. These VaR forecast results are consistent with the in-sample fits of the *t*-GARCH and MixN(3,2) models, as shown in Table 3.1.

In addition to an excellent in-sample fit, the mixture model in this context allows for a potential interpretation of the components. Clearly, the third component is used to pick up the zeros and embed them adequately in a stochastic process. The remaining two GARCH mixture components can be viewed as capturing the result of the two major groups of market participants: affected units who buy and sell permits based primarily on their current and forecasted needs (i.e., the permits are viewed as a factor of production), and speculative traders or simply non-affected agents (i.e., banks and investment funds).⁹ Of course, these two groups could possibly interact. Observe that the general mixed normal GARCH model (3.2) and (3.3) allows for a type of dynamic interaction between the components, though for the SO₂ data set, the diagonal model was statistically superior to the full model.

⁹Using a parallelism with Chapter 4, the two major groups may be firms which expect the market being in extreme shortage, and those which expect precisely the opposite, i.e. extreme excess of emission permits.

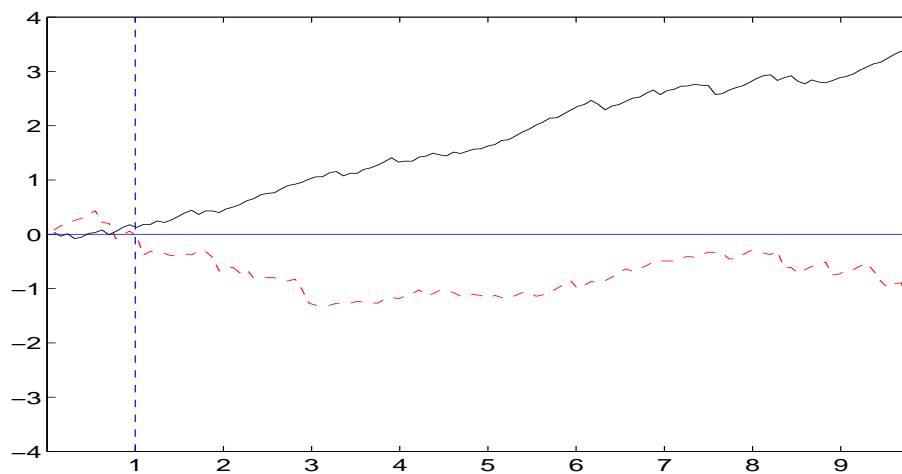


Figure 3.5: Deviation probability plot; solid is for t -GARCH, dashed is for MixN(3,2). Plotted values are “Deviation” $:= 100(F_U - \hat{F})$ (vertical axis) versus $100\hat{F}$ (horizontal axis), where F_U is the cdf of a uniform random variable; \hat{F} refers to the empirical cdf formed from evaluating the 1,280 one-step, out-of-sample distribution forecasts at the true, observed, next-day return.

3.2 Analysis of CO₂ Returns

For the CO₂ price series, we have only 454 daily returns; these are plotted in Figure 3.6. The larger presence of covered companies under the EU ETS translates into a higher daily traded emission volume and therefore into a much lower presence of zeros in the return time series (only 8). Without the zeros-problem, all conventional GARCH models can, in theory, be entertained, though the small sample size effectively prohibits use of the more elaborate models. In this study, we consider an AR(1)-GARCH(1,1) model with the following innovation distributions: Student’s t , symmetric and asymmetric stable, and the

generalized asymmetric t distribution. The latter, abbreviated GAt, has pdf

$$f_{\text{GAt}}(z; d, \nu, \theta) = K \times \begin{cases} \left(1 + \frac{(-z \cdot \theta)^d}{\nu}\right)^{-(\nu + \frac{1}{d})}, & \text{if } z < 0, \\ \left(1 + \frac{(z/\theta)^d}{\nu}\right)^{-(\nu + \frac{1}{d})}, & \text{if } z \geq 0, \end{cases} \quad (3.6)$$

for $d, \nu, \theta \in \mathbb{R}_{>0}$. Expressions for the integrating constant K , the cdf, moments, and the expected shortfall are given in Paoletta (2007, p. 273).

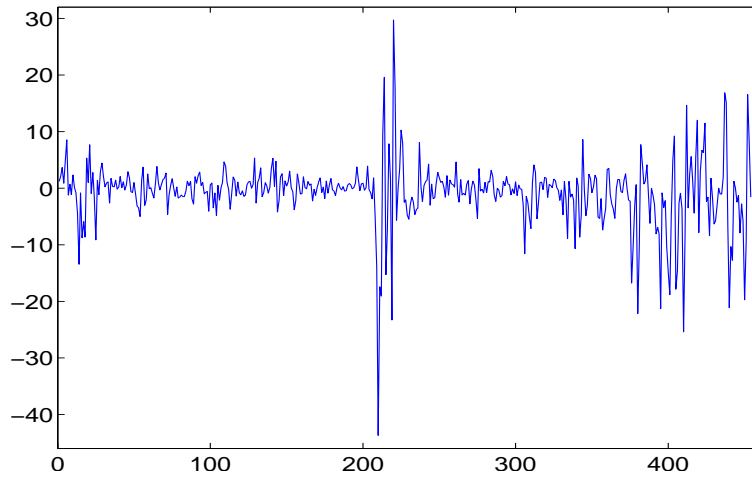


Figure 3.6: Daily CO₂ returns

The in-sample fits are summarized in Table 3.2. The best fitting model for all criteria is the GAt-GARCH.

We conduct a model comparison for assessing the forecasting quality of the models by repeating the exercise done for the SO₂ data. This resulted in Figure 3.7, which is similar to Figure 3.5 but corresponds to the CO₂ returns and is based on moving window of size 250 (and yielding 204 cdf values).

Model	K	L	AIC	BIC
normal-GARCH	5	-1260.8	2531.6	2552.2
t -GARCH	6	-1192.5	2397.1	2421.8
GAt-GARCH	8	-1181.9	2379.9	2412.8
$S_{\alpha,0}$ -GARCH	6	-1218.2	2448.5	2473.2
$S_{\alpha,\beta}$ -GARCH	7	-1192.0	2398.1	2426.9
MixN(3,2)	13	-1190.1	2406.2	2459.7

Table 3.2: Likelihood-based goodness-of-fit for the CO₂ return series. The best values for each criteria are marked in boldface.

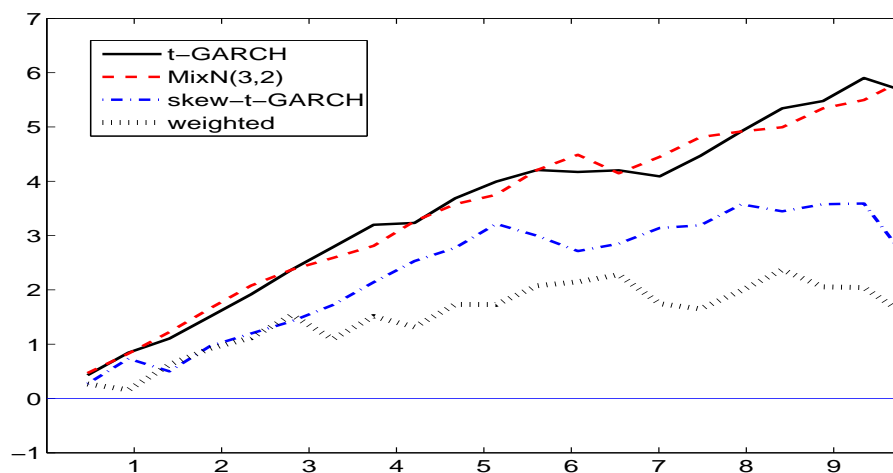


Figure 3.7: Deviation probability plot for CO₂ returns based on moving window of size 250; see caption in Figure 3.5 for description. Solid is for t -GARCH, dashed is MixN(3,2), dash-dot is GAt-GARCH, and dotted is GAt-GARCH using a weighted likelihood.

To avoid an overly cluttered graph, the figure just shows the VaR forecasting results for the t -GARCH, MixN(3,2), GAt-GARCH, and GAt-GARCH using a weighted likelihood,

the latter to be discussed below. The plots for the $S_{\alpha,0}$ -GARCH and $S_{\alpha,\beta}$ -GARCH (not shown) lie mostly between t and GAt-GARCH, with the latter having performed overall slightly better than $S_{\alpha,\beta}$ -GARCH. This agrees with the in-sample results reported above. We see that, of all the models so far discussed, all perform reasonably well near the 1% level, but none performs adequately for other tail values, lowering our confidence in their realistic applicability in practice. In this case, the MixN(3,2) performs worse than the stable and GAt models, but still on par with the t -GARCH. Similar results hold for MixN(2,2) and MixN(3,3) models.

The GAt-GARCH model has been found by other authors such as [56], [36] and [50] to deliver competitive VaR predictions. In our case, it performed overall best for the CO₂ data set, though still not adequately, and finding a better model seems necessary. One might argue that the short window size of 250, which corresponds to just one year of trading-day data, is too short for reliably estimating and forecasting any GARCH-type model. However, in their large study using NASDAQ index returns data, [50] shows that this need not be the case and VaR forecasts based on 250 observations can be highly accurate, and even outperform those based on windows of length 500 and 1000. As such, it is not necessarily the short window size, but rather that all the models we entertained are highly misspecified. To remedy this, one could try a full battery of GARCH and related models and choose the best-performing one, though such a data-snooping exercise may not be successful, and even if it is, it could lead to a model choice which is “too trained” to the current, very small, dataset and perform poorly in the future. To avoid this problem, we attempt to directly address the reason why the models do not perform well; namely, that the true data generating process (DGP) is changing too quickly over time. The CO₂ emissions market is rather new, and it is expected to be evolving, as do all emerging markets. Indeed, the particular dependencies of this market on political and regulatory uncertainties amplify the effects of news arrivals, such as the announcements in May 2006 that the market was overall

in emission excess. As such, it seems plausible that the DGP will experience significant changes in the short term.¹⁰

To deal with an unknown and changing DGP, we suggest choosing a tractable parametric structure which reasonably captures the salient features of the DGP and estimating it with more weight given to recent observations. Observe that, if the DGP were a GARCH model with parameters which vary smoothly over time in an unknown fashion, then this method acts as a way of balancing the tradeoff between using only very recent observations (and inducing a high variance in the estimates) and using all the data equally weighted (and thus delivering highly biased estimates). To negotiate this tradeoff optimally, we use the criteria of one-step VaR prediction with a geometric weighting scheme with weights $\tau_t \propto \rho^{T-t}$ (and then scaled to sum to one), where the single parameter ρ dictates the shape of the weighting function. Values of ρ with $0 < \rho < 1$ cause more recent observations to be given relatively more weight than values further in the past; $\rho = 1$ corresponds to standard ML estimation. Each value in a large grid of ρ values is used in the forecasting exercise and the one which delivers the best forecasts is used. This weighting scheme was shown in [56] to be effective in a financial returns density prediction context. It is interesting to note that the idea of weighting recent events more heavily is embedded by the decision weight function in Prospect Theory; see, e.g., [43].

Another source of inaccuracy in the prediction of downside risk is likely to stem from asymmetries in the data which are not adequately captured by the chosen model, even when the model allows for both an asymmetric response to shocks in the GARCH equation and a flexible asymmetric innovations distributional assumption. Indeed, if instead of accurate forecasting of the entire density, interest is restricted to just VaR, then it would seem wise to place more weight on the negative observations in the sample. This is demonstrated in

¹⁰To overcome the loose results of the econometric model for CO₂, we develop an equilibrium model for the price of the emission permits in Chapter 4.

[63], who investigate numerous schemes of placing more weight on negative returns. By combining the two weighting schemes, they show, using a variety of financial returns data sets and asymmetric GARCH models, that considerable improvement in forecast accuracy can be obtained which exceeds use of only one (or none) of the weighting schemes. In this paper, we use the following weighting scheme for placing more weight on the negative returns: first construct $\pi_t = 1 + r_t / \max_t(-r_t)$, and, for $\gamma \geq 0$, let

$$\omega_t = \begin{cases} \gamma(1 - F_{\text{Beta}}(\pi_t; p, q)) + 1, & \text{if } r_t < 0, \\ 1, & \text{otherwise.} \end{cases}$$

where F_{Beta} is the cumulative distribution function of a beta random variable with parameters p and q . The choice of γ determines how much weight is assigned to negative returns, with a value of zero yielding the default case of equal weights. In the final step the weights are standardized so that the sum of all weights (negative and positive) equals one, $\sum_{t=1}^T \omega_t = 1$. The appropriate value of γ is determined in the same fashion as parameter ρ in the time-weighting scheme.

The two weighting methods are easily combined by multiplying the two weight vectors elementwise, and then scaling so it sums to one. For a matrix of pairs of (ρ, ω) values, the forecasting exercise is conducted and that pair which yields the smallest VaR forecast deviation for VaR-probability values lower than 10% is chosen. Based on the 204 forecast values, we found that $\rho = 0.7$ and $\gamma = 4$ were optimal. Use of these values gives rise to the dotted plot in Figure 3.7, which offers an improvement over all the non-weighted models at all risk levels, notably 1% and 5%.

3.3 Conclusions

In this chapter, we rely on an empirical approach for the modeling and forecasting of the returns on the emissions permits. Because of the relatively high presence of zeros in the SO₂ return series, we argue and show that standard GARCH models cannot be used in this context. To handle this peculiarity, we propose and operationalize the use of the mixed-normal GARCH model, because it inherently can allocate a component to the extra mass around zero of the return distribution, unlike standard and even highly sophisticated GARCH models. Besides “solving the zeros problem”, the mixed-normal GARCH model provides both an excellent in-sample fit and out-of-sample VaR forecasts, and also allows for economic interpretation not shared by most GARCH models. As such, the model appears to be very well suited for modeling and forecasting the risk of the SO₂ emission market.

For the CO₂ data, no model which assumes a constant data generating process resulted in accurate VaR forecasts at all risk levels. By using a parametric model which places more weight, via the likelihood function, both on more recent returns and on negative returns, improved forecasting ability is achieved. However, since knowledge of the (unconditional and conditional) distribution of emission permit prices is essential for constructing optimal hedging and risk management strategies in the carbon market, further investigations are required.

One might suggest that the inclusion of a longer (up to the submission of this thesis, August 2008) and more updated time series could lead to some improvements. This would generate a fundamental mistake. The existence of banking restrictions (of the emission permits -see Chapter 2 for a detailed description of the market mechanisms and references) between the two first phases make *spot* contracts in phase I different from *spot* contracts in phase II. Therefore, a convenient merge of the two time series is simply not possible. Furthermore, since the emission spot price was overing above zero from mid 2007, the

inclusion of the rest of phase I time series into the analysis would have not enhance the results about the true DGP of the CO₂.

An interesting direction for future research would be to consider implementing the fundamental analysis (fuel prices are considered to be the primary approximate fundamentals for the emission permits - see section 4.2 for a discussion) into the mean equation of the return process and use GARCH-type structures, as discussed in this chapter, for the variance equation.

Chapter 4

Continuous Time Model: An Equilibrium Approach

*”Of course it was ambitious to set up a market for something you can’t see
and to expect to see immediate changes in behavior . . .”*

-Jacqueline McGlade, executive director, European Environment Agency

(June - 2008)

In Chapter 2 we discuss the environmental economics literature which focuses on the economic and policy aspects of marketable permits starting from the seminal paper of [60].¹ Theoretical aspects that Montgomery (1972) does not consider have been addressed

¹The original version of this chapter was co-written with Marc Chesney. It is included in the NCCR-FinRisk Working Paper Series under the title “The Endogenous Price Dynamics of Emission Allowances: An Application to CO₂ Option Pricing” . The paper has been presented at the following conferences: “IX-Workshop on Quantitative Finance” , January 2008, University of Rome “Tor Vergata” - Roma, Italy; “8th Ritsumeikan International Symposium on Stochastic Processes and Application to Mathematical Finance and 8th Columbia-Jafee Conference on Mathematical Finance” , April 2008, Ritsumeikan University, Kyoto - Japan; “EAERE - European Association of Environmental and Resource Economists” , June 2008, School of Business, Economics and Law, Goteborg - Sweden; and “35th European Finance Association meeting” , August 2008, Athens - Greece.

by the studies listed below and which are more extensively discussed in Chapter 2. The influence of uncertainty, regarding the regulation policy of public utility commissions on the behavior of regulated firms, have been discussed by [8] and [22]. Concentration in the permit market and in the output market have been exploited respectively by [38], [55] and by [52]. [73], [30] and [59] have developed models to include transaction costs in the theoretical frame. Finally, [16, 17] has conducted an analysis on auction and rules design.

Though such literature is fairly extensive, an explicit study of the dynamic of the price of the emission permits in the presence of market uncertainty is an almost unexplored area. Most of the present research relies on the key result - largely discussed in Chapter 2 - that, in an efficient market, the equilibrium price of the emission allowances is equal to the marginal costs of the cheapest pollution abatement solution. This statement underpins the belief that a high price level for emission permits brings about the relevant companies with lower marginal abatement costs in order to exploit consequent price differences. Such companies make profits by lowering the level of offending gases more than is necessary to comply with regulations and subsequently sell their spare permits. Through the use of optimal-control theory, [69] extends the discrete-time and deterministic setting of [77] and [26] and provides a trading model for permits in continuous time. The author, introducing the possibility of banking and borrowing permits, demonstrates that the discounted marginal costs of abatement are, theoretically, constant over time. As a consequence the permit price grows in equilibrium with discount rates (i.e. risk-free interest rates).

Recently, in an effort to bridge the gap between theory and observed market-price behavior, an increasing number of empirical studies has been investigating the historical time series of the permit price. In [28] several different diffusion and jump-diffusion processes were fitted to the European CO₂ futures time series. [6] analyze the short-term spot price behavior of CO₂ permits employing a Markov-switching model to capture the heteroskedastic behavior of the return time series. In contrast, in Chapter 3 we advocate the

use of a new GARCH-type structure for the analysis of inherent heteroskedastic dynamics in the returns of SO₂ in the U.S. and of CO₂ emission permits in the EU ETS.

As discussed in Chapter 2, with a precise focus on the European emission market and in an attempt to develop a valid dynamic price model, [71] and [34] elaborate a quantitative analysis of the CO₂ permits price founded on the pivotal results from environmental economics literature. In particular, [71] consider one representative agent who decides whether or not to spend money on lowering emission levels. The model is based on the optimal abatement decision of an affected company, therefore it very much depends on its total expected emissions. With a distinction between long-term and short-term abatement measures, [34] concentrate on the energy sector considering n affected utilities which decide their abatement levels by relying on the cheapest possible abatement option in the short-term, i.e. so-called fuel-switching.²

In common with the last-mentioned paper, we differentiate short-term and long-term abatement measures. As extensively discussed in section 4.2, a few options are available to the majority of affected companies and even fewer fall into the list of so-called short-term abatement possibilities. [18] prove that most abatement technologies in the U.S. energy sector are perceived as durable and irreversible investments whereas emission permits provide a greater flexibility in adapting to changing conditions. As a result, the compliance aim in the short run becomes a market problem, rather than a strictly engineering one. Accordingly, in our model we assume that the companies' pollution dynamics are exogenous processes. Relevant companies optimize their cost function by continuously adjusting their permit portfolio allocations and by choosing the optimal permit amount to purchase (in the

²It involves the replacement of high-carbon (sulfur) fuels with low-carbon (sulfur) alternatives. The most common form of fuel switching in the U.S. is the replacement of high-sulfur coal with a low-sulfur coal. In Europe, coal is typically replaced by natural gas.

shortage permit case) or to sell (in the excess permit case). The main objective is to develop an endogenous model for the permit-price dynamics. Remarkably, the underlying of the price of the emission permits are the net accumulated pollution processes of the relevant companies. Permit prices are sensitive to the strategic purchases and sales of all relevant companies. Moreover, each specific position in permits determines the final value of the emission price, as opposed to [34] where the boundary condition comes from the aggregate market position in permits.

The organization of the remaining sections of the chapter is as follows. In section 4.1, we briefly introduce the formal design of the markets for emission permits and we recall the EU ETS market characteristics. Section 4.2 addresses the fundamental distinction between long-term and short-term abatement policies. In section 4.3 we present the model and its formulation for the basic case of one company with emission-trading opportunity only at time zero. Then, in section 4.4, we extend the model to account for two-firm interactions that may continuously trade permits which are coupled with asymmetric information. In section 4.5 we give a proposition for the model with multi-firms. Finally, section 4.6 concludes with an extensive numerical exercise to derive the equilibrium price for CO₂ permits.

4.1 Marketable Permits for Air-control

A marketable permits scheme for air pollution control is constructed as follows. Marketable permits are issued to relevant facilities. These emission permits (or allowances) are denominated in units of a specific pollutant (for example in tons of CO₂) and in amounts proportional to their size and emissions according to a referred year as baseline. For a detailed discussion of initial allocation criteria see [3] and references therein.

At regular intervals, facilities submit emission reports for their compliance period, at

the end of which facilities must own sufficient permits to cover their emissions. This implies that each facility must hold at least as many valid credits as emissions during the compliance period. A penalty is levied if a facility does not deliver a sufficient amount of permits at the end of the compliance period. The payment of a fine does not remove the obligation to achieve compliance, which means that undelivered permits have to be handed in. Having been used to cover emissions, these “credits” are then deleted from the regulatory compliance system, preventing subsequent use or transfer. The compliance date marks the end of each period for which a facility has to file an emissions report, which is due on the certification date.

The largest and most important emission-trading program has been developed by the European Union to facilitate implementation of the Kyoto Protocol. The EU ETS covers five different industrial sectors and almost 12,000 installations in 27 countries, responsible for nearly half of the EU’s CO₂ emissions. They have been allocated permits giving them the right, over the first phase (2005-2007), to emit 6.6 billion tons of CO₂. The second phase coincides with the first Kyoto commitment period, beginning in 2008 and continuing through 2012. At the time of writing, ongoing negotiations are specifying the details of the imminent third phase.³ The EU ETS has created *de facto* property rights for emissions that are freely tradable. All permits are transferable, i.e. a facility that generates excess permits by reducing emissions below its allocated levels can sell those extra “credits” to other relevant entities. In addition to the so-called spatial trading,⁴ both schemes allow for

³It worth to mention that beside emission permits, relevant companies can also use “certificates” acquired from outside the European Union, via the Joint Implementation (JI) or the Clean Development Mechanism (CDM) to meet their obligations under the EU ETS. The Kyoto protocol allows the utilization of so-called *flexible mechanisms*. Through JI, developed countries can receive emissions reduction units whenever they finance projects that reduce net pollution emissions in other developed countries. Through CDM, developed countries may finance GHG emission reduction or removal projects in developing countries, and receive credits for doing so. Interested readers may refer to the Intergovernmental Panel on Climate Change (IPCC) for relevant scientific, technical and socio-economic information related to these flexible mechanisms.

⁴According to environmental terminology, spatial trading means that a unit can reduce its emissions below

inter-temporal trading, so that companies can save their permits for use in the future. This is reflected by a larger time flexibility for pollution-control investments. More precisely, the EU ETS allows only *within-phase* banking, i.e. permits can be banked from one year to the next. Unused allowances, however, are not valid during the following phase.

The economic incentives embedded in the tradable permits are designed to force companies to participate in the permits market. As discussed in Chapter 2, this leads to a theoretical equalization of marginal abatement costs across different pollution sources. However, in the short run, the observed permit price does not coincide with the expected theoretical level.⁵ Though this might be ascribed to a market which is in the initial stage of development, in the next section we will attempt to address directly the reasons why this mismatch is present.

4.2 Abatement Opportunities in the Short Term

According to the market-based approach which we have described in Chapter 2, a generating unit is endowed with high flexibility in determining the best strategy of achieving compliance under a cap-and-trade program: each firm faces a basic choice between buying (or selling) permits, and reducing emissions through use of alternative technologies. Three general classes of techniques for the physical reduction of emissions are available. Firstly, emissions can be reduced by lowering the output scale. Secondly, the production process or the inputs used - for example, fuels - can be altered. Finally, tail-end cleaning equipment can be installed to remove pollutants from effluent streams before they are released into the

its allocated number of permits, transferring those unused to other units within the same company or selling them to other companies or brokers. Conversely, it can decide not to abate its emissions but to purchase permits covering emissions above its allocation. Spatial trading along with temporal trading are one key characteristics of marketable permits, see Chapter 2.

⁵It should be noted that in the long run too, the equality permit price and cheapest marginal abatement cost hold only in the presence of a non-evident permit-excess situation - see section 4.6 for numerical results.

environment.

European firms, in order to accomplish Europe's severe environmental regulations, have mostly achieved high environmental standards either in production processes or in the reduction of offending gases released as a byproduct into the air. Due to this advanced technological situation, and coupled with a typical inelastic demand for particular products – such as electricity or ore-mining materials⁶ – the first abatement alternative can be considered as the exception rather than the rule, see [44] and [76] for a more comprehensive discussion.

Although in the EU ETS the largest-affected sector, which received the lower amount of initial permits, i.e. the fuel-burning energy producers, has one of the cheapest abatement alternatives (so-called fuel-switching), it is worth noticing that fuel-switching is rarely implemented at present. A possible reason for this is that medium-sized to large utilities purchase fuel signing-up contracts with long maturities in order to lock in a particular price premium, providing in such a way an element of irreversibility to fuel-switching decisions -see [45]. More plausibly, since fuel switching is generally implemented whenever there exists a sensible fuel-input price difference, the actual CO₂ price/cost ratio does not currently trigger the hypothetical *daily* fuel-switching. In fact, considering the marginal optimization procedures, fuel-switching is mainly dependent on the marginal generation costs. CO₂ is also an important cost-component but a plausibly more relevant component is the cost associated to the change of fuel. Moreover, a firm cannot fuel-switching instantaneously: the process requires implementation-time. We concentrate our attention, therefore, on the short-term period. A reference to the possible incorporation of the typical production management decisions based on daily CO₂ price movements is available in section 4.3.

⁶The largest permit allocations correspond respectively to iron and steel producers, non-metallic mineral producers and energy producers.

A market-based approach leads to an efficient allocation of abatement costs across different pollution sources, as shown by [60]. However, this heavily depends on the implicit assumption that any technological abatement solution, for instance the installation of scrubbers on smokestacks to extract noxious fumes as solid residues,⁷ is perceived as a perfect substitute for emission allowances. This only holds true in an efficient market with no uncertainty. Those facilities which are affected, on the contrary, face considerable uncertainty. [18] show that companies perceive abatement technologies - in particular scrubber plants for SO₂ - as inferior substitutes for emission allowances. In contrast to emission permits, investments in pollution-reduction infrastructures are irrevocable commitments which last for decades and typically need some lead time in order to become effective. For a more extensive discussion refer to [33] and [80]. The purchase of permits is adjustable to changing market conditions whereas a scrubber might be under-utilized if demand falls. Moreover, the cost of a scrubber might be excessive following a fall in permit price. Hence, since pollution abatement technologies are often expensive, durable and irreversible investments, they are not commonly deemed to be perfect substitute for emission permits. Plausibly, other sources of uncertainties - regulatory uncertainty, for instance - can distort the theoretical equilibrium price, but the overall effect would always be a mismatch.

Following this line of reasoning, we develop an equilibrium model for the short-term permit price and we propose possible model extensions for the inclusion of general technological abatement measures coupled with long-run management strategies.

⁷It is important to note that currently there is no commercially available *end-of-the-stack* technology to extract CO₂.

4.3 The Formal Model: "Wait-and-see" for One Company

In the tradable permit price modeling, as outlined by [60], the existence of an efficient market has been generally assumed. This leads to an equalization of marginal abatement costs across the different pollution emitters and to an emergence of an alignment of companies' interests with those of a representative agent (as in [71]), or with a social planner (as in [34]).⁸ Employing the existence of a single representative firm in the market (as in [71]), we model the permit price process in an elementary situation where trading is only possible at the inception of an environmental program that has a finite length T . To simplify matters, we do not account for the possibility of trading the emission certificates generated by JI and CDM projects. The study of their stochastic impact on the emission market is left for future research. Additionally, in addressing the cost minimization problem, we derive the permit price in analytic form.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space, $\mathcal{F} = (\mathcal{F}_t)$ the filtration where $\mathcal{F}_0 = \sigma(Q_0)$. We denote with Q_0 the initial pollution level and with X_0 the quantity of permits that the company buys ($X_0 > 0$) or sells ($X_0 < 0$) at time zero, and with N the initial permits endowment. We label δ_0 the overall net amount of permits for the company at initial time, where $\delta_0 = N + X_0$ and it gives the company the right to emit a volume of offending gases up to such a level. We assume that the firm continuously emits offending gas according to a stochastic exogenous process over the period $[0, T]$. The process evolves accordingly to a geometric Brownian motion:⁹

$$\frac{dQ_t}{Q_t} = \mu dt + \sigma dW_t, \quad \text{or equivalently} \quad Q_t = Q_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t} \quad (4.1)$$

⁸In [34] the coincidence of the equilibrium permit price with the solution of social planner problem is a result of the model since fuel-switching is considered as a perfect substitute of emission permits.

⁹The assumption of an (on average) increasing pollution process is aligned with the United Nation Framework on Climate Change (UNFCCC) estimations on future consumption of energy, cement and steel.

where μ and σ are respectively the instantaneously constant drift term and the constant volatility of the pollution process. Q_0 can be interpreted as the business-as-usual (BAU) emission level, while the drift and the volatility characterize the gas emission process with respect to the BAU pollution level. Using Equation (4.1), the accumulated pollution volume at time t is simply $\int_0^t Q_s ds$. Thereby, a negative μ implies a lower (expected) accumulated pollution level - maybe due to a technological improvement - that is reflected in future GHG emissions; whereas σ measures the uncertainty about the accumulated pollution volume. A natural extension of the model would be the introduction of an endogenous pollution process. This means that a company can modify its production process according to the level of accumulated pollution and consequently impact the permit price evolution. We leave this exploration for future research.

As described in section 4.1, in order to pollute *legally*, the company must have enough permits by the end of the period T . If the firm fails to achieve compliance, it will pay a penalty equal to P . More precisely, in the EU ETS penalty costs may occur at the end of every year. However, the European Directive allows a one-year borrowing within a trading period. This means that companies are allowed to use permits with future maturity for compliance in the current year without having to buy the permits in the market. It is thus not unreasonable to assume that companies will not pay penalties for a shortfall within a particular trading period.

At the end of the period we expect either a shortage or a surplus situation (or possibly a perfect match) between the issued emission permits and the verified pollution level. Inevitably, the company will either be holding worthless emission permits or paying the price for being uncovered - i.e. the penalty P times the number of uncovered tons - or be totally and perfectly hedged. Yet, as this last possibility is quite unlikely, the final cash outflow boils down to a binary outcome. In fact, the company's final cost in a *wait-and-see*

situation without any trading opportunity during the period $[0, T]$ is:

$$\max \left\{ 0, \left(\int_0^T Q_s ds - \delta_0 \right) \right\} \cdot P, \quad (4.2)$$

where $\int_0^T Q_s ds$ is the firm final accumulated pollution level. Expression (4.2) recalls a typical option payoff. From here it is obvious that emission permits - like many other tradable permits - are to all intents option contracts. Several features shared with standard options contracts are discussed in the forthcoming numerical section.

Given the initial endowment of permits and the expected future permits net position, a firm minimizes its costs at the inception of the period. The total cost is simply the sum of the cash-flows at initial time (or minus the proceeds from permits sales) and the potential penalties at the end of the program. Therefore, the resulting minimization problem is:

$$\min_{\{X_0\}} \left\{ S_0 \cdot X_0 + e^{-\eta T} \mathbb{E}_{\mathbb{P}} \left[\left(\int_0^T Q_s ds - \delta_0 \right)^+ \cdot P \mid \mathcal{F}_0 \right] \right\} \quad (4.3)$$

where the expectation is taken under the historical probability measure \mathbb{P} ,¹⁰ η is the discount rate - the weighted average cost of capital - and S_0 is the permit price (known) at time $t = 0$.

In order to express the permit price in analytic form, we rely on [35] and write the objective function as follows:

$$\mathbf{H} \equiv \left\{ S_0 \cdot X_0 + e^{-\eta T} \mathbb{E}_{\mathbb{P}} \left[\left(\int_0^T Q_s ds - N - X_0 \right)^+ \cdot P \right] \right\}$$

¹⁰The historical probability \mathbb{P} refers to the pollution processes, a non-tradable *asset*. There is no need to construct a risk-neutral probability measure for the pollution processes since a corresponding risk-neutral pollution dynamics has no reason to be evaluated.

$$\text{with } \int_0^T Q_s ds = \frac{4}{\sigma^2} \cdot Q_0 \int_0^{\sigma^2 T/4} e^{2(\widetilde{W}_u + zu)} du = \frac{4}{\sigma^2} \cdot Q_0 \cdot A_{\sigma^2 T/4}^z$$

$$z := \frac{2\nu}{\sigma}, \quad \nu := \frac{1}{\sigma} \cdot \left(\mu - \frac{\sigma^2}{2}\right), \quad \text{and } \widetilde{W}_u := \frac{\sigma}{2} W_{4u/\sigma^2} \text{ is a Brownian motion.}$$

Finally, we denote $A_T^\nu = \int_0^T e^{2(W_s + \nu s)} ds$.

Computing the first order condition (FOC), X_0 satisfies the following equation (the detailed derivation is in Appendix A.1):

$$S_0 = e^{-\eta T} \cdot P \cdot \int_{\delta_0 \cdot \sigma^2/4 Q_0}^{\infty} \mathbb{P} \left[A_{\sigma^2 T/4}^z \in dx \right]. \quad (4.4)$$

It is observable that the emission allowance spot price is a function of the penalty level and the probability of a permit–shortage situation. The functional form of such probability is known, but unfortunately is problematic to evaluate numerically. For illustrative purposes, therefore, we let T be an arbitrary small-time interval ($\Delta t = T$) and then compute the discrete approximation of $\int_0^T Q_s ds$. This enables us to derive a more intuitive analytical form for the permit spot price (the detailed derivation is in Appendix A.1):

$$S_0 = e^{-\eta T} [P \cdot \Phi(d_-)], \quad \text{where } d_- = \frac{\ln(Q_0 \cdot \Delta t / \delta_0) + (\mu - \frac{\sigma^2}{2}) \Delta t}{\sigma \sqrt{\Delta t}}, \quad (4.5)$$

and $\Phi(x)$ is the standard cumulative distribution function $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$.

Equation (4.5) is the discounted value of future expected expenses, i.e. the *contingent claim* payoff, where the claim is the emission permit whose value is contingent on the shortage event ω , with $\omega = \{\int_0^T Q_s ds > \delta_0\}$.¹¹

In figure 4.1 we give a graphical interpretation of Equation (4.5). In the right picture the solid line is the permit price as a function of X_0 in a zero-volatility situation. This

¹¹It is worth noting that Equation (4.5) corresponds to the equilibrium permit price described in Theorem 1 of [34].

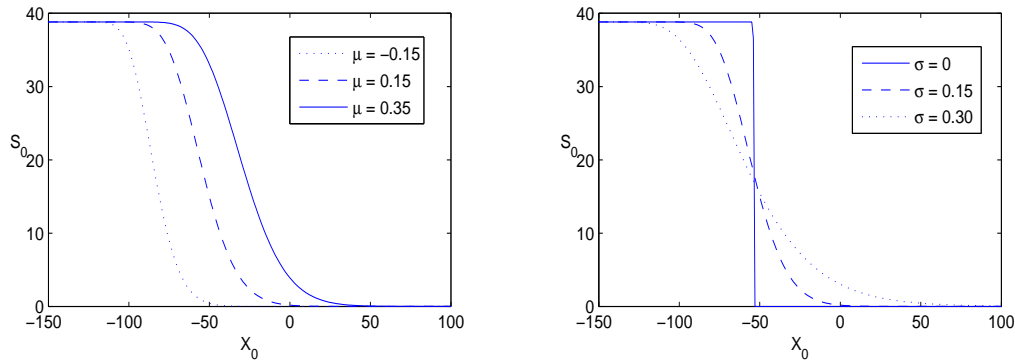


Figure 4.1: Plot of S_0 , the emission permit price at initial time, as a function of X_0 , the amount of permits sold (negative values) or bought (positive values) at time $t = 0$. We plot the permit price for different $\{\mu : \mu \in \mathbb{R}\}$, left picture, and different $\{\sigma : \sigma \in \mathbb{R}^+\}$, right picture, keeping all other parameters constant. In this example, $N = 170$, $P = 40$ and $Q_0 = 100$.

is equivalent to the exclusion of uncertainty about the final accumulated pollution level of the firm. More precisely, in this particular example, if the company sells less than 50 units of permits, it surely ends up in a permits-excess situation. As a result, the firm achieves compliance - it has no penalty to pay. However, the permits that are left over have no longer any monetary value (this potential loss is correctly taken into account in the extended model version). Conversely, if the firm sells more than 50 units of permits, ending up with a permits-shortage, it pays the penalty for each uncovered ton of offending gas. In this last situation, the permit has a value equal to the amount the firm is obliged to pay: the discounted penalty P . Theoretically, in a world of certainty, we can evaluate a quantity-threshold: on the left-hand side the emission allowance has a value equal to $e^{-\eta T} \cdot P$, while on the right-hand side any permit is worthless.

As any pollution process is marked by uncertainty, the emission allowance price lies

somewhere between zero and the discounted penalty level, i.e. $S_0 \in [0, e^{-\eta T} \cdot P]$. As figure 4.1 shows, the difference between the accumulated pollution volume and the net amount of permits leads either to a shortage or to a surplus. This depends on the the probability of such an event, and determines where the price lies. Keeping all other parameters constant, we plot the permit price S_0 for different drift parameter values on the right-hand side. On the left-hand side we use different instantaneous volatility values. As expected, the higher is μ or σ the larger has to be the amount of permits to be bought at time $t = 0$. Similarly the lower has to be the amount of permits to be sold - in order to reduce the risk of facing a penalty. It is worth noticing that this figure resembles the graphical results of the equilibrium spot price in [71].

4.4 Two-companies and Multi-periods Trading

A market for tradable permits is clearly different from the oversimplified situation described above. Not just one representative agent, but different companies operate at the same time on the market. The resulting interaction of the companies' optimization strategies must be properly taken into account as anticipated in section 2.2. In addition, the inherent uncertainty associated with almost all emission levels affects the market efficiency, as discussed in section 4.2. Several technical, commercial, and operational factors contribute to the uncertainty observed in emission levels and to the perception of a larger flexibility for the emission permits compared to other abatement measures. These factors include uncertainty in the demand for companies' goods and services. This results in a variation in the production activity levels, measurement and monitoring uncertainty. These, coupled with imperfect information regarding emission levels, typically lead to either the facilities ending up short of, or in excess of emission permits. Both of these are highly undesirable scenarios. The former results in excessive emissions in the environment in conjunction

with high violation penalties for the facilities while the latter represents unrealized productive and/or market value for the firm. As a result, facilities are forced to participate in the market in order to reconcile their emission credit accounts. They do this either by selling or buying permits. Historical price evidence suggests that many of the affected firms dynamically adjust their positions, thus ensuring compliance. On the one hand some companies continuously purchase/sell the difference between their permit allocation and their expected net future emission. On the other hand, other players take a more speculative approach by selling off permits when the permit price is high, and purchasing them back later on if in a permit need situation or if the permit price is conveniently low.¹² In what follows we extend the basic model, accommodating it to the interaction of two firms that trade in a multi-period setting and to the presence of asymmetric information.¹³

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ be the probability space, $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ is the filtration where $\mathcal{F}_t = \sigma(\cup_{i \in \mathcal{I}} Q_s^i, s \in [0, t])$, and $\mathcal{I} = \{1, 2\}$. Each firm continuously emits offending gas accordingly to an exogenous process:

$$\frac{dQ_{i,t}}{Q_{i,t}} = \mu_i dt + \sigma_i dW_{i,t}.$$

where we assume $dW_{1,t} \cdot dW_{2,t} = 0$. We denote with $X_{i,t}$ and $N_{i,0}$ respectively the quantity of permits that the i -th company buys or sells and the initial permits endowment. In a cap-and-trade, as the EU ETS, the GHG reduction target is settled at the inception of each phase, therefore the supply side of pollution permits is indeed fixed: $N = N_{1,0} + N_{2,0}$.

¹²An analysis of the interests of the various players in the market (governments, financial institutions, industrials and energy companies and NGOs) might lead to a different interpretation of permit price dynamics in the EU ETS.

¹³Asymmetric information means that firms found their permit trading strategies based on different information sets.

The net amount of permits that the i -th company possesses at time t is denoted by

$$\delta_{i,t} := N_{i,0} + \sum_{s=0}^t X_{i,s} = N_{i,t-1} + X_{i,t}, \quad \forall \quad t = 1, 2, \dots, T-1,$$

where we define $N_{i,t-1}$ as the sum of the marginal quantities of emission permits bought and sold by company i -th including the initial permit endowment.

Given that the total number of permits is fixed, the market clearing condition is:

$$\delta_{1,t} + \delta_{2,t} = N \quad \text{or in another form} \quad X_{1,t} = -X_{2,t} \quad \forall \quad t = 0, 1, \dots, T-1. \quad (4.6)$$

Condition (4.6) implies that in equilibrium the permit positions are in zero net supply. Hence, it satisfies the competitive equilibrium condition that requires equality between supply and demand for pollution rights in the market.

We label the i -th net accumulated pollution volume at time t as $\int_0^t Q_{i,s} ds - \delta_{i,t-1}$. At time $t \in [0, T]$, company i has complete knowledge about its own net accumulated pollution volume, $\int_0^t Q_{i,s} ds - \delta_{i,t-1}$, and partial knowledge about the net accumulated pollution volumes of company j , $\int_0^{t-1} Q_{j,s} ds - \delta_{j,t-1}$. In other words, the presence of asymmetric information imposes a lag-effect on the expected future net-emission levels of the other company. A ready extension of the model to I companies is possible splitting the set $\mathcal{I} = \{1, 2, \dots, I\}$ into two parts, $I^- := \mathcal{I} - i$ and i , and assuming equal information among the companies in I^- . Moreover, using constant drift and volatility terms, $\{\mu, \sigma\} \in \mathbb{R}^{I-1}$, and relying on standard technique of the methods of moments, one can approximate the cumulative pollution process, $Q_{I^-,t} = \sum_{j=1, j \neq i}^I Q_{j,t}$, with a new geometric Brownian motion, see [12]. However, we focus on the case of $\mathcal{I} = \{1, 2\}$ and we formalize the extension of the model in the following theorem.

At time T , if neither of the company is in a permit need, all left-over permits have

zero value. Conversely, if at least one of the firms is in permit shortage, since by law all covered companies have to surrender sufficient credits at time T , the permit has a value equal to the penalty level P . This holds assuming that there is no presence of market-power and each company in shortage is indifferent to purchase permits and to penalty payments. Analytically, the permit value at time T is:

$$S_T = \begin{cases} 0 & \text{if } \forall i \in I \quad \int_0^T Q_{i,s} ds \leq \delta_{i,T-1} \\ P & \text{if } \exists i \in I \quad \int_0^T Q_{i,s} ds > \delta_{i,T-1} \end{cases} \quad (4.7)$$

In accordance with the emission market construction at time T , if company i is in permit excess, it can sell to company j what the latter wants to buy:

$$\min \left\{ \left(\delta_{i,T-1} - \int_0^T Q_{i,s} ds \right)^+, \left(\int_0^T Q_{j,s} ds - \delta_{j,T-1} \right)^+ \right\} =: \Gamma \quad (4.8)$$

On the other hand, if company i is in permit shortage, it can buy from company j what the latter wants to sell:

$$\min \left\{ \left(\int_0^T Q_{i,s} ds - \delta_{i,T-1} \right)^+, \left(\delta_{j,T-1} - \int_0^T Q_{j,s} ds \right)^+ \right\} =: \Pi \quad (4.9)$$

However, if $\Pi < \left(\int_0^T Q_{i,s} ds - \delta_{i,T-1} \right)$, by law company i has to pay P for each uncovered ton of offending gas emitted. Thus, combining equations (4.8) and (4.9), a severe asymmetry between buyer and seller positions is detectable and the boundary conditions for the permit-quantity at time T can be simplified to:

$$X_{i,T} = \left(\int_0^T Q_{i,s} ds - \delta_{i,T-1} \right)^+ - \Gamma, \quad \forall i \in I.$$

Let us consider Δt to be the unit time. As in the previous section, given the initial permit

endowments and expectations on the accumulated pollution volumes, each firm minimizes its total costs at every time $t \in [0, T - \Delta t]$. The minimization problem for company $i = 1$ at time $T - \Delta t$ is:

$$\min_{\{X_{1,T-\Delta t}\}} \left\{ S_{T-\Delta t} \cdot X_{1,T-\Delta t} + e^{-\eta\Delta t} \mathbb{E}_{\mathbb{P}} [S_T \cdot X_{1,T} | \mathcal{F}_{T-\Delta t}] \right\}.$$

And deriving the FOC:

$$\begin{aligned} \bar{S}_{T-\Delta t} &= e^{-\eta\Delta t} \cdot P \cdot \mathbb{E}_{\mathbb{P}} \left[\mathbb{1}_{\int_0^T Q_{1,s} ds > \delta_{1,T-\Delta t}} | \mathcal{F}_{T-\Delta t} \right] \\ &+ e^{-\eta\Delta t} \cdot P \cdot \mathbb{E}_{\mathbb{P}} \left[\mathbb{1}_{\delta_{1,T-\Delta t} > \int_0^T Q_{1,s} ds} \cdot \mathbb{1}_{\int_0^T Q_{2,s} ds > \delta_{2,T-\Delta t}} | \mathcal{F}_{T-\Delta t} \right]. \end{aligned} \quad (4.10)$$

Since $\int_0^t Q_{i,s} ds$ is a monotonically non-decreasing function in t , it follows

$$\mathbb{E}_{\mathbb{P}} \left[\mathbb{1}_{\int_0^T Q_{1,s} ds > \delta_{1,T-\Delta t}} | \mathcal{F}_{T-\Delta t} \right] = \begin{cases} \Phi(d_{1,T-\Delta t}) & \text{if } \int_0^{T-\Delta t} Q_{1,s} ds - \delta_{1,T-\Delta t} \leq 0 \\ 1 & \text{else} \end{cases}$$

$$\text{where } d_{1,T-\Delta t} = \frac{\ln \left(\frac{Q_{1,T-\Delta t} \cdot \Delta t}{N_{1,T-2\Delta t} + X_{1,T-\Delta t} - \int_0^{T-\Delta t} Q_{1,s} ds} \right) + \left(\mu_1 - \frac{\sigma_1^2}{2} \right) \cdot \Delta t}{\sigma_1 \cdot \sqrt{\Delta t}}.$$

And, by independence:¹⁴

$$\mathbb{E}_{\mathbb{P}} \left[\mathbb{1}_{\delta_{1,T-\Delta t} > \int_0^T Q_{1,s} ds} \cdot \mathbb{1}_{\int_0^T Q_{2,s} ds > \delta_{2,T-\Delta t}} | \mathcal{F}_{T-\Delta t} \right] = \Phi(-d_{1,T-\Delta t}) \cdot \Phi(d_{2,T-\Delta t}^{\text{lag}}),$$

$$\text{where } d_{2,T-\Delta t}^{\text{lag}} = \frac{\ln \left(\frac{Q_{2,T-2\Delta t} \cdot 2\Delta t}{N_{2,T-2\Delta t} + X_{2,T-\Delta t} - \int_0^{T-2\Delta t} Q_{2,s} ds} \right) + \left(\mu_2 - \frac{\sigma_2^2}{2} \right) \cdot 2\Delta t}{\sigma_2 \cdot \sqrt{2\Delta t}}.$$

¹⁴From a practical point of view, the EU ETS covers five different industrial sectors and almost 12.000 installations in 27 European countries. So, it is plausible that two companies, although belonging to the same industrial sector, are affected by different technical, commercial and operational factors.

Moving on from this, we can then express the emission spot price analytically for company one at time $T - \Delta t$ as the discounted penalty level weighted by the shortage probabilities (for the computations see Appendix A.2):

$$S_{T-\Delta t} = e^{-\eta\Delta t} \cdot P \cdot [1 - \mathbb{P}_{T-\Delta t}^1] \quad (4.11)$$

where $\mathbb{P}_{T-\Delta t}^1 := \Phi(-d_{1,T-\Delta t}) \cdot \Phi(-d_{2,T-\Delta t}^{\text{lag}})$ is the probability of non-shortage future situations for both companies from the point of view of company one.

Similarly, solving the optimization problem for company two, it follows:

$$\bar{S}_{T-\Delta t} = e^{-\eta\Delta t} \cdot P \cdot [1 - \mathbb{P}_{T-\Delta t}^2], \quad (4.12)$$

where $d_{2,T-\Delta t}$ and $d_{1,T-\Delta t}^{\text{lag}}$ are defined similarly as above. $\mathbb{P}_{T-\Delta t}^2 := \Phi(-d_{2,T-\Delta t}) \cdot \Phi(-d_{1,T-\Delta t}^{\text{lag}})$ is the probability of non-shortage future situations for both companies from the point of view of company two. For the sake of simplicity, we use the same discounting factor η for both companies. A generalization taking two different discounting factors is straightforward.

Moving backwards and repeating the optimization procedure at each time step $k \in [1, 2, \dots, T/\Delta t]$, we obtain a pair ($i \neq j$) of emission price equations (see Appendix A.3):

$$\bar{S}_{T-k\Delta t} = e^{-\eta k\Delta t} \cdot P \cdot \left\{ 1 - \mathbb{E}_{\mathbb{P}} \left[\Phi(-d_{i,T-\Delta t}) \cdot \Phi(-d_{j,T-\Delta t}^{\text{lag}}) \middle| \mathcal{F}_{T-k\Delta t} \right] \right\}. \quad (4.13)$$

With these two equations and the market clearing condition (4.6), at each time step we determine the equilibrium permit price by numerically evaluating the quantity of permits

that satisfies the following equality:

$$\mathbb{E}_{\mathbb{P}} \left[\Phi(-d_{i,T-\Delta t}) \cdot \Phi(-d_{j,T-\Delta t}^{\text{lag}}) | \mathcal{F}_{T-k\Delta t} \right] = \mathbb{E}_{\mathbb{P}} \left[\Phi(-d_{j,T-\Delta t}) \cdot \Phi(-d_{i,T-\Delta t}^{\text{lag}}) | \mathcal{F}_{T-k\Delta t} \right], \quad (4.14)$$

for a given set of parameters $(\{\mu, \sigma, Q_0, N_0\} \in \mathbb{R}^2)$ that characterize the two pollution processes.

4.5 Multi-firm and Multi-periods Trading

Along similar lines of the previous sub-section, and splitting the set \mathcal{I} into two parts ($\mathcal{I} = I^- \cup i$), we can generalize the model to I companies in a multi-period setting. The equilibrium permit price result from the solution of a system of I equations (see Appendix A.4).

Proposition 4.5.1 *Given the exogenous pollution processes $\{Q_{i,t}\}_{t=0}^T$ for company $i = 1, 2, \dots, I$ the price process $\bar{S} = \{\bar{S}_t\}_{t=0}^T$ is called equilibrium permit price process, if there exists $\{\bar{X}_{i,t}\}_{t=0}^{T-\Delta t}$ for $i = 1, 2, \dots, I$ such that for all $i = 1, 2, \dots, I$ and $t = 0, \dots, T - \Delta t$*

$$\mathbb{E}_{\mathbb{P}} \left[\Phi(-d_{i,t}) \cdot \Phi(-d_{I^-,t}^{\text{lag}}) | \mathcal{F}_t \right] = \mathbb{E}_{\mathbb{P}} \left[\Phi(-d_{j,t}) \cdot \Phi(-d_{I^-,t}^{\text{lag}}) | \mathcal{F}_t \right], \quad \mathcal{I} = I^- \cup i \quad (4.15)$$

and the market clearing condition is satisfied $\sum_{i=1}^I \bar{X}_{i,t} = 0$ for all $t = 0, \dots, T - \Delta t$.

It is remarkable to notice that, at each time step both the permits traded-quantity and the permit price, in equilibrium, are the result of the companies' continuous adjusting of emission portfolio allocations based on the accumulated pollution processes and the available information about net permit positions. In the next section we delve deeper into these

aspects by means of an extensive numerical exercise.

4.6 Numerical Evaluation

For illustrative purposes we consider a situation where $\mathcal{I} = 2$. Based on equation (4.13) and the market clearing condition (4.6), we simulate several paths of the emission permit price. In each simulation exercise the time period T is fixed at one year (i.e. 250 trading days), the weighted average cost of capital is set at 10% and the penalty, P , is equal to 40.

Starting at $t = 0$, and using equation (4.1), we simulate a pair of independent pollution processes: one for each company i -th, $i \in I$. Then, according to the initial permits amount $N_{i,0}$, we evaluate numerically the quantity of equilibrium permits such that equation (4.14) holds. We then calculate S_0^1 , the *implied* equilibrium permit price. This procedure is repeated n -times to evaluate the expected equilibrium permit price $\bar{S}_0 := \sum_{j=1}^n S_0^j/n$. At time $t = \Delta t$, the resulting net-permits positions $(\delta_{i,0}; i = 1, 2)$ are evaluated using \bar{S}_0 and a fixed pair of accumulated pollution volumes, randomly chosen among the n pairs of pollution simulations. Repeating n -times the procedure described above, we compute the expected equilibrium permit price $\bar{S}_{\Delta t}$. Reiterating this at each time step up to $T - \Delta t$ we obtain the simulated equilibrium permit price history depicted in the figures below.

Figures 4.2 and 4.3 illustrate the equilibrium permit price evolution stopped at three different time steps (50, 150 and 200 days) of the described procedure. In the first figure, we depict a situation where both companies' pollution processes have a positive quick-paced drift of 15% and 10% respectively, and a mild volatility level, set at 10% for both. While the second company has been equipped with an initial permit endowment close to its total expected emissions, i.e. $N_{2,0} \approx Q_{2,0} \cdot \int_0^T e^{\mu_2 s} ds$ where $T = 1$, the first company has been allocated an initial amount of permits slightly larger than its total expected emissions, i.e. $N_{1,0} > Q_{1,0} \cdot \int_0^T e^{\mu_1 s} ds$ where $T = 1$. As such, we would expect an upward-moving

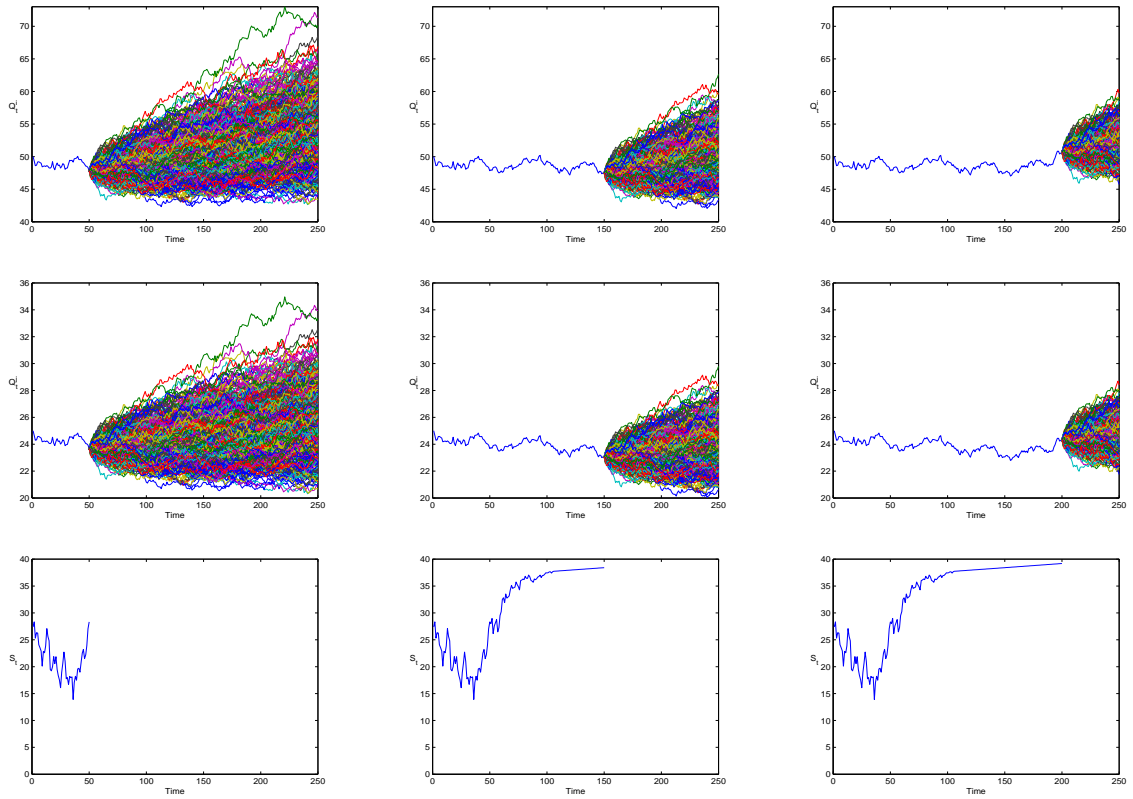


Figure 4.2: \overline{S}_t permit price evolution (bottom-part) for the pollution parameters $\mu = (0.15; 0.10)$, $\sigma = (0.10; 0.10)$, $Q_0 = (50; 25)$, $N_0 = (52; 25)$. The simulated pollution processes are depicted in the upper ($Q_{1,t}$) and middle-part ($Q_{2,t}$).

equilibrium permit price: the price evolution in the last row of figure 4.2 confirms this. Conversely, modifying the pollution drift terms and setting, respectively, a negative value for the first company, $\mu_1 = -0.15$, and a negligible drift term for the second one, $\mu_2 = 0.001$, we would expect a reverse effect, other things being equal. The last row of figure 4.3 supports the expectation of a downward-moving equilibrium permit price.

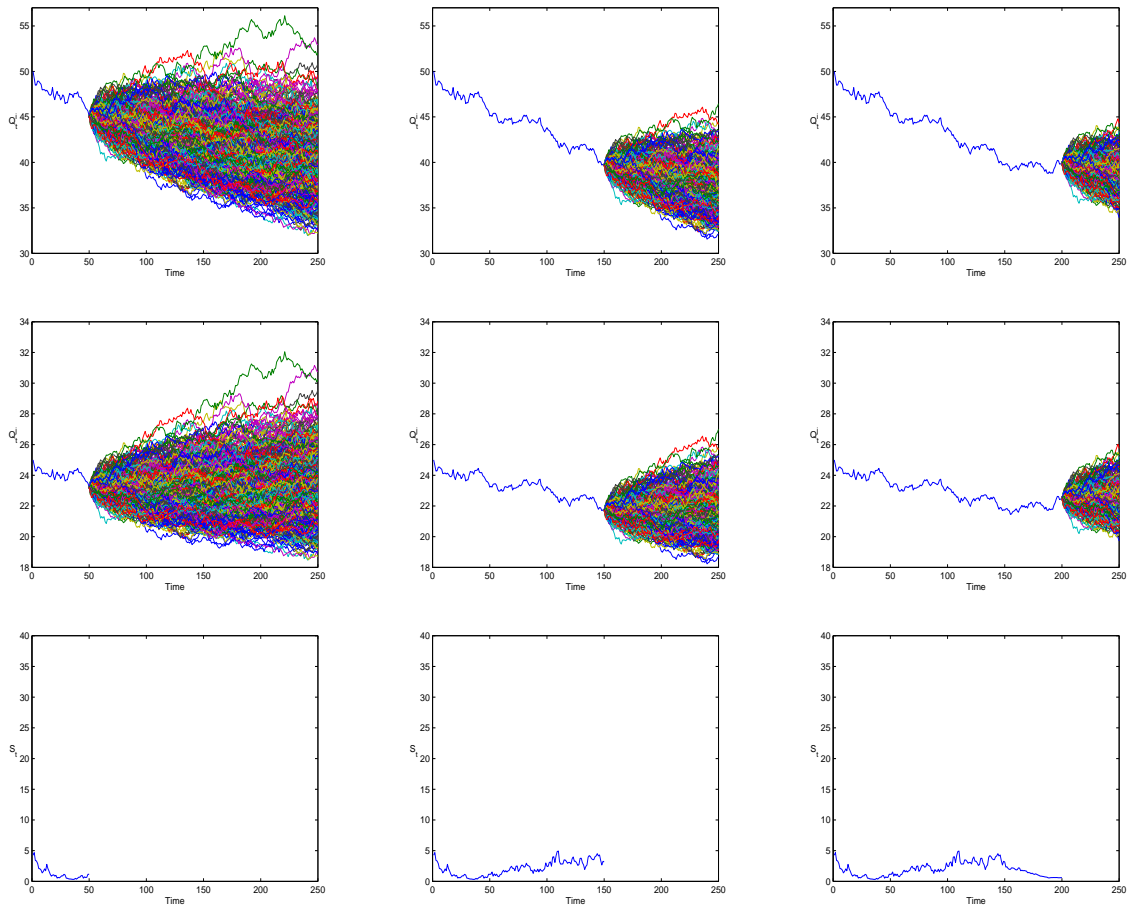


Figure 4.3: \bar{S}_t permit price evolution (bottom-part) for the pollution parameters $\boldsymbol{\mu} = (-0.15; 0.001)$, $\boldsymbol{\sigma} = (0.10; 0.10)$, $\mathbf{Q}_0 = (50; 25)$, $\mathbf{N}_0 = (52; 25)$. The simulated pollution processes are depicted in the upper ($Q_{1,t}$) and middle-part ($Q_{2,t}$).

Figure 4.4 depicts a brief sensitivity analysis of the equilibrium permit price with respect to the parameters of the companies' pollution processes. Starting from a set of conveniently chosen parameters, i.e. $\boldsymbol{\mu} = (0.25; 0.20)$, $\boldsymbol{\sigma} = (0.15; 0.40)$, $\mathbf{Q}_0 = (50; 25)$,

$\mathbf{N}_0 = (60; 40)$, we let the drift and volatility terms of company one vary, both in the first and in the second picture, keeping all the other parameters constant. As expected, the larger μ_1 , the higher the probability of being in shortage by the end of the period, i.e. T . This reasonably implies an upward trend in the permit price. However, in the particular simulated case, for each employed drift term except where $\mu_1 = 0.50$, as time moves forward and uncertainty is resolved, the initial permit endowments are sufficiently large to reverse such a trend (see upper part of figure 4.4).

Similarly, the larger σ_1 , the higher the uncertainty about $\int_0^T Q_{1,s} ds - \delta_{1,T-\Delta t}$, i.e. the net permit position before the compliance date, and consequently about the probability of a non-shortage situation in the future for both companies, i.e. $\mathbb{P}_t^i, t \in [0, T - \Delta t]$. As can be observed, higher volatility uncertainty is reflected in a higher permit (option) price. However, in our particular simulated example, while more information about the accumulated pollution volumes is collected, the initial permit amount value takes precedence over the overall uncertainty level. This, in turn leads to a price decrease (see middle part of figure 4.4). Finally, the impact of different pairs of initial permit endowments is observable in the last picture. The upper line depicts a clear shortage situation. After some trading time, the shortage status becomes a fact and the permit price is simply the discounted penalty level. The lower line depicts the opposite situation. Both companies have been allocated an amount of permits that is over-generous and the permit price hovers slightly above zero (see lower part of figure 4.4). It is extremely interesting to observe that the yellow price path very closely resembles the empirical spot permit price of CO₂ in the European market during 2005 and 2006. After a period of slow but continuous upward movement, due to purchasers being convinced of a shortage, the price plummeted by almost 70%, thereafter

drifting towards zero. This price reverse can be attributed to the disappearance of asymmetric information among market players in terms of their net permits positions.¹⁵ By the end of 2007, the emission permit spot price for phase I is almost nil; however it would have been zero only if the probability of an excess situation had been exactly one. This feature, along with the described price reaction to drift and volatility movement, is common to standard financial option contracts.

Optimal strategies are readily computable in a static and deterministic framework as those described in section 2.3. Conversely, regulatory uncertainties and uncertainties in the evolution of the pollution processes make an identification of the best strategy, which is less straightforward in the short-term. Apart from technological issues and physical constraints, financial concerns are also beginning to creep in. Anecdotal evidence of extreme volatility in the European and U.S. permit markets suggests an urgent need for the development of effective hedging techniques.¹⁶ In addition, the numerous risks related to market-based products highlight the importance of developing appropriate risk-management tools for those companies which are subject to environmental programs, as well as to specialized traders. More importantly, a valid price model is required for pricing any financial instruments or project whose value *derives* from the future CO₂ spot permit price. Extremely relevant examples are project-based investments (see next chapter), that at regular intervals, return emission reduction certificates, yielding a payoff that depends on the CO₂ permit market price. Other important examples are technological abatement investments or production process modifications that can be valued in terms of costs saved from purchasing emission

¹⁵Referring to the discussion in the subsection 3.1e and considering the model framework in Chapter 3, the price reverse may be a result of a larger (density) weight placed on the group (of firms) which expect the market being in extreme shortage with respect to the (density) weight of those expecting a market in excess of permits.

¹⁶Hedging strategies can be constructed by means of futures contracts or by introducing option instruments (the first option contract on CO₂ was traded in October, 2005 between the French electricity company EDF and the Amsterdam based company Statkraft). Futures are traded both over-the-counter and on several exchanges.

permits or revenue from the sales of extra unused permits.

4.7 CO₂ Option Pricing

In the EU ETS there are two major groups of players: that one which is covered by the regulations and the one that trades permits with the purpose to speculate.¹⁷ This last group consists of brokers, hedge funds, banks and insurers. Due to the nature of such institutions, it is clear these players have been attracted by the high volatility of the market of emission permits. Moreover, the emergence of international exchanges and the consequent enhancement of the liquidity of spot and futures contracts encouraged even more the participation of financial players in the EU ETS. At the time of writing, most of the trades concern futures traded on the London-based European Climate Exchange, the largest pan-European platform. Banks and insurance invest in emission permits largely through funds. As a matter of fact, in the last 2 years several funds emerged in the private sector, raising the apprehension for the existence of a speculative bubble. In sum, the financial sector is providing liquidity by trading futures contracts; it is sharpening the risk-sharing products by offering ad-hoc insurance contracts; it is serving the market by creating new service-solutions; and, more importantly, it is pushing for the creation of a solid market for option contracts where the underlying is the spot price of the emission permits.

The other group consists of all industrial companies which are covered by the EU ETS. Industry can act as seller or buyer in the market of emission permits like the financial sector. However, the main difference between these two sectors concerns the market perception. Financial companies look at emission permits as a new market (a new opportunity), while industry faces it as an entirely new set of regulations. To some extent, today the industrial

¹⁷The original version of this chapter is part of the paper “The Endogenous Price Dynamics of Emission Allowances: An Application to CO₂ Option Pricing” that was co-written with Marc Chesney.

sector can be split into two groups in respect to the employed strategy. More precisely, one can distinguish between “passive” and “active” strategy. In the “passive” strategy the firm only takes action at the beginning of the trading period (buying if it expects to be in need or selling if it expects to be in excess - but maintaining a sufficient reservoir), holds the rest of the permits on the registry account, and waits for all verified emission data. This strategy corresponds to the situation described in section 4.3 and it is not far from what has been experienced by some industries in Western Europe. In fact, producers of cement, ferrous metals, building materials and pulp and paper have generally limited ability to reduce their emissions but, having received fairly generous initial amount of permits in phase I, they acted passively on average. The “active” strategy can have several variations. In the simplest form, a company sells all permits once they appear on the registry account but immediately buys them back on the forward market. As such, permits are temporarily transferred into cash which is later on transferred back into permits at a known (forward) price. The net annual surplus or shortage of permits is then settled every year on the spot market. In a sophisticated variant of the “active” approach, a company could cover the buy-back leg of the transaction by insurance or a call option.¹⁸ Trivially if, at expiry of the option, the price of the permit is higher than the strike price of the option, the option is exercised. If, on the other hand the permit price is below the strike price of the option, the option is not exercised and the emission permit can be bought at the lower market price. Therefore, also active industrial traders are interested in the development of a large option market - especially call options.

Based on this discussion, it's clear why a CO₂ option market is slowly growing and attracting a wide variety of industrials, utilities and financial institutions of various nature.

¹⁸A more advanced variant of the “active” strategy corresponds to the dynamic problem described in section 4.4.

With regard to the different scopes listed above, the importance of such a market is two-fold. First, CO₂ option contracts satisfy the primary need of risk transfer from those who wish to reduce the risk of a permits shortage situation, namely the risk of financial exposure, to those willing to accept it. By allowing European covered companies to reduce their exposure to price risk, buyers and sellers can better plan their (active) trading strategies and their businesses. Second, writing option contracts financial institutions can take a position on the market.

It worth to mention that the development of a CO₂ option pricing approach is not limited to price standard financial contracts. Any project-based investment, i.e. investments committed under the so-called CDM and JI mechanisms, which at regular intervals returns CO₂ emission reduction certificates yielding a payoff that depends on the CO₂ permit market price, can be considered as (real) option contracts.¹⁹ It is natural to interpret such projects as contracts whose value *derives* from the future CO₂ spot permit price. Similarly, any technological abatement investment or production process modification can be valued in terms of saved costs from purchasing emission permits or revenue from the sales of extra unused permits. As mentioned in section 4.2, [18] used this argument in order to identify a plausible reason for the difference between the marginal cost of running abatement technologies such as scrubbers and the emission allowance price. They called this difference the option premium. This is the first paper that discovers the option-value implicitly embedded in the value of an emission permit. In line with this consideration, an option where the underlying is any sort of marketable permit is in fact a compound option. Finally, all industries which undertake any kind of technological abatement investments in order to free-up emission permits and sell them are interested in standard put options. Recalling the discussion in

¹⁹It has to be noted that each certificate has to succeed a completely regulated verification and certification procedure before being eligible as CER (for CDM) or as ERU (for JI).

section 4.2, physical abatement investments often entails a multi-year time horizon. Therefore, companies are generally quite reluctant to sell forward excess emission permits. Put options satisfy the need to lock in a certain price in the case the implemented abatement measure delivers the promised amount of permits.

In what follows we develop a CO₂ option pricing approach based on the equilibrium price dynamics obtained in the previous sections. A comparison with a benchmark model, i.e. the Black-Merton-Scholes formula, follows as an exercise.

4.8 CO₂ Option Pricing: A Comparison

In this brief last section we carry over a European-style (financial and real) option pricing model comparison. In particular, we attempt a comparison of CO₂ option-pricing models evaluating plain vanilla European options (Call and Put) with strike price K and maturity T . The benchmark model is that of Black, Merton and Scholes (1973) (called BMS), where one assumes the permit price evolves according to a geometric Brownian motion

$$\frac{d\tilde{S}_t}{\tilde{S}_t} = \alpha dt + \beta dW_t,$$

where α and β are the constant drift and constant volatility respectively. The second option approach relies on the endogenous equilibrium price procedure described in the previous section, thus the underlying price follows a dynamics like (A.9) in the Appendix A.5. European options are priced numerically by means of Monte Carlo simulations. However, in order to obtain a fair option pricing comparison under the same probability measure, we consider the risk adjusted version of (A.9) such that the discounted equilibrium price is a martingale, i.e. the dynamics (A.10).

To be consistent with the numerical results exposed in section 4.6, we maintain the time

period T fixed at one year, the discount factor is 10% and the penalty P is equal to €40. For illustrative purposes, we consider six possible general situations where pollution drift and volatility terms are varying but we make the initial pollution levels and the initial permits endowments fixed, i.e. $Q_0 = (50; 25)$ and $N_0 = (55; 25)$. For all models, the risk-free rate used is $r = 0.03$. Similarly, to obtain comparable and meaningful results, the parameters of the exogenous geometric Brownian motion used in the BMS option pricing model are estimated from the simulated prices obtained applying the equilibrium price procedure. More precisely, the starting permit price at time $t = 0$ is the mean of all first points of the simulated paths, i.e. $\tilde{S}_0 = \sum_{i=1}^A \bar{S}_0^i / A$ where A is the number of simulations. In addition, the constant volatility term $\hat{\beta}$ is the mean of A annualized historical volatilities estimated by the squared difference of the log-returns of each simulated path.

In the first situation, called "Positive–Positive", both companies are characterized by a mild positive drift term of 10% and a rather high volatility level of 20%. In the second situation, the first company has a positive drift term, $\mu_1 = 10\%$ and a relatively mild constant volatility, $\sigma_1 = 10\%$, whereas the second company has zero-drift and a volatility double compared to the previous, $\sigma_2 = 20\%$. With an identical volatility level of 15% and an equal drift term of 10%, but of opposite sign, we compute the option price in the situation labelled "Positive–Negative". In the fourth case, called "Stable–Stable", both pollution drifts are negligible and the volatility levels are 15% and 20% respectively. The "Stable–Negative" situation is characterized by identical volatility, 25%, negligible drift for the first company and a negative one, -10%, for the second company. Finally, in the last case, drift parameters are both -10% and volatilities are 20% and 15% respectively.

The option prices are consistent across all six described situations. The higher the pollution drift terms the higher the Call options values (sorted for decreasing strike price K); the lower the pollution drift terms, the higher the Put options values (sorted for increasing strike price K). Given the unboundedness of a geometric Brownian motion, one would

Parameters	Strike	K=10		K=20		K=30	
		Call	Put	Call	Put	Call	Put
$\tilde{S}_0, \hat{\beta}; [\mu, \sigma]$	Model						
$\tilde{S}_0 = 18, 46, \hat{\beta} = 0.41$ [(0.10, 0.10)(0.20, 0.20)]	PP BMS PP Endogenous	8.89 7.40	0.14 2.65	2.63 2.18	3.58 7.13	0.63 0.00	11.28 14.65
$\tilde{S}_0 = 11, 33, \hat{\beta} = 0.76$ [(0.10, 0.01)(0.10, 0.20)]	PS BMS PS Endogenous	4.00 2.46	2.37 4.92	1.54 0.01	9.61 12.17	0.69 0.00	18.46 21.86
$\tilde{S}_0 = 11, 23, \hat{\beta} = 0.69$ [(0.10, -0.10)(0.15, 0.15)]	PN BMS PN Endogenous	3.66 2.37	2.13 4.66	1.21 0.01	9.38 12.00	0.46 0.00	18.34 21.69
$\tilde{S}_0 = 8, 80, \hat{\beta} = 0.76$ [(0.01, 0.01)(0.15, 0.20)]	SS BMS SS Endogenous	2.34 1.06	3.24 5.43	0.76 0.00	11.36 14.07	0.30 0.00	20.61 23.77
$\tilde{S}_0 = 6, 52, \hat{\beta} = 0.79$ [(0.01, -0.01)(0.25, 0.25)]	SN BMS SN Endogenous	1.20 0.27	4.38 6.15	0.33 0.00	13.21 15.58	0.12 0.00	22.71 25.28
$\tilde{S}_0 = 6, 60, \hat{\beta} = 0.61$ [(-0.10, -0.10)(0.20, 0.15)]	NN BMS NN Endogenous	0.79 0.20	3.88 5.23	0.11 0.00	12.91 14.73	0.02 0.00	22.53 24.43

Table 4.1: European Call and Put option prices according to two different option pricing models. The table reports the results for six possible general situations where we allow the pollution drift and volatility terms to vary. Initial pollution levels, $\mathbf{Q}_0 = (50; 25)$, and initial permits endowments, $\mathbf{N}_0 = (55; 25)$, are fixed. The risk-free rate is $r = 0.03$, maturity is $T = 1$; and the penalty $P = 40$. The simulated pollution processes n are 500 for each company and $A = 100$.

correctly expect the Call prices valuated with BMS model to be higher compared to the second approach. This is the case in all six situations and for every strike price. Put prices valuated with BMS model are lower. Indeed, the possibility for the market to be in excess at maturity generates a non-negligible probability for the event $S_T = 0$. In other words, in the BMS model, contrary to our model, the probability to be in excess does not play a significant role. However, such a price-rank is not always the rule and it is intimately linked to the initial amount of permits allocated to each company. For instance, an emission permits

market over-allocation (N_0 is significantly larger than Q_0), would impact the value-range for \tilde{S}_0 shifting it down and implicitly reducing or deleting the option price difference. An emission permits market under-allocation (N_0 is extremely smaller than Q_0), instead could reverse the price-rank and make BMS Call options cheaper.

Those that attribute the existence of such a difference to a model specification (in fact the historical volatility is quite high), might suggest employing a more general exogenous dynamics to describe the CO₂ permit price.²⁰ However, this will not fill the price gap inasmuch as the emission permits market positions are not taken into account. In general, in a market for emission permits, the lower (higher) the initial allocation the more valuable would be a Call (Put) option.

²⁰The estimation of several popular continuous-time processes via maximum likelihood in [28] confirms that the jump-diffusion model proposed by [54] delivers the best CO₂ spot-returns description.

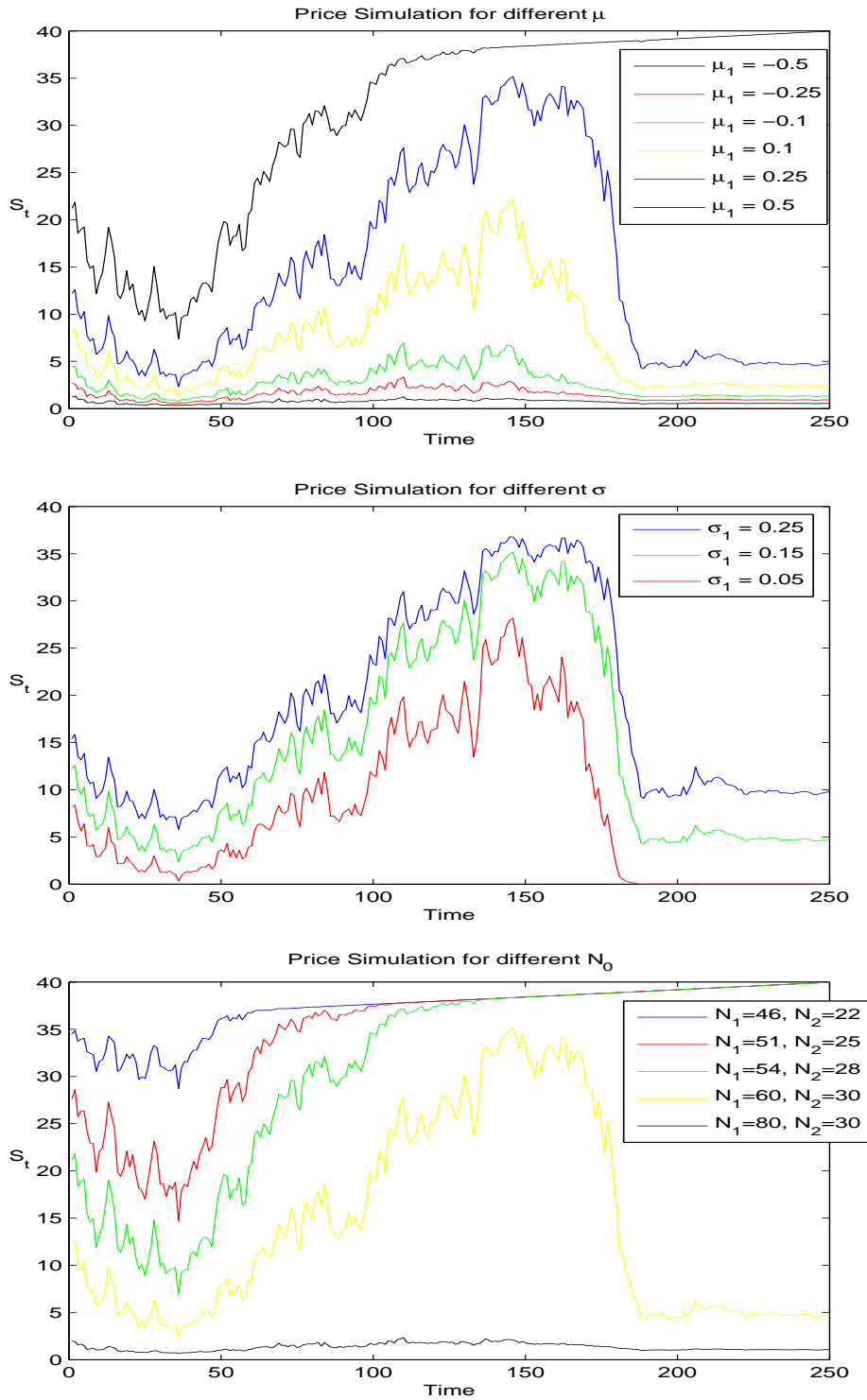


Figure 4.4: \bar{S}_t permit price evolution letting vary the drift and the volatility terms for company one, respectively upper and middle picture, and both the initial permits endowments, lower picture. The common starting pollution parameters are $\mu = [-0.15; 0.001]$, $\sigma = [0.10; 0.10]$, $Q_0 = [50; 25]$, $N_0 = [52; 25]$.

Appendix A

A.1 Appendix A.1

The following objective function has to be minimized with respect to X_0

$$H \equiv \left\{ S_0 \cdot X_0 + e^{-\eta T} \mathbb{E}_{\mathbb{P}} \left[\left(\frac{4}{\sigma^2} \cdot Q_0 \cdot A_{\sigma^2 T/4}^z - N_0 - X_0 \right)^+ \cdot P \right] \right\} \quad (\text{A.1})$$

and denoting $A_T^\nu = \int_0^T e^{2(W_s + \nu s)} ds$, $\nu := \frac{1}{\sigma} \cdot (\mu - \frac{\sigma^2}{2})$.

The law of A_t^z is $\mathbb{P}(A_t^z \in dx) = \varphi(t, x) dx$ where

$$\varphi(t, x) = x^{\nu-1} \frac{1}{(2\pi^3 t)^{1/2}} e^{\left(\frac{\pi^2}{2t} - \frac{1}{2x} - \frac{\nu^2 t}{2}\right)} \int_0^\infty y^\nu e^{-\frac{1}{2}xy^2} \Upsilon_y(t) dy,$$

$$\Upsilon_r(t) = \int_0^\infty e^{-\frac{y^2}{2t}} \cdot e^{-r(\cosh y)} \cdot \sinh(y) \cdot \sin\left(\frac{\pi y}{t}\right) dy.$$

Computing the first order condition (FOC) the following is obtained:

$$\bar{S}_0 = e^{-\eta T} \cdot P \cdot \mathbb{P} \left[A_{\sigma^2 T/4}^z > \frac{\delta_0 \cdot \sigma^2}{4Q_0} \right]$$

Therefore we can express the emission allowance price as a function of the penalty and

the probability of permit shortage:

$$\bar{S}_0 = e^{-\eta T} \cdot P \cdot \int_{\delta_0 \cdot \sigma^2 / 4 Q_0}^{\infty} \mathbb{P} \left[A_{\sigma^2 T / 4}^z \in dx \right].$$

For a simple analytical interpretation of the problem we can assume $T = \Delta t$, where Δt is a small time interval, and approximate the cumulative pollution process with its discrete representation:

$$\int_0^T Q_s ds = Q_0 e^{(\mu - \frac{\sigma^2}{2})\Delta t + \sigma W_{\Delta t}} \cdot \Delta t$$

Substituting in the objective function it follows:

$$\mathbf{H} \equiv \left\{ S_0 \cdot X_0 + e^{-\eta T} \mathbb{E}_{\mathbb{P}} \left[\left(Q_0 e^{(\mu - \frac{\sigma^2}{2})\Delta t + \sigma W_{\Delta t}} \cdot \Delta t - N_0 - X_0 \right)^+ \cdot P \right] \right\} \quad (\text{A.2})$$

Computing the FOC it follows:

$$\begin{aligned} \bar{S}_0 &= e^{-\eta T} \cdot P \cdot \mathbb{E}_{\mathbb{P}} \left[\mathbb{1}_{Q_0 e^{(\mu - \frac{\sigma^2}{2})\Delta t + \sigma W_{\Delta t}} \cdot \Delta t > N_0 + X_0} \right] \\ &= e^{-\eta T} \cdot P \cdot \mathbb{P} \left[Q_0 e^{(\mu - \frac{\sigma^2}{2})\Delta t + \sigma W_{\Delta t}} \cdot \Delta t > N_0 + X_0 \right], \end{aligned}$$

moving on from this, we express the price as a function of the penalty and the probability of permit shortage and the results of equation (4.5) are obtained.

A.2 Appendix A.2

The following objective function has to be minimized with respect to $X_{1,T-\Delta t}$:

$$H \equiv \left\{ S_{T-\Delta t} \cdot X_{1,T-\Delta t} + e^{-\eta \Delta t} \mathbb{E}_{\mathbb{P}} [S_T \cdot X_{1,T} | \mathcal{F}_{T-\Delta t}] \right\}.$$

Deriving the first order conditions, we arrive at equation (4.10). Introducing asymmetric information as explained in the section 4.4, we consider the discrete approximation for the pollution processes and obtain

$$\begin{aligned} \mathbb{P} \left(\int_0^T Q_{1,s} ds > \delta_{1,T-\Delta t} \right) &= \mathbb{P} \left(Q_{1,T-\Delta t} \cdot e^{\left(\mu_1 - \frac{\sigma_1^2}{2}\right) \cdot \Delta t + \sigma_1 W_{\Delta t}} \cdot \Delta t > \delta_{1,T-\Delta t} - \int_0^{T-\Delta t} Q_{1,s} ds \right) \\ &= \Phi(d_{1,T-\Delta t}), \end{aligned}$$

where $d_{1,T-\Delta t}$ is defined in section 4.4. Similarly

$$\mathbb{E}_{\mathbb{P}} \left[\mathbb{1}_{\delta_{1,T-\Delta t} > \int_0^T Q_{1,s} ds} | \mathcal{F}_{T-\Delta t} \right] = \mathbb{P} \left(\delta_{1,T-\Delta t} > \int_0^T Q_{1,s} ds \right) = \Phi(-d_{1,T-\Delta t}), \quad \text{and}$$

$$\mathbb{E}_{\mathbb{P}} \left[\mathbb{1}_{\int_0^T Q_{2,s} ds > \delta_{2,T-\Delta t}} | \mathcal{F}_{T-\Delta t} \right] = \mathbb{P} \left(\int_0^T Q_{2,s} ds > \delta_{2,T-\Delta t} \right) = \Phi(d_{2,T-\Delta t}^{\text{lag}}),$$

where $d_{2,T-\Delta t}^{\text{lag}}$ is defined in section 4.4. It follows that:

$$\begin{aligned} \bar{S}_{T-\Delta t} &= e^{-\eta \Delta t} \cdot P \cdot \left[\Phi(d_{1,T-\Delta t}) + \Phi(-d_{1,T-\Delta t}) \cdot \Phi(d_{2,T-\Delta t}^{\text{lag}}) \right] \\ &= e^{-\eta \Delta t} \cdot P \cdot \left[1 - \Phi(-d_{1,T-\Delta t}) \cdot \Phi(-d_{2,T-\Delta t}^{\text{lag}}) \right] \end{aligned}$$

The same computation holds for equation (4.12).

A.3 Appendix A.3

The following objective function has to be minimized with respect to $X_{1,T-2\Delta t}$:

$$H \equiv \left\{ S_{T-2\Delta t} \cdot X_{1,T-2\Delta t} + e^{-\eta \Delta t} \mathbb{E}_{\mathbb{P}} \left[\bar{S}_{T-\Delta t} \cdot \bar{X}_{1,T-\Delta t} + e^{-\eta \Delta t} \cdot S_T \cdot X_{1,T} | \mathcal{F}_{T-2\Delta t} \right] \right\}.$$

Computing the FOC, the following is obtained:

$$0 = S_{T-2\Delta t} + e^{-\eta\Delta t} \mathbb{E}_{\mathbb{P}} \left[\bar{S}_{T-\Delta t} \cdot \frac{\partial \bar{X}_{1,T-\Delta t}}{\partial X_{1,T-2\Delta t}} + \bar{X}_{1,T-\Delta t} \cdot \frac{\partial \bar{S}_{T-\Delta t}}{\partial X_{1,T-2\Delta t}} \right]$$

because by equation (4.7), $S_T = \{0, P\}$, hence $X_{1,T} \cdot \frac{\partial S_T}{\partial X_{1,T-2\Delta t}} = 0$.

Moreover, considering the existence of a lag-effect due to the presence of asymmetric information and assuming that

$$\frac{\partial \bar{X}_{1,T-(j-1)\Delta t}}{\partial X_{1,T-j\Delta t}} = -1, \quad \frac{\partial \bar{X}_{1,T-(j-k)\Delta t}}{\partial X_{1,T-j\Delta t}} = 0 \quad \text{where } k \in [2, j] \quad k \in \mathbb{N}^+, \quad (\text{A.3})$$

it follows

$$\frac{\partial \bar{X}_{1,T}}{\partial X_{1,T-2\Delta t}} = 0.$$

The previous assumptions are introduced for the sake of tractability of the model. A rigorous mathematical approach requires the introduction of backward-forward stochastic differential equations (BFSDEs) in order to model the decision problem. In fact, it is not sufficient to solve a stochastic dynamic programming problem since at each time-step ($T - j\Delta t$) the control variable (the quantity of permits to buy or to sell) is a function of the previous quantity of permits traded ($(T - (j + h)\Delta t)$, where $h \in [1, T/\Delta t - j] \quad h \in \mathbb{N}^+$) and of the future quantity of permits that will be traded ($(T - (j - k)\Delta t)$, where $k \in [1, j] \quad k \in \mathbb{N}^+$).

Let us define:

$$a_1 = \left(N_{1,T-2\Delta t} + \bar{X}_{1,T-\Delta t} - \int_0^{T-\Delta t} Q_{1,s} ds \right),$$

$$b_2^{\text{lag}} = \left(N_{2,T-2\Delta t} + \bar{X}_{2,T-\Delta t} - \int_0^{T-2\Delta t} Q_{2,s} ds \right),$$

since $X_{1,s} = -X_{2,s} \quad \forall \quad s \in [0, T - 1]$.

Recalling equation (4.11), we expand $\partial \bar{S}_{T-\Delta t} / \partial X_{1,T-2\Delta t}$:

$$\begin{aligned} \frac{\partial \bar{S}_{T-\Delta t}}{\partial X_{1,T-2\Delta t}} &= \frac{\partial}{\partial X_{1,T-2\Delta t}} \left[e^{-\eta \Delta t} \cdot P \left[1 - \Phi(-d_{1,T-\Delta t}) \cdot \Phi(-d_{2,T-\Delta t}^{\text{lag}}) \right] \right] \quad (\text{A.4}) \\ &= e^{-\eta \Delta t} \cdot P \cdot \phi(-d_{1,T-\Delta t}) \cdot \frac{\partial d_{1,T-\Delta t}}{\partial X_{1,T-2\Delta t}} \cdot \Phi(-d_{2,T-\Delta t}^{\text{lag}}) \\ &\quad + e^{-\eta \Delta t} \cdot P \cdot \Phi(-d_{1,T-\Delta t}) \cdot \phi(-d_{2,T-\Delta t}^{\text{lag}}) \cdot \frac{\partial d_{2,T-\Delta t}^{\text{lag}}}{\partial X_{1,T-2\Delta t}}. \end{aligned}$$

Using conditions (A.3), the following equations are obtained:

$$\frac{\partial d_{1,T-\Delta t}}{\partial X_{1,T-2\Delta t}} = \frac{1}{\sigma_1 \sqrt{\Delta t}} \cdot \frac{-1}{(Q_{1,T-\Delta t} \cdot \Delta t) / a_1} \cdot (Q_{1,T-\Delta t} \cdot \Delta t) \cdot (a_1)^{-2} \cdot \frac{\partial a_1}{\partial X_{1,T-2\Delta t}} = 0,$$

$$\frac{\partial d_{2,T-\Delta t}^{\text{lag}}}{\partial X_{1,T-2\Delta t}} = \frac{1}{\sigma_2 \sqrt{2\Delta t}} \cdot \frac{-1}{(Q_{2,T-2\Delta t} \cdot 2\Delta t) / b_2^{\text{lag}}} \cdot (Q_{2,T-2\Delta t} \cdot 2\Delta t) \cdot (b_2^{\text{lag}})^{-2} \cdot \frac{\partial b_2^{\text{lag}}}{\partial X_{1,T-2\Delta t}} = 0;$$

and hence $\frac{\partial \bar{S}_{T-\Delta t}}{\partial X_{1,T-2\Delta t}} = 0$.

Combining the previous result and conditions (A.3), the spot price of the emission allowances at time $T - 2\Delta t$ is:

$$\begin{aligned} \bar{S}_{T-2\Delta t} &= e^{-\eta \Delta t} \cdot \mathbb{E}_{\mathbb{P}} \left[\bar{S}_{T-\Delta t} | \mathcal{F}_{T-2\Delta t} \right] \\ &= e^{-\eta 2\Delta t} \cdot P \cdot \left\{ 1 - \mathbb{E}_{\mathbb{P}} \left[\Phi(-d_{1,T-\Delta t}) \cdot \Phi(-d_{2,T-\Delta t}^{\text{lag}}) | \mathcal{F}_{T-2\Delta t} \right] \right\}. \quad (\text{A.5}) \end{aligned}$$

Similarly, solving the minimization problem corresponding to company $i = 2$, it follows:

$$\begin{aligned} \bar{S}_{T-2\Delta t} &= e^{-\eta \Delta t} \cdot \mathbb{E}_{\mathbb{P}} \left[\bar{S}_{T-\Delta t} | \mathcal{F}_{T-2\Delta t} \right] \\ &= e^{-\eta 2\Delta t} \cdot P \cdot \left\{ 1 - \mathbb{E}_{\mathbb{P}} \left[\Phi(-d_{2,T-\Delta t}) \cdot \Phi(-d_{1,T-\Delta t}^{\text{lag}}) | \mathcal{F}_{T-2\Delta t} \right] \right\}. \quad (\text{A.6}) \end{aligned}$$

We generalize the proof for the time step $T - j\Delta t$ considering the following objective function that has to be minimized with respect to $X_{1,T-j\Delta t}$:

$$H \equiv \left\{ S_{T-j\Delta t} \cdot X_{1,T-j\Delta t} + e^{-\eta\Delta t} \mathbb{E}_{\mathbb{P}} \left[\sum_{h=1}^j e^{-\eta(h-1)\Delta t} \bar{S}_{T-(j-h)\Delta t} \cdot \bar{X}_{1,T-(j-k)\Delta t} \middle| \mathcal{F}_{T-j\Delta t} \right] \right\},$$

Computing the FOC, it follows:

$$\begin{aligned} S_{T-j\Delta t} \cdot \frac{\partial \bar{X}_{1,T-j\Delta t}}{\partial X_{1,T-j\Delta t}} = \\ -e^{-\eta\Delta t} \mathbb{E}_{\mathbb{P}} \left[\sum_{h=1}^j e^{-\eta(h-1)\Delta t} \bar{S}_{T-(j-h)\Delta t} \cdot \frac{\partial \bar{X}_{1,T-(j-h)\Delta t}}{\partial X_{1,T-j\Delta t}} + \bar{X}_{1,T-(j-h)\Delta t} \cdot \frac{\partial \bar{S}_{T-(j-h)\Delta t}}{\partial X_{1,T-j\Delta t}} \middle| \mathcal{F}_{T-j\Delta t} \right] \end{aligned}$$

After some computation and using conditions (A.3) and equation (A.4), the following equation is obtained:

$$\bar{S}_{T-j\Delta t} = e^{-\eta\Delta t} \mathbb{E}_{\mathbb{P}} \left[\bar{S}_{T-(j-1)\Delta t} \middle| \mathcal{F}_{T-j\Delta t} \right],$$

hence

$$\begin{aligned} \bar{S}_{T-j\Delta t} &= e^{-\eta\Delta t} \mathbb{E}_{\mathbb{P}} \left[e^{-\eta(j-1)\Delta t} \cdot P \cdot \left\{ 1 - \mathbb{E}_{\mathbb{P}} \left[\Phi(-d_{1,T-\Delta t}) \cdot \Phi(-d_{2,T-\Delta t}^{\text{lag}}) \middle| \mathcal{F}_{T-(j-1)\Delta t} \right] \right\} \middle| \mathcal{F}_{T-j\Delta t} \right] \\ &= e^{-\eta j\Delta t} \cdot P \cdot \left\{ 1 - \mathbb{E}_{\mathbb{P}} \left[\Phi(-d_{1,T-\Delta t}) \cdot \Phi(-d_{2,T-\Delta t}^{\text{lag}}) \middle| \mathcal{F}_{T-j\Delta t} \right] \right\}. \end{aligned}$$

A.4 Appendix A.4

Let us define $\mathcal{S} = \{1, 2, \dots, I\}$ the set of relevant companies. The existence of asymmetric information is modeled assuming that each company i observes its accumulated

pollution process and the accumulated (and aggregated) pollution process of the I^- companies with a lag, where $I^- := \mathcal{I} - i$. Modeling the emission permit price in a multi-period and multi-firm framework requires solving I minimization problems at each time step $k \in [1, 2, \dots, T/\Delta t]$. Along the line of [12], one can approximate the cumulative pollution process, $Q_{I^-,t} = \sum_{j=1, j \neq i}^I Q_{j,t}$, with a new geometric Brownian motion and obtain I emission price equations as described in section 4.4:

$$\bar{S}_{T-k\Delta t} = e^{-\eta k \Delta t} \cdot P \cdot \left\{ 1 - \mathbb{E}_{\mathbb{P}} \left[\Phi(-d_{i,T-\Delta t}) \cdot \Phi(-d_{I^-,T-\Delta t}^{\text{lag}}) \mid \mathcal{F}_{T-k\Delta t} \right] \right\},$$

where

$$d_{i,T-\Delta t} = \frac{\ln \left(\frac{Q_{i,T-\Delta t} \cdot \Delta t}{N_{i,T-2\Delta t} + X_{i,T-\Delta t} - \int_0^{T-\Delta t} Q_{i,s} ds} \right) + \left(\mu_i - \frac{\sigma_i^2}{2} \right) \cdot \Delta t}{\sigma_i \cdot \sqrt{\Delta t}},$$

and

$$d_{I^-,T-\Delta t}^{\text{lag}} = \frac{\ln \left(\frac{Q_{I^-,T-2\Delta t} \cdot 2\Delta t}{N_{I^-,T-2\Delta t} + X_{I^-,T-\Delta t} - \int_0^{T-2\Delta t} Q_{I^-,s} ds} \right) + \left(\mu_{I^-} - \frac{\sigma_{I^-}^2}{2} \right) \cdot 2\Delta t}{\sigma_{I^-} \cdot \sqrt{2\Delta t}}.$$

Using constant drift and volatility terms, $\{\mu, \sigma\} \in \mathbb{R}^I$, and relying on the standard technique of the methods of moments, we can determine the parameters of the new approximated geometric Brownian motion $Q_{I^-,t}$,

$$\frac{dQ_{I^-,t}}{Q_{I^-,t}} = \mu_{I^-} dt + \sigma_{I^-} dW_{I^-,t}$$

where W_{I^-} is a Brownian motion and

$$\mu_{I^-} = \frac{1}{t} \ln \left(\frac{\sum_{j=1, j \neq i}^I Q_{j,0} e^{\mu_j t}}{\sum_{j=1, j \neq i}^I Q_{j,0}} \right), \quad \sigma_{I^-}^2 = \frac{1}{t} \ln \left(\frac{\sum_{j,k=1, j,k \neq i}^I Q_{k,0} Q_{j,0} e^{(\mu_k + \mu_j + \rho_{k,j} \sigma_k \sigma_j) t}}{\left(\sum_{j=1}^I Q_{j,0} e^{\mu_j t} \right)^2} \right).$$

Hence, we determine the equilibrium permit price solving a system of I equations. More precisely, we numerically evaluate the quantity of permits that satisfies the following $I - 1$ equalities at each time step $k \in [1, 2, \dots, T/\Delta t]$:

$$\mathbb{E}_{\mathbb{P}} \left[\Phi(-d_{i,T-\Delta t}) \cdot \Phi(-d_{I^-,T-\Delta t}^{\text{lag}}) | \mathcal{F}_{T-k\Delta t} \right] = \mathbb{E}_{\mathbb{P}} \left[\Phi(-d_{j,T-\Delta t}) \cdot \Phi(-d_{I^-,T-\Delta t}^{\text{lag}}) | \mathcal{F}_{T-k\Delta t} \right], \quad (\text{A.7})$$

(for $\{i, j\} \in \mathcal{I}$ and $i \neq j$) and the market clearing condition $\sum_{i=1}^I \bar{X}_{i,T-k\Delta t} = 0$, for a given set of parameters $(\{\mu, \sigma, \mathbf{Q}_0, \mathbf{N}_0\} \in \mathbb{R}^I)$ that characterize the I pollution processes.

A.5 Appendix A.5

An interesting finding is the comparison of the resulting permit price dynamics with a standard geometric Brownian motion. Starting from equation (4.13) (from the point of view of company 1), we introduce the following more convenient notation $H =: \Phi(-d_{1,T-\Delta t}) \cdot \Phi(-d_{2,T-\Delta t}^{\text{lag}})$.

$$\begin{aligned} \frac{\Delta S_{T-(k+1)\Delta t}}{S_{T-(k+1)\Delta t}} &\approx \frac{1 - \mathbb{E}_{\mathbb{P}} [H | \mathcal{F}_{T-k\Delta t}] - (1 - \eta\Delta t) \cdot \{1 - \mathbb{E}_{\mathbb{P}} [H | \mathcal{F}_{T-(k+1)\Delta t}]\}}{(1 - \eta\Delta t) \cdot \{1 - \mathbb{E}_{\mathbb{P}} [H | \mathcal{F}_{T-(k+1)\Delta t}]\}} \\ &\approx \left\{ \eta\Delta t + \frac{\Delta \{1 - \mathbb{E}_{\mathbb{P}} [H | \mathcal{F}_{T-(k+1)\Delta t}]\}}{1 - \mathbb{E}_{\mathbb{P}} [H | \mathcal{F}_{T-(k+1)\Delta t}]} \right\} (1 - \eta\Delta t)^{-1} \\ &\approx \left\{ \eta\Delta t + \frac{\Delta \mathbb{P}_{T-(k+1)\Delta t}}{\mathbb{P}_{T-(k+1)\Delta t}} \right\}, \end{aligned} \quad (\text{A.8})$$

where Δt is infinitesimally small and \mathbb{P}_t here is the probability of shortage at time t from the point of view of company 1. Rewriting equation (A.8) for a general time instant t , it follows:

$$\frac{\Delta S_t}{S_t} \approx \frac{\Delta \mathbb{P}_t}{\mathbb{P}_t} + \eta \Delta t. \quad (\text{A.9})$$

Equation (A.9) describes the endogenous dynamics of the emission permit price from the point of view of company 1. A similar equation can be obtained from the point of view of company two. We determine the equilibrium permit amount equating these two quantities (along the lines of section 4.4) and then obtain the endogenous permit price. Not surprisingly, an increase in the probability of shortage ($\frac{\Delta \mathbb{P}_t}{\mathbb{P}_t} > 0$) would lead to an increase in the price of the emission permits. The corresponding risk-neutral dynamic is simply:

$$\frac{\Delta S_t}{S_t} \approx r \Delta t + \left(\frac{\Delta \mathbb{P}_t}{\mathbb{P}_t} - \mathbb{E}_{\mathbb{P}} \left[\frac{\Delta \mathbb{P}_t}{\mathbb{P}_t} \right] \right). \quad (\text{A.10})$$

Bibliography

- [1] C. Alexander and E. Lazar. Normal mixture GARCH(1,1): Applications to exchange rate modelling. *Journal of Applied Econometrics*, 21:307–336, 2006.
- [2] S. Atkinson and T. H. Tietenberg. Market failure in incentive-based regulation: The case of emission trading. *Journal of Environmental Economics and Management*, 21:17–31, 1991.
- [3] O. Bahn, B. Büler, S. Kypreos, and H. J. Lüthi. Modelling an international market of CO₂ emission permits. *International Journal of Global Energy Issues*, 12:283–291, 1997.
- [4] W. J. Baumol and W. E. Oates. The use of standards and prices for protection of the environment. *Swedish Journal of Economics*, 73(1):42–54, 1971.
- [5] W. J. Baumol and W. E. Oates. *The Theory of Environmental Policy*. Cambridge University Press, Cambridge, 1988.
- [6] E. Benz and S. Trück. Modeling the price dynamics of CO₂ emission allowances. Working paper, University of Bonn, 2006.
- [7] F. Black and M. S. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 7:637–654, 1973.

-
- [8] D. R. Bohi and D. Burtraw. Utility investment behavior and the emission trading market. *Resource and Energy Economics*, 14:129–153, 1992.
- [9] C. Böringer and A. Lange. On the design of optimal grandfathering schemes for emission allowances. *European Economics Review*, 49:2041–2055, 2005.
- [10] V. Bosetti, C. Carraro, and E. Massetti. International energy R&D spillovers and the economics of greenhouse gas atmospheric stabilization. Working Papers from University of Venice "Ca' Foscari", Department of Economics, 2007.
- [11] K. Brännäs and N. Nordman. Conditional skewness modelling for stock returns. *Applied Economics Letters*, 10:725–728, 2003.
- [12] D. Brigo, F. Mercurio, F. Rapisarda, and R. Scotti. Approximated moment-matching dynamics for basket-options simulation. *Quantitative Finance*, 4:1–16, 2004.
- [13] D. Bunn and C. Fezzi. Interaction of European carbon trading and energy prices. Fondazione Eni Enrico Mattei - Working Paper 123, 2007.
- [14] K. P. Burnham and D. Anderson. *Model Selection and Multi-Model Inference*. second edition, 2003.
- [15] D. Burtraw, K. Palmer, B. Bharvirkar, and A. Paul. The effect of allowance allocation on the cost of carbon emission trading. Working paper, Resources for the Future, Washington D.C., 2001.
- [16] T. J. Cason. Seller incentive properties of EPA's emission trading auction. *Journal of Environmental Economics and Management*, 25:177–195, 1993.
- [17] T. J. Cason. An experimental investigation of the seller incentive in EPA's emission trading auction. *American Economic Review*, 85:905–922, 1995.

-
- [18] H. Chao and R. Wilson. Option value of emission allowances. *Journal of Regulatory Economics*, 5:233–249, 1993.
- [19] M. Chesney and L. Taschini. The endogenous price dynamics of emission allowances: An application to CO₂ option pricing. Swiss Banking Institute, University of Zurich, Switzerland, 2008.
- [20] A. Christiansen, A. Arvanitakis, K. Tangen, and H. Hasselknippe. Price determinants in the EU emissions trading scheme. *Climate Policy*, 5:15–30, 2005.
- [21] R.H. Coase. The problem of social cost. *Journal of Law and Economics*, 1960.
- [22] J. S. Coggins and J. R. Swinton. The price of pollution: A dual approach to valuing SO₂ allowances. *Journal of Environmental Economics and Management*, 30:58–72, 1996.
- [23] F. J. Convery and L. Redmond. Market and price developments in the European Union Emissions Trading Scheme. *Review of Environmental Economics and Policy*, 1:88–111, 2007.
- [24] P. Cramton and S. Kerr. Tradable carbon permits auctions: How and why to auction not grandfather. *Energy Policy*, 30:333–345, 2002.
- [25] T.D. Crocker. *The Structuring of Atmospheric Pollution Control Systems*, volume 1 of *The Economics of Air Pollution*. Harold Wolozin, New York, 1966.
- [26] M.B. Cronshaw and J.B. Kruse. Regulated firms in pollution permit markets with banking. *Journal of Regulatory Economics*, 9:179–189, 1996.
- [27] J.H. Dales. *Pollution Property and Prices*. University of Toronto Press, Toronto, 1968.

-
- [28] G. Daskalakis, D. Psychoyios, and R. N. Markellos. Modeling CO₂ emission allowance prices and derivatives: Evidence from the european markets. Working paper, Athens University of Economics and Business, 2007.
- [29] C. de Jong and K. Walet. *Compliance strategies in the US Acid Rain program*, pages 201–217. A Guide to Emissions Trading. Risk Books, London, 2004.
- [30] J. A. Doucet and T. Strauss. On the bundling of coal and sulfur dioxide emissions allowances. *Energy Policy*, 22:764–770, 1994.
- [31] A. D. Ellerman, P. L. Joskow, and D. Harrison. Emissions trading in the u.s.: Experience, lessons and considerations for greenhouse gases. Technical report, Pew Center for Climate Change, 2003.
- [32] A.D. Ellerman. Lessons from phase 2 compliance with the us acid rain program. Technical document, Center for Energy and Environmental Policy Research CEEPR - MIT, 2003.
- [33] Y. H. Farzin and P. M. Kort. Pollution abatement investment when environmental regulation is uncertain. *Journal of Public Economic Theory*, 2:183–212, 2000.
- [34] M. Fehr and J. Hinz. A quantitative approach to carbon price risk modeling. Institute of Operations Research, ETH Zurich., 2006.
- [35] H. Geman and M. Yor. Bessel processes, asian options, and perpetuities. *Mathematical Finance*, 3:349–375, 1993.
- [36] P. Giot and S. Laurent. Modelling daily Value-at-Risk using realized volatility and arch type models. *Journal of Empirical Finance*, 11:379–398, 2004.

-
- [37] M. Haas, S. Mittnik, and M. S. Paolella. Mixed normal conditional heteroskedasticity. *Journal of Financial Econometrics*, 2(2):211–250, 2004.
- [38] R. W. Hahn. Market power and tradable property rights. *Quarterly Journal of Economics*, 99:753–765, 1983.
- [39] R. W. Hahn and G. L. Herster. Where did all the markets go? *Yale Journal of Regulations*, 6:109–153, 1989.
- [40] J. D. Hamilton. A quasi-bayesian approach to estimating parameters for mixtures of normal distributions. *Journal of Business and Economic Statistics*, 9(1):21–39, 1991.
- [41] B. E. Hansen. Autoregressive conditional density estimation. *International Economic Review*, 35:705–730, 1994.
- [42] C. R. Harvey and A. Siddique. Autoregressive conditional skewness. *Journal of Financial and Quantitative Analysis*, 34(4):465–487, 1999.
- [43] R. Hertwig, G. Barron, E. U. Weber, and I. Erev. Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, 15:534–539, 2004.
- [44] I. Hidalgo, L. Szabö, J. C. Ciscar, and A. Soria. Technological prospects and CO₂ emission trading analysis in the iron and steel industry: A global model. *Energy*, 30:583–610, 2005.
- [45] M. C. Insley. On the option to invest in pollution control under a regime of tradable emission allowances. *Canadian Journal of Economics*, 35:860–883, 2003.
- [46] C. Jung, K. Krutilla, and R. Boyd. Incentives for advanced pollution abatement technology at the industry level: An evaluation of policy alternatives. *Journal of Environmental Economics and Management*, 30(1):95–111, 1996.

-
- [47] M. Kamien, N. Schwartz, and F. T. Dolbear. Asymmetry between bribes and charges. *Water Resources Research*, 2(1):147–157, 1966.
- [48] A. G. Keeler. Non compliant firms in transferable discharge permit market: Some extensions. *Journal of Environmental Economics and Management*, 21:180–189, 1991.
- [49] M. Kirchler and J. Huber. Fat tails and volatility clustering in experimental asset markets. *Journal of Economic Dynamics & Control*, 31:1844–1874, 2007.
- [50] K. Kuester, S. Mittik, and M. S. Paolella. Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, 4(1):53–89, 2005.
- [51] T. Lux and M. Marchesi. Volatility clustering in financial markets: A microsimulation of interacting agents. *International Journal of Theoretical and Applied Finance*, 3:675–702, 2000.
- [52] D. A. Malueg. Welfare consequences of emission credit trading programs. *Journal of Environmental Economics and Management*, 18:66–77, 1990.
- [53] R. C. Merton. The theory of rational option pricing. *Bell Journal*, 4:141–183, 1973.
- [54] R. C. Merton. Option pricing when the underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.
- [55] W. S. Misolek and H. W. Elder. Exclusionary manipulation of markets for pollution rights. *Journal of Environmental Economics and Management*, 16:156–166, 1989.
- [56] S. Mittnik and M. S. Paolella. Conditional density and value-at-risk prediction of asian currency exchange rates. *Journal of Forecasting*, 19:313–333, 2000.

-
- [57] S. Mittnik and M. S. Paolella. Prediction of financial downside risk with heavy tailed conditional distributions. In S. T. Rachev, editor, *Handbook of Heavy Tailed Distributions in Finance*. Elsevier Science, Amsterdam, 2003.
- [58] S. Mittnik, Marc S. Paolella, and S. T. Rachev. Stationarity of stable power–GARCH processes. *Journal of Econometrics*, 106:97–107, 2002.
- [59] J. P. Montero. Marketable pollution permits with uncertainty and transaction costs. *Resource and Energy Economics*, 20:27–50, 1998.
- [60] W.D. Montgomery. Markets in licenses and efficient pollution control programs. *Journal of Economic Theory*, 5, 1972.
- [61] M. S. Paolella. Testing the stable paretian assumption. *Mathematical and Computer Modelling*, 34:1095–1112, 2001.
- [62] M. S. Paolella. *Intermediate Probability: A Computational Approach*. Chichester, 2007.
- [63] M. S. Paolella and S. C. Steude. Risk prediction: A DWARF-like approach. Swiss Banking Institute of the University of Zurich, 2007.
- [64] M. S. Paolella and L. Taschini. An econometric analysis of emission–allowances prices. *Journal of Banking and Finance – forthcoming*, 2008.
- [65] A.C. Pigou. *The Economics of Welfare*. Macmillan Press, UK, 1918.
- [66] S. T. Rachev and S. Mittnik. *Stable Paretian Models in Finance*. Chichester, 2000.
- [67] M. Roberts and M. Spence. Effluent charges and licenses under uncertainty. *Journal of Public Economic*, 5(3):193–208, 1976.

- [68] M. Rockinger and E. Jondeau. Entropy densities with an application to autoregressive conditional skewness and kurtosis. *Journal of Econometrics*, 106:119–142, 2002.
- [69] J.D. Rubin. A model of intertemporal emission trading, banking, and borrowing. *Journal of Environmental Economics and Management*, 31:269–286, 1996.
- [70] S. M. Schennach. 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, 2000.
- [71] J. Seifert, M. Uhrig-Homburg, and M. Wagner. Dynamic behavior of CO₂ spot prices: Theory and empirical evidence. Chair of Financial Engineering and Derivatives, University of Karlsruhe, 2006.
- [72] R. N. Stavins. Transaction costs and the performance of markets for pollution control. Discussion Paper 93-02, Kennedy School of Government, Harvard University, 1993.
- [73] R. N. Stavins. Transaction costs and tradable permits. *Journal of Environmental Economics and Management*, 29:133–148, 1995.
- [74] R. N. Stavins. Correlated uncertainty and policy instrument choice. *Journal of Environmental Economics and Management*, 30:218–232, 1996.
- [75] K. J. Sunnevag. Auction design for the allocation of emission permits in the presence of market power. *Journal Environmental and Resource Economics*, 26:385–400, 2003.
- [76] L. Szabö, I. Hidalgo, J. C. Ciscar, and A. Soria. CO₂ emission trading within the European Union and Annex B countries: the cement industry case. *Energy Policy*, 34:72–87, 2006.

- [77] T. H. Tietenberg. Emission trading: An exercise in reforming pollution policy. Working paper, Resources for the Future, Washington D.C., 1985.
- [78] T. H. Tietenberg. Economic instruments for environmental regulation. *Oxford Review of Economic Policy*, 6(1):17–33, 1990.
- [79] M. L. Weitzman. Prices vs. quantities. *Review of Economic Studies*, 41(4):683–691, 1974.
- [80] J. Zhao. Irreversible abatement investment under cost uncertainties: Tradable emission permits and emissions charges. *Journal of Public Economics*, 87:2765–2789, 2003.