Valuation of Carbon Emission Allowances and Sustainable Investment

by

Mingyu Fang

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Actuarial Science

Waterloo, Ontario, Canada, 2019

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Ilias Tsiakas Professor Department of Economics and Finance University of Guelph
Supervisors	Ken Seng Tan Professor Department of Statistics and Actuarial Science University of Waterloo
	Tony Wirjanto Professor Department of Statistics and Actuarial Science University of Waterloo
Internal Member	Ming Bin Feng Assistant Professor Department of Statistics and Actuarial Science University of Waterloo
	David Saunders Associate Professor Department of Statistics and Actuarial Science University of Waterloo
Internal-external Member	Olaf Weber Professor School of Environment, Enterprise and Development University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Rising awareness of the impacts of climate change is leading to a rapid development of emission trading schemes (ETS) globally as a market-based means of emission control. Under a typical ETS, emission allowances are issued to firms under predetermined quotas at the beginning of each year by regulators, a sufficient number of which must be surrendered at the end of year to cover their emission amounts or a penalty must be paid. At the same time, allowances may be traded between firms, forming a secondary market of the instrument that is of interest to both the financial and actuarial industry seeking sustainable asset and liability portfolios. Research on an efficient allowance valuation model is therefore urgently called for.

In this thesis, we present valuation frameworks and models for allowances and options under closed and open trading phases. A trading phase consists of multiple years. Within each phase, banking and borrowing of allowances are permitted. Existing studies on allowance valuation fail to differentiate between the two types of phases and their implications on the modeling approaches. An ETS operating in a closed phase imposes a lapse of unused allowances at the end of phase, introducing terminal conditions of allowance prices for which structural models are required. We present three closed-phase allowance valuation models based on different specifications of the aggregate emission process: an Arithmetic Brownian Motion (ABM) emission rate, a Vasicek emission amount, and a Vasicek emission rate. In contrast, an ETS operating in an open phase permits inter-phase banking of unused allowances, the price of which is conveniently captured by reduced-form models. We investigate three open-phase allowance option valuation models based on different specifications of the allowance price process: a Lognormal model, a skewness kurtosis adjusted Lognormal model, and a Mixture Lognormal model. Closed-form expressions are derived for allowances, futures, and options as applicable. Market completeness and model applications are discussed, supplemented by numerical illustrations of model performance using actual market price data.

Furthermore, we extend the results of the open-phase models presented to the design and modeling of variable annuities backed by allowance-based managed funds and indexed annuities linked to the allowance markets. These products present an innovative yet feasible step forward for insurers and pensioners, as the allowance's non-lapsing feature and low return correlations with the stock market make it a suitable alternative asset for portfolio diversification. Analytical pricing, valuation, and hedging schemes are presented and numerically illustrated through semi-hypothetical examples.

Finally, we introduce a systematic framework for the management of sustainable equity investment portfolios under climate change, where metrics and methods are proposed to quantify investment-related climate change risks. Inferior historical risk-adjusted performance of carbon-intensive sectors are empirically demonstrated to justify the recommended divestment from these industries. A method to generate equity return impact scenarios is proposed, which can be incorporated into scenario-based calculations and models in various actuarial practices. Using the generated scenarios as key inputs, we present and illustrate the construction of a sustainable portfolio under a hypothetical asset universe.

Acknowledgements

I would like to express my utmost thanks to my supervisors, Professor Ken Seng Tan and Professor Tony Wirjanto, for their guidance and support throughout my research work and the completion of this thesis. I would also like to thank Professor Ming Bin Feng, Professor David Saunders, Professor Ilias Tsiakas, and Professor Olaf Weber for their kind suggestions that facilitate the progress in bringing this thesis to its completion.

The studies covered in this thesis are partially funded by the Society of Actuaries Hickman Scholarship and the Society of Actuaries Center of Actuarial Excellence Research Grant on the topic "Maintaining Financial Stability in an Era of Changing Climate and Demographics".

Dedication

This thesis is dedicated to my parents and loved ones.

Table of Contents

Chapter 1

Introduction

Rising awareness of the impacts of climate change and global warming triggered initiatives designated to control the output of greenhouse gases, commonly measured as carbon dioxide equivalents. Among these efforts, the establishment of emission trading schemes (ETS) is arguably the most interesting one. A typical ETS adopts a cap-and-trade strategy by using emission allowances. At the beginning of a compliance year, participating firms receive from the regulators an initial grant of allowance contracts, each of which allows outputting one equivalent tonne of carbon dioxide. At the end of a given year, firms must undergo compliance filing and submit sufficient allowances to cover their emissions over the year, or a non-compliance penalty must be paid on the excess emissions. The allowances can be traded freely on the secondary market along with their corresponding futures and option contracts.

A formal emission allowance market did not exist until the establishment of the European Union Emission Trading Scheme (EU ETS) in 2005. The EU ETS is divided into several trading phases, at the beginning of which the European Commission sets new regulations based on the preceding experience. Within each trading phase, allowances can be banked and borrowed between years: unused allowances at the end of a year can be carried over to future years, while any shortages can be covered by using the next year's allowances since the new allowance grants arrive prior to current year's compliance filing deadline. Over the past decade, the EU ETS has gone through 2 trading phases and is currently in Phase 3 spanning the period from 2013 to 2020. In the meantime, the market has also grown significantly in size, stability, and liquidity, making emission allowances an attractive asset class for many institutional investors with a preference toward sustainable carbon-neutral investment portfolios, as no other asset classes provide such a direct carbon

offset. Research on an efficient allowance valuation model is therefore urgently called for.

Recent developments on allowance valuation models can be broadly categorized into four streams of research. The first stream derives the allowance value using stochastic control techniques with the emission abatement mechanisms, pioneered by Seifert et al. (2008)[59]. This study assumes a stochastic emission process and the existence of a central planner who strives to minimize the total cost of emission abatements and non-compliance penalties by altering the abatement strategies. By solving this stochastic control problem with boundary conditions, the risk neutral allowance price is obtained. However, it lacks an analytical expression in general. Other studies in this stream includes Carmona et al. (2009)[20] and Carmona and Hinz (2011)[22], where the latter focuses on modeling the risk neutral emission abatement processes, whose marginal value is theoretically equal to the value of an allowance contract. The second stream of studies are reduced-form diffusion models for the spot allowance price dynamics. Literature in this stream varies widely in the assumptions made and in the scope of applications concerning market completeness. As an example, Cetin and Verscheure (2009)[25] study the EU ETS allowance prices and present a single-period model by assuming that the price follows a geometric Brownian Motion with a binary Markovian drift term representing the effect of new information arrival on the market's net emission position, which dictates the terminal spot price. This approach does not lead to an analytical allowance value expression under incomplete information. Other studies in this stream include Daskalakis et al. (2009)[31] who provide a comprehensive comparative analysis between alternative diffusion and jump-diffusion models for allowance futures prices and Manif (2012)[75] who presents both discrete and continuous time models of allowance prices under incomplete markets. The third stream of literature adopts econometric techniques in modeling allowance price returns. This includes, for example, Paollela and Taschini (2008)[81], Benz and Truck (2009)[90], Sun (2010)[88], and Shi (2014)[86]. Such models are less helpful for a valuation purpose as the underlying price drivers are often ignored completely. The fourth stream of literature consists of hybrid and structural models that derive the allowance value from underlying emission processes, where a common approach starts by specifying a distribution of the cumulative emission processes, which can be modeled using approximations (Grull and Kiesel, 2009[47]) or jump diffusions (Borokov et al., 2010[10]). Among the unpublished studies, Huang (2010)[55] proposes an allowance valuation framework under a single-period closed trading phase for a single-agent market, for which structural models are derived by specifying general forms of the emission dynamics.

In this thesis, we present several valuation models for emission allowances and related derivatives as well as the formulation of sustainable investment portfolios. The theories

and analysis are heavily based on the EU ETS directives where the respective emission allowance is termed the EU allowance (EUA). Our approach differs from existing studies in that we propose two fundamentally different frameworks based on the prevailing trading phase being open or closed, which has key implications that the existing studies fail to emphasize. In summary, a closed trading imposes expiration of unused allowances at the end of phase, and thereby induces a dichotomous terminal condition on allowance price that is best captured by structural models. In contrast, an open trading phase permits the banking of excess allowances into future phases, under which reduced-form allowance price models are appropriate. The proposed models balance complexity, accuracy, and implementability by yielding tractable closed-form expressions for allowance options values, which set them apart from other existing models. Note that throughout our discussion, the terms "price" and "value" may be used interchangeably as deemed appropriate in the contexts, though the models in this thesis are intended for valuation with potential applications for pricing purposes. This subtle point is worth being emphasize, and is discussed in more detail in Figlewski (1989)[44].

This thesis is organized as follows. Chapter two presents the allowance valuation framework under a closed trading phase, where structural models are built based on different specifications of the aggregate emission process. Numerical implementations and analysis are performed using Phase 1 spot market data. Chapter three discusses allowance option valuation under an open trading phase utilizing reduced-form allowance price models. Stylized facts of open phase allowance returns are analyzed and considered in the assessment of the models, which are calibrated to Phase 3 allowance futures option price data. Chapter four presents the design and modeling of allowance-linked variable and fixed indexed annuity products, supplemented with numerical demonstrations. Chapter five introduces the construction and management of sustainable equity portfolios under climate change, where the results can be easily adapted to include other asset classes. Chapter six concludes the thesis with a discussion on avenues for future studies.

Chapter 2

Models under a closed trading phase

A trading phase is defined as closed if banking and borrowing of allowances are permitted within each phase but prohibited across phases. More specifically, for any compliance year, the unused allowances can be carried over without any limit to future years within the phase, while any shortages can be covered using the next year's allowances since the new allowance grants arrive prior to the current year's compliance filing deadline. A perfect example of a closed trading phase is Phase 1 of the EU ETS scheme which spanned the period of 2005-2007. This three-year phase served as a pilot phase where deficiencies are revealed with criticisms received from the market and public. As the key distinguishing feature of a closed-trading phase, the payoff of an allowance converges to one of the two terminal values, inflating the price volatilities that respond to the arrival of information on the aggregate emission positions.

There are several reasons why we begin our discussion in this chapter with closed trading phases. First, studies targeted at closed trading phases are relatively few, where many past studies fail to emphasize the implications of a closed trading phase. The prohibition to banking the allowances to subsequent phases is a unique feature which induces terminal conditions for the allowance price. In addition, the effect of information incompleteness is more pronounced in a closed phase. Empirical analysis of allowance valuation under this setting can provide key insights to policy makers on improving information flow and completeness, benefiting thereby both the investors and the society at large. Thirdly, demonstrating the fundamental mechanics of a cap-and-trade scheme, closed trading phases remain relevant to date for emission markets such as the Kazakhstan ETS. Finally, from a forward-looking perspective, closed phases potentially serve as the piloting period of ETS with planned implementations in countries such as Ukraine and many developing coun-

tries in Asia and South America. Conclusions reached in this study can usefully serve as guidance for regulators and institutions currently operating under these emerging schemes.

Our contributions to the literature in this chapter include both theoretical and practical aspects. First, we propose a complete-market [1] allowance valuation framework under a closed trading phase, with extended applications under incomplete markets. A set of regularity assumptions are made and partially validated so that the framework accommodates multi-agent activities across years in the trading phase as in the real world situation. Second, we present structural models based on three different emission dynamics with time-homogeneous parameters[2], where closed-form expressions are derived for allowances, allowance futures, and options. The numerical efficiency resulted from analytical valuation formula is imperative in many scenario-based actuarial calculations for asset-liability management, as the nested stochastic runs otherwise required are often expensive. Thirdly, we discuss in detail the completeness of the allowance market and the source of incompleteness supported by empirical evidence, which is often omitted in the existing literature. While identifying limitations of the models, we also present in detail their applications under information incompleteness, with a focus on deriving implied emission values from actual market prices. Finally, all of the models and applications presented in this chapter are implemented and demonstrated using real world allowance data, based on which comparative analysis and conclusions are made. This provides further contribution to the literature than the existing works where simulation studies under subjectively assigned parameters are prominent and drive the numerical conclusions.

2.1 Framework specifications

Denote the filtered probability space by $(\Omega, \{\mathcal{F}_t\}, \mathcal{P})$, where $t \in [0, T]$ in units of years is the time index and T is the ending time of the trading phase. Assume that the emission market is complete and information-complete, whose justifications will be discussed later in this chapter. To motivate the idea, first consider a hypothetical single-period trading phase (e.g. $T = 1$) with two market participants: a firm trading allowances for compliance

[1] An allowance market is complete if and only if all derivatives traded have replicating portfolios (i.e. all risks are hedgeable). This strong assumption guarantees the existence and uniqueness of the risk neutral measure. See Bjork (2004)[6] for comprehensive discussions on market completeness

[2] Two of the emission dynamics were introduced with time-inhomogeneous parameters in Huang (2010)[55], which do not lead to closed-form valuation expressions in general

and a bank as the firm's counterparty. The total number of allowances is fixed and the bank has full information over the firm's emission process. We refer to the result as a simplified framework, which serves as the basis for all the models we develop. The terminal payoff of each allowance at the end of the trading phase, denoted by $S(T)$, is given by:

$$S(T) = \begin{cases} G & Q(0,T) > A(T) \\ 0 & Q(0,T) \leq A(T), \end{cases} \tag{2.1}$$

where:

1. $A(t)$ denotes the total number of allowances available at time t, which is the sum of the initial grant and the net allowance transactions in $(0, t)$.

2. $Q(t_1, t_2)$ is the firm's accumulated net emission amount during the period between t_1 and t_2, $0 \leq t_1 \leq t_2 \leq T$, including all emission abatements and eligible carbon offsets.

3. G is the penalty of non-compliance.

Due to such boundary conditions, it is improper to apply reduced-form models to the allowance values directly. Instead, the allowance should be viewed as a derivative of the net emission amounts. The two-participant setting leads to the ideal result that the spot price of the allowance naturally equals its marginal value to the firm. Thus we use $S(t)$ to denote both the time-t spot price and the value of an allowance contract, where the term "price" and "value" are used interchangeably as deemed appropriate in this chapter. To make the scope of the framework manageable, future trading activities are ignored at valuation points.

Assuming a continuously compounded risk free rate r, risk neutral valuation is applied to derive the expression for $S(t)$ under the equivalent martingale measure $\mathcal{Q}$ defined by

$$\frac{d\mathcal{Q}}{d\mathcal{P}} = \exp\left(- \int_0^T \Lambda(t) dZ(t) - \frac{1}{2} \int_0^T \Lambda(t)^2 dt \right),$$

where $Z(t)$ is the standard Brownian Motion under $\mathcal{P}$ and $\Lambda(t)$ is an $\mathcal{F}_t$-adapted process representing the market price of risk. The existence and uniqueness of $\mathcal{Q}$ is secured by market-completeness. From now on, we work under $\mathcal{Q}$ unless otherwise specified, where the filtered probability space becomes $(\Omega, \{\mathcal{F}_t\}, \mathcal{Q})$. The spot allowance price is given by:

$$S(t) = e^{-r(T-t)} E\left[G 1_{Q(t,T) > A(t) - Q(0,t)} \mid \mathcal{F}_t \right], \tag{2.2}$$

where 1_D denotes the binary variable that takes the value of 1 if condition D is satisfied and 0 otherwise. Stochasticity is driven by the random variable $Q(t,T)$ given $\mathcal{F}_t$, whose distribution must be specified and calibrated. Notice that the allowance at time t can be viewed as a binary call option on the future emission amount $Q(t,T)$ maturing at T with a strike price $A(t) - Q(0,t)$.

This simplified framework is the setting used in most existing studies. It is extended by Nel (2009)[77], and subsequently by Chesney and Taschini (2012)[26] to accommodate multiple market participants and compliance years by explicitly modeling the market clearings, which nonetheless requires intensive numerical evaluations for the allowance price. To maintain analytical tractability, we take an alternative approach to achieve such extension using two regularity assumptions discussed below.

Assumption 2.1: *There is a zero probability that the aggregate allowance supply equals the demand at T.*

Assumption 2.1 aligns the allowance values for all firms in the market despite their varying emission processes, and hence adapting the framework to accommodate multiple market participants. The allowance value under the assumption is solely driven by the aggregate emission process as given in Proposition and Corollary 2.1 below.

Proposition 2.1: *Consider a hypothetical single-period closed trading phase in the time interval $[0,T]$ with N participants indexed by $i, i \in \{1,2,...N\}$, where each firm i is subject to a non-compliance penalty of G and an initial emission allowance grant of A_g^i. Let $Q^i(t_1,t_2)$ denote the net emission amount of firm i over the period $[t_1,t_2]$. Assume there is a zero probability that the aggregate allowance supply equals the demand at T. That is:*

$$Pr\left(Q^C(0,T) = A_g^C\right) = 0 \quad a.s.$$

where

$$Q^C(0,T) = \sum_{i=1}^{N} Q^i(0,T).$$

$$A_g^C = \sum_{i=1}^{N} A_g^i.$$

Then, the terminal payoff of an allowance contract is determined by the market's aggregate emission as:

$$S(T) = \begin{cases} G & Q^C(0,T) > A_g^C \\ 0 & Q^C(0,T) \leq A_g^C \end{cases}$$

Proof: Define the market position in allowances to be short if the aggregate emission amount exceeds the aggregate number of allowances and long otherwise. In addition, let $A^i(t)$ denote the total number of allowances available for firm i at time t. At $t = T$, since $Pr\left(Q^C(0,T) = A_g^C\right) = 0 \quad a.s.$, one of the followings must take place:

1. $Q^C(0,T) > A_g^C$. The market is short in allowance. This implies that at least one market participant is short and must purchase allowances to cover the excess emission. Transactions take place until all sellers have just enough allowances to cover their emissions, after which transactions will only be made at the penalty level G. Consider the allowance's terminal payoff to firm i, where two possible scenarios may take place:
   (a) $Q^i(0,T) > A^i(T)$. The firm is short. For each ton of excess emission, it either pays the penalty G or buys an allowance at the terminal price G for compliance. The payoff of each allowance is G.
   (b) $Q^i(0,T) < A^i(T)$. The firm is long. It sells each excess allowance at the terminal price G. The payoff of each allowance is G.
2. $Q^C(0,T) < A_g^C$. The market is long in allowance. This implies that at least one market participant is long and must sell the excess allowances which will otherwise be worthless. Transactions take place until all buyers have just enough allowances to cover their emissions, after which transactions can only be made at a price of 0. Consider the allowance's terminal payoff to firm i, where two possible scenarios may take place:
   (a) $Q^i(0,T) < A^i(T)$. The firm is long. It either lapses the extra allowances or sells each at the price 0. The payoff of each allowance is 0
   (b) $Q^i(0,T) > A^i(T)$. The firm is short. For each ton of excess emission, it buys an allowance at the terminal price 0 for compliance. The payoff of each allowance is 0.

Therefore,the payoff of an emission allowance is as described in Proposition 2.1. **End of proof**.

Since the payoff structure is, loosely speaking, independent[3] of each firm's own emission process, a corollary of Proposition 2.1 directly follows.

Corollary 2.1: *Consider a hypothetical market with the same settings and assumptions as in Proposition 2.1, the value of an emission allowance is common to all market participants at each time point* $t \in [0, T]$ *and is equal to the spot allowance price at time* t*.*

Next, to accommodate multiple compliance years, we define emission solvency and make Assumption 2.2 as described below.

Definition 2.1: *An emission market participant is defined as emission-solvent over the trading phase* $[0, T]$ *if it remains compliant during the period* $[0, T-1]$ *without purchasing allowances. More formally, a firm indexed by* $i, i \in \{1, 2, ...N\}$*, is emission-solvent if and only if:*

$$\sum_{t=1}^{\tau} Q^i\left(t-1, t\right) \leq \sum_{t=1}^{\tau+1} A_g^i(t) - \sum_{t=1}^{\tau} X_s^i\left(t-1, t\right) \quad \forall \tau \in 1, 2, ...T-1,$$

where $A_g^i(t)$ *is the allowance grant in year* t *to the firm, and* $X_s^i\left(t_1, t_2\right)$ *is the number of allowances sold by the firm in the period* (t_1, t_2)*.*

Assumption 2.2: *All market participants remain emission-solvent over* $[0, T]$*.*

Definition 2.1 follows from the intra-phase borrowing mechanism in closed phases as in the EU ETS Phase 1, where allowances for the new year are granted before the compliance filing date for the current year. This makes compliance easier since it is very unlikely that the firm's annual emission exceeds the sum of two years' allowance quota. On the other hand, sales of allowances usually occur only on solid expectations that such borrowing provides sufficient room for future emissions. According to the study by Sevendson and Vesterdal (2003)[85], large emitters are price takers in the market and the risk of having to buy back the allowances at a much higher price forces the firms to make prudential and conservative selling decisions. Hence, Assumption 2.2 is reasonable, under which allowances are not subject to surrendering until T when it pays off.

[3]The term "independent" is used here to primarily emphasize that the firms' own emission processes, when viewed individually, are not key drivers of the allowance price. Obviously, since the aggregate emission is the sum of individual firm-level emissions, there is inherently some dependence.

The valuation framework is therefore extended under the two regularity assumptions to serve a general market with N participants and multiple compliance years, where $T \geq 1$:

$$S(T) = \begin{cases} G & Q^C(0,T) > \alpha \\ 0 & Q^C(0,T) \leq \alpha, \end{cases} \tag{2.3}$$

where

$$\alpha = \sum_{t=1}^{T} \sum_{i=1}^{N} A_g^i(t) = \sum_{t=1}^{T} A_g^C(t).$$

The spot price of the allowance at time t is given by:

$$S(t) = e^{-r(T-t)} E\left[G 1_{Q^C(t,T) > \alpha - Q^C(0,t)} \mid \mathcal{F}_t \right]. \tag{2.4}$$

Structural valuation models are built under the framework by specifying the aggregate emission process. To further simplify the notation, we drop the superscript "C" that denote the market aggregate variable whenever the context is clear.

2.2 Allowance and futures valuation

In this section, three allowance valuation models are built under the framework based on different specifications of the emission dynamics. The valuation expressions and related key results are derived for allowances and allowance futures.

2.2.1 ABM emission rate model

We first present a model in which the market's aggregate rate of emission follows an Arithmetic Brownian Motion (ABM) with time-homogeneous drift and volatility parameters. As we will see later on, the model yields a good fit while having the advantage of maintaining analytical tractability compared to its time-inhomogeneous counterpart introduced in Huang (2010)[55]. Let $y(t)$ denote the emission rate described by the following diffusion process:

$$dy(t) = \mu dt + \sigma dW(t), \quad y(0) = y_0, \tag{2.5}$$

where μ and σ are respectively the drift and volatility of the emission rate process, and y_0 is the initial emission rate at time 0. $W(t)$ is a standard Brownian Motion under $\mathcal{Q}$ for $t \in [0, T]$. Given the σ-field $\mathcal{F}_s, 0 \leq s < t < T$:

$$y(t) = y(s) + \mu(t-s) + \sigma(W(t) - W(s)). \tag{2.6}$$

That is, the emission rate is a normally distributed random variable whose support includes negative values. This captures situations where carbon offsets are obtained or qualifying green projects are taken at times to reduce the accumulated emission amounts, which are treated as negative emission-flows. Given $\mathcal{F}_s$, $0 \leq s < t \leq T$:

$$Q(s,t) = \int_s^t y(u)du = \beta(s,t) - \sigma W(s)(t-s) + \sigma \int_s^t W(u)du,$$

where,

$$\beta(s,t) = y(s)(t-s) + 0.5\mu\left(t^2 - s^2\right) - \mu s(t-s). \tag{2.7}$$

Not surprisingly, the accumulated emission over any period is an increasing function of the period-start emission rate. Furthermore, it can be shown that

$$\int_s^t (W(u))du \sim N\left(W(s)(t-s), \sqrt{\frac{(t-s)^3}{3}}\right), \tag{2.8}$$

where $N(.,.)$ denotes a Normal distribution with the two arguments being the mean and standard deviation parameters respectively. Appendix A shows the proof of this result, which leads to:

$$Q(s,t) \mid \mathcal{F}_s \sim N\left(\beta(s,t), \sigma\sqrt{\frac{(t-s)^3}{3}}\right). \tag{2.9}$$

The accumulated emission also has a support including the negative portion of the real line, which reflects the possible situations that the emission offsets exceed the emission amounts over particular time periods. Such a feature also applies to the other two models to be introduced later in this chapter.

Following the above results, the spot allowance price under the ABM emission rate model is given by Proposition 2.2 below.

Proposition 2.2: *Consider a closed trading phase spanning the period* $(0, T)$ *with a non-compliance penalty of* G *and a total emission allowance grant of* α*, where Assumption 2.1*

and 2.2 apply. Assume that the aggregate emission rate $y(t)$ follows the arithmetic Brownian Motion governed by (2.5). The risk neutral price of the allowance at time t is given by:

$$S(t) = e^{-r(T-t)} G\Phi\left(\frac{\beta(t,T) - \alpha(t)}{\sigma\sqrt{\frac{(T-t)^3}{3}}}\right),$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of a standard Normal random variable and

$$\beta(t,T) = y(t)(T-t) + 0.5\mu\left(T^2 - t^2\right) - \mu t(T-t),$$
$$\alpha(t) = \alpha - Q(0,t).$$

Proof: see Appendix B.

In reality, each company has full information on its own emission status but, at best, only partial information on other companies'. The actual emission amounts are reported to the national registries on an annual basis during compliance filing at the end of each April, which is subsequently released to the public[4]. Therefore, although $y(t)$ and $Q(0,t)$ are $\mathcal{F}_t$-measurable, the information is not accessible to the general public except on the release dates. The implication of such information incompleteness is discussed in Section 2.4. At this stage we maintain the assumption of information and market completeness. A tractable expression for the allowance price dynamics is given by:

$$dS(t) = rS(t)dt + \frac{\sqrt{3}G}{\sqrt{2\pi(T-t)}} e^{-\left(r(T-t) + 0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2\right)} dW(t), \quad S(0) = S_0, \quad (2.10)$$

where S_0 is the initial allowance price at time $t = 0$. The terminal condition of the process follows the dichotomy in (2.3). The general solution to this stochastic differential equation is established algebraically for any $\mathcal{F}_s$, $0 \leq s < t < T$:

$$S(t) = e^{-r(T-t)} G\Phi\left(\frac{\sqrt{\frac{(T-s)^3}{3}}\Phi^{-1}\left(\frac{S_s e^{r(T-s)}}{G}\right) + \lambda(s,t)}{\sqrt{\frac{(T-t)^3}{3}}}\right), \quad (2.11)$$

[4]Nowadays, technology advancement has provided cost-efficient means of continuous emission monitoring, which renders a high frequency emission reporting possible.

where

$$\lambda(s,t) = T\left(W(t) - W(s)\right) + sW(s) - tW(t) + \int_s^t W(u)du. \tag{2.12}$$

Several important observations can be readily made here. First, the allowance price process does not contain any emission parameters or variables. This desirable property is uncommon, as shall be demonstrated in the other two structural models in this Chapter. The only unknown parameter is the risk free rate, which can be specified exogenously. It follows that numerical procedures can be conveniently used to perform simulation studies of the allowance price movements without having to estimate the emission-specific parameters, which in turn allows us to study the behavior of more exotic allowance derivatives and customized contracts. In addition, the process retains its natural Markovian property. This suggests efficiency in parameter estimation and numerical simulation as well as applicability of many established methodologies. Lastly, under discrete approximations to the stochastic integrals, a systematic algorithm can be designed to discover the emission information from allowance prices. As expected from the defining property of a risk neutral measure, the discounted allowance price process is shown to be a martingale.

$$d(e^{-rt}S(t)) = \frac{\sqrt{3}G}{\sqrt{2\pi(T-t)}} e^{-\left(rT + 0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2\right)} dW(t). \tag{2.13}$$

Next, denote by $F(t,\tau')$ the time-t price of an allowance futures contract to be settled at τ' where $0 \leq t < \tau' \leq T$. Assume that the market is arbitrage free and that cash is borrowed and invested at the risk free rate. Using arbitrage-free pricing and Proposition 2.2:

$$F(t,\tau') = e^{r(\tau'-t)}S(t) = e^{-r(T-\tau')}G\Phi\left(\frac{\beta(t,T) - (\alpha(t))}{\sigma\sqrt{\frac{(T-t)^3}{3}}}\right), \tag{2.14}$$

where $\beta(t,T)$ and $\alpha(t)$ are as defined in Proposition 2.2. The corresponding futures price dynamics is given by:

$$dF(t,\tau') = \frac{\sqrt{3}G}{\sqrt{2\pi(T-t)}} e^{-\left(r(T-\tau') + 0.5\Phi^{-1}\left(\frac{F(t,\tau')e^{r(T-\tau')}}{G}\right)^2\right)} dW(t) \tag{2.15}$$

with the initial condition $F(0,\tau') = F_0$ and the terminal condition $F(\tau',\tau') = S(\tau')$. This shows that the futures price is also a martingale under $\mathcal{Q}$. The general solution to (2.15)

conditioning on $\mathcal{F}_s$ is obtained algebraically using the results of (2.11) and (2.14), giving:

$$F(t,\tau') = e^{-r(T-\tau')} G\Phi \left(\frac{\sqrt{\frac{(T-s)^3}{3}}\Phi^{-1}\left(\frac{F_s e^{r(T-\tau')}}{G}\right) + \lambda(s,t)}{\sqrt{\frac{(T-t)^3}{3}}} \right), \tag{2.16}$$

where $\lambda(s,t)$ is as defined in (2.12). These results will be used in calibrating the model to allowance futures data.

2.2.2 Vasicek emission model

Next, we propose a structural model assuming that the accumulated emission amount $Q(t)$ follows a Vasicek process. The motivation for assuming mean-reverting emissions ties to emission planning: following the release of new directives at the beginning of a trading phase, each firm has the opportunity to budget its emission activities based on the business forecasts, the allowance grant, and the penalty for excess emission. Thus, there exists an emission budget for each firm over the phase, which displays a mean-reverting behavior in the emission market. Although the empirical analysis reported in Sun (2010)[88] pointed to some evidence of mean-reversion embedded in the Phase 1 allowance returns based on econometric models, studies of structural mean-reversion in the current literatures are limited and do not seem to have reached clear consensus on this issue.

The Vasicek emission model specifies a time-homogeneous mean-reverting emission model. The reversion level parameter is a constant representing a long-term target emission amount independent of the allowance market activities. From a broader scope, this emission level can be interpreted as the ultimate target of the emission control scheme, at which additional emissions are precisely offset by abatement activities. Numerically, it introduces concavity in the emission amount over time, which increases at a diminishing rate during the phase. This captures a long term feedback process that differs from the ABM model under which the emission amount follows a generally linear trend.

Assume that the accumulated emission $Q(0,t)$ follows a Vasicek process:

$$dQ(0,t) = \kappa(\theta - Q(0,t))dt + \sigma dW(t), \tag{2.17}$$

with the initial condition $Q(0,0) = 0$. This is an Ornstein-Uhlenceck process and it can be easily shown that the general solution to the equation given $\mathcal{F}_s$ is:

$$Q(0,t) = Q(0,s)e^{-\kappa(t-s)} + \theta\left(1 - e^{-\kappa(t-s)}\right) + \sigma e^{-\kappa t}\int_s^t e^{\kappa u} dW(u). \tag{2.18}$$

Conceptually, θ represents the long-term target emission level. The parameter κ represents a reversion rate reflecting the market participants' adjustments to bring the aggregate emission amount to the target. The stochastic integral term is conditionally normally distributed with a mean of 0, where the variance term can be found using Ito's isometry, giving the key result:

$$Q(s,t) \mid \mathcal{F}_s \sim N\left(\beta^*(s,t), \sigma\sqrt{\frac{\left(1 - e^{-2\kappa(t-s)}\right)}{2\kappa}}\right), \tag{2.19}$$

where

$$\beta^*(s,t) = (\theta - Q(0,s))\left(1 - e^{-\kappa(t-s)}\right). \tag{2.20}$$

A proof of this result is presented in Appendix E.

Following the above results, the spot allowance price under Vasicek emission is given by Proposition 2.3 below.

Proposition 2.3: *Consider a closed trading phase spanning the period $(0,T)$ with a non-compliance penalty of G and a total emission allowance grants of α, where Assumption 2.1 and 2.2 apply. Assume that the aggregate emission rate $Q(t)$ follows the Vasicek process governed by (2.17). The risk neutral price of the allowance at time t is given by:*

$$S(t) = e^{-r(T-t)}G\Phi\left(\frac{(\beta^*(t,T) - \alpha(t))}{\sigma\sqrt{\frac{\left(1 - e^{-2\kappa(T-t)}\right)}{2\kappa}}}\right),$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of a standard Normal random variable and

$$\beta^*(t,T) = (\theta - Q(0,t))\left(1 - e^{-\kappa(T-t)}\right),$$
$$\alpha(t) = \alpha - Q(0,t).$$

Proof: see Appendix F.

This result can be equivalently expressed as:

$$S(t) = e^{-r(T-t)} G\Phi\left(\frac{(\beta(t,T)-\alpha)}{\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}}}\right), \tag{2.21}$$

where

$$\beta(t,T) = Q(0,t)e^{-\kappa(T-t)} + \theta\left(1-e^{-\kappa(T-t)}\right). \tag{2.22}$$

In fact, the equation in (2.21) is more convenient to use in practice. By differentiating both sides of (2.21), we obtain the risk neutral allowance price dynamics:

$$dS(t) = rS(t)dt + \frac{G\sqrt{\kappa}}{\sqrt{\pi\left(e^{2\kappa(T-t)}-1\right)}} e^{-\left(r(T-t)+0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2\right)} dW(t). \tag{2.23}$$

The process has the initial condition $S(0) = S_0$ and the terminal condition specified by the dichotomy in (2.3). A proof of this result is given by Appendix G. A closed-form general solution to this equation given $\mathcal{F}_s$ is established algebraically for $0 \le s < t \le T$:

$$S(t) = e^{-r(T-t)} G\Phi\left(\frac{\sigma e^{\kappa(t-s)}\sqrt{\frac{\left(1-e^{-2\kappa(T-s)}\right)}{2\kappa}}\Phi^{-1}\left(\frac{S_s e^{r(T-s)}}{G}\right) + \nu(s,t)}{\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}}}\right), \tag{2.24}$$

where

$$\begin{aligned}\nu(s,t) = &\left(e^{\kappa(t-s)}-1\right)\left(\alpha-(1-e^{-\kappa T})\theta\right) \\ &+ \sigma e^{-\kappa T}\left(\int_s^t e^{\kappa u}dW(u) + \left(1-e^{\kappa(t-s)}\right)\int_0^s e^{\kappa u}dW(u)\right).\end{aligned} \tag{2.25}$$

Unlike the case under the ABM emission rate model, the allowance price process under the Vasicek emission model contains emission-specific parameters. This introduces complexity to model calibration, whereby the algorithm for fitting the ABM emission rate cannot be applied to the Vasicek emission model directly. The differences will be addressed later,

when we discuss the model implementations and numerical results. Similar to the ABM emission rate model, the discounted allowance price under Vasicek emission is also a martingale under $\mathcal{Q}$:

$$\begin{aligned} d(e^{-rt}S(t)) &= -re^{-rt}S(t)dt + e^{-rt}dS(t) \\ &= \frac{G\sqrt{\kappa}}{\sqrt{\pi\left(e^{2\kappa(T-t)}-1\right)}} e^{-\left(rT+0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2\right)} dW(t). \end{aligned} \tag{2.26}$$

Now consider the price of an allowance futures contract to be settled at τ', where $0 \leq s < t < \tau' \leq T$. By arbitrage-free pricing and Proposition 2.3, we get:

$$F(t,\tau') = e^{r(\tau-t)}S(t) = e^{-r(T-\tau')}\left(\frac{(\beta^*(t,T)-\alpha(t))}{\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}}}\right), \tag{2.27}$$

where $\beta^*(t,T)$ is defined in Proposition 2.3. The futures price dynamics is given by:

$$dF(t,\tau') = \frac{G\sqrt{\kappa}}{\sqrt{\pi\left(e^{2\kappa(T-t)}-1\right)}} e^{-\left(r(T-\tau')+0.5\Phi^{-1}\left(\frac{F(t,\tau')e^{r(T-\tau')}}{G}\right)^2\right)} dW(t), \tag{2.28}$$

with the initial condition $F(0,\tau') = F_0$ and the terminal condition $F(\tau',\tau') = S(\tau')$. This shows that under Vasicek emissions, the futures price remains a martingale under $\mathcal{Q}$. The general solution to (2.28) given $\mathcal{F}_s$ is obtained algebraically using the results of (2.24) and (2.27):

$$F(t,\tau') = e^{-r(T-\tau')}G\Phi\left(\frac{\sigma e^{\kappa(t-s)}\sqrt{\frac{\left(1-e^{-2\kappa(T-s)}\right)}{2\kappa}}\Phi^{-1}\left(\frac{F_s e^{r(T-\tau')}}{G}\right)+\nu(s,t)}{\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}}}\right), \tag{2.29}$$

where $\nu(s,t)$ is defined in (2.25). These results are useful for calibrating the model to the allowance futures market data.

2.2.3 Vasicek emission rate model

Finally, we present a model where the market's aggregate rate of emission follows a Vasicek process with time-homogeneous drift and volatility parameters. Similar to the Vasicek emission model, such a specification captures the feedback process induced jointly by the ETS allowance quota and the production budget, where periods of rapid emission trigger heavier abatement to reduce the likelihood of non-compliance, and those of slow emission cause accelerated production toward the equilibrium level. Since adjustments taken by a firm are directly reflected in the rate of emission as opposed to the emission amount, the mean-reverting specification to a target emission rate is more reasonable and intuitive compared to a long term target emission amount given a closed trading phase.

Assume that the emission rate $y(t)$ is described by the following stochastic differential equation:

$$dy(t) = \kappa(\theta - y(t))dt + \sigma dW(t), \quad y(0) = y_0, \tag{2.30}$$

where θ represents the target emission rate, and κ is the reversion rate reflecting the market's adjustments to bring the emission rate to the target. This is an Ornstein-Uhlenbeck process with a general solution given $\mathcal{F}_s$, $0 \leq s < t < T$:

$$y(t) = y(s)e^{-\kappa(t-s)} + \theta(1 - e^{-\kappa(t-s)}) + \sigma e^{-\kappa t} \int_s^t e^{ku} dW(u). \tag{2.31}$$

Conditioning on $\mathcal{F}_s$, $0 \leq s < t \leq T$:

$$Q(s,t) = \int_s^t y(u)du = \beta(s,t) + \frac{\sigma}{\kappa} \int_s^t \left(1 - e^{-\kappa(t-u)}\right) dW(u), \tag{2.32}$$

where

$$\beta(s,t) = \frac{y(s)}{\kappa}\left(1 - e^{-\kappa(t-s)}\right) + \theta\left(t - s - \frac{1}{\kappa}\left(1 - e^{-\kappa(t-s)}\right)\right).$$

Furthermore, it is shown in Appendix I that:

$$Q(s,t) \mid \mathcal{F}_s \sim N\left(\beta(s,t), \frac{\sigma}{\kappa}\sqrt{t - s - \frac{2}{\kappa}\left(1 - e^{-\kappa(t-s)}\right) + \frac{1}{2\kappa}\left(1 - e^{-\kappa(t-s)}\right)}\right). \tag{2.33}$$

Hence conditional normality of the emission amount variable is maintained under the Vasicek emission rate specification, making the rest of the analysis methodologically similar

to those for the previous two models.

Following the above results, the spot allowance price is given by Proposition 2.4 below.

Proposition 2.4: *Consider a closed trading phase spanning the period* $(0, T)$ *with a non-compliance penalty of* G *and a total emission allowance grants of* α*, where Assumption 2.1 and 2.2 apply. Assume that the aggregate emission rate* $y(t)$ *follows the Vasicek process in (2.30). The risk neutral price (and value) of the allowance at time* t *is given by:*

$$S(t) = e^{-r(T-t)} G\Phi\left(\frac{\beta(t,T) - \alpha(t)}{\frac{\sigma}{\kappa}\sqrt{T-t-\frac{2}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-t)}\right)}}\right),$$

where $\Phi(\cdot)$ *denotes the cumulative distribution function of a standard Normal random variable and*

$$\begin{aligned}\beta(t,T) &= \frac{y(t)}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\theta\left(T-t-\frac{1}{\kappa}\left(1-e^{-\kappa(T-t)}\right)\right),\\ \alpha(t) &= \alpha - Q(0,t).\end{aligned}$$

The proof for Proposition 2.4 is given in Appendix J.

Similar to the cases for the previous two models, $y(t)$ and $Q(0,t)$ are $\mathcal{F}_t$-measurable but not generally accessible to the public except at compliance release dates. Being derived in Appendix K, the allowance price dynamics follows:

$$dS(t) = rS(t)dt + \frac{G\left(1-e^{-\kappa(T-t)}\right)e^{-\left(r(T-t)+0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2\right)}}{\sqrt{2\pi\left(T-t-\frac{2}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-t)}\right)\right)}}dW(t), \quad (2.34)$$

with the initial condition $S(0) = S_0$. The terminal condition of the process follows the dichotomy in (2.3). The solution to this stochastic differential equation is established

algebraically in Appendix L for any $\mathcal{F}_s$, $0 \leq s < t < T$:

$$S(t) = \frac{G}{e^{r(T-t)}}\Phi\left(\frac{\sqrt{T-s-\frac{2}{\kappa}\left(1-e^{-\kappa(T-s)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-s)}\right)}\Phi^{-1}\left(\frac{S_s e^{r(T-s)}}{G}\right)+\xi(s,t)}{\sqrt{T-t-\frac{2}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-t)}\right)}}\right), \tag{2.35}$$

where

$$\xi(s,t) = \int_s^t \left(1-e^{-\kappa(T-u)}\right) dW(u). \tag{2.36}$$

Unlike the ABM emission rate model but similar to the Vasicek emission, the process under the Vasicek emission rate model contains the emission parameter κ, adding complexity to model calibrations. However, numerical efficiency is largely retained by the absence of emission variables in the process. In addition, by examining the drift and diffusion terms in (2.34), it is easy to see that the discounted allowance price process remains a martingale under $\mathcal{Q}$ as in the previous two models presented.

Following Proposition 2.4 and arbitrage free pricing, the time-t price for an allowance futures contract maturing at τ', $0 \leq t < \tau' \leq T$ is given by.

$$F(t,\tau') = Ge^{-r(T-\tau')}\Phi\left(\frac{\beta(t,T)-\alpha(t)}{\frac{\sigma}{\kappa}\sqrt{T-t-\frac{2}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-t)}\right)}}\right), \tag{2.37}$$

where $\beta(t,T)$ and $\alpha(t)$ are defined in Proposition 2.4. The futures price dynamics is given by:

$$dF(t,\tau') = \frac{G\left(1-e^{-\kappa(T-t)}\right)e^{-\left(r(T-\tau')+0.5\Phi^{-1}\left(\frac{F(t,\tau')e^{r(T-\tau')}}{G}\right)^2\right)}}{\sqrt{2\pi\left(T-t-\frac{2}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-t)}\right)\right)}}dW(t), \tag{2.38}$$

with the initial condition $F(0,\tau') = F_0$ and the terminal condition $F(\tau',\tau') = S(\tau')$. Hence the futures price under the Vasicek emission rate is also a martingale under $\mathcal{Q}$. The general

solution to (2.38) given $\mathcal{F}_s$ is obtained algebraically to arrive at:

$$F(t,\tau') = \tag{2.39}$$
$$\frac{G}{e^{r(T-\tau')}}\Phi\left(\frac{\sqrt{T-s-\frac{2}{\kappa}\left(1-e^{-\kappa(T-s)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-s)}\right)}\Phi^{-1}\left(\frac{F_s e^{r(T-\tau')}}{G}\right)+\xi(s,t)}{\sqrt{T-t-\frac{2}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-t)}\right)}}\right),$$

where $\xi(s,t)$ is defined in (2.36).

2.3 Option valuation

In this section, we present the valuation results for allowance options and allowance futures options under the three structural models specified in Section 2.2.

2.3.1 General results for allowance option valuation

Allowance options traded under the EU ETS framework and other major emission regulatory directives are European style. As discussed previously, each allowance can be considered a binary call option over the future emission amounts. Hence, an allowance option can be viewed as a compound option on the binary call maturing at T. Allowance option valuation has been discussed in Carmona and Hinz (2011)[22], Huang (2010)[55], Cetin et al.(2009)[25], and Borokov et al.(2010)[10]. We contribute to the literature by presenting a set of key analytical results for allowance option prices under the structural models we build. Let $C(t,K,\tau)$ and $P(t,K,\tau)$ denote respectively the time-t values of allowance call and put options with a strike price K and maturity τ where $0 \leq t < \tau \leq T$. The following model-free propositions are conveniently introduced.

Proposition 2.5: *Consider a closed trading phase spanning the period* $(0,T)$ *with a non-compliance penalty of* G *and a total emission allowance grant of* α *where Assumption 2.1 and 2.2 apply. If the time-*τ *allowance price given* $\mathcal{F}_t$ *can be expressed as follows*

$\forall 0 \leq t < \tau \leq T$:

$$S(\tau) = Ge^{-r(T-\tau)}\Phi\left(cX(t,\tau) + \zeta\right)$$

and

$$X(t,\tau) \mid \mathcal{F}_t \sim N(\mu_x, \sigma_x),$$

where c, ζ, μ_x, σ_x *are* $\mathcal{F}_t$*-measurable functions, then the time-*t *risk neutral price of a European allowance call option with strike price* K *maturing at time* τ *is given by:*

$$C(t,K,\tau) = Ge^{-r(T-t)}\left(\Phi_2\left(\frac{\mu_x - x^*}{\sigma_x}, \frac{c\mu_x + \zeta}{\sqrt{1+c^2\sigma_x^2}}, \rho\right) - K^*\Phi\left(\frac{\mu_x - x^*}{\sigma_x}\right)\right),$$

where:

$$K^* = \frac{e^{r(T-\tau)}}{G}K,$$
$$x^* = \frac{\Phi^{-1}(K^*) - \zeta}{c},$$
$$\rho = \frac{c\sigma_x}{\sqrt{1+c^2\sigma_x^2}},$$

and the notation $\Phi_2(x_1, x_2, \rho)$ *denotes the bivariate standard Normal distribution function with correlation coefficient* ρ *evaluated at* (x_1, x_2).

The prices for allowance futures options can be derived in a similar manner, or using the result of Proposition 2.5 and Corollary 2.2 below. Let $C_F(t,K,\tau,\tau')$ and $P_F(t,K,\tau,\tau')$ denote respectively the time-t prices of an allowance futures call and put option with a strike price K and a maturity τ, written on an allowance futures to be settled at time τ', $0 \leq t < \tau \leq \tau' \leq T$. Holders of this call/put have the right (but not the obligation) to enter the long/short position of the underlying futures contract at time τ with a settlement price K. Allowance futures options are standardized and exchange-traded contracts with higher transaction volumes than the allowance options. Payoffs are realized through cash settlements, benefiting from and contributing to the high liquidity in the underlying futures market. Corollary 2.2 below gives the relationship between the spot and futures option values.

Corollary 2.2: *Consider a closed trading phase spanning the period* $(0,T)$ *with a non-compliance penalty of* G *and a total emission allowance grant of* α *where Assumption 2.1 and 2.2 apply. Then:*

$$C_F(t,K,\tau,\tau') = e^{r(\tau'-\tau)}C(t,K',\tau),$$

$$P_F(t, K, \tau, \tau') = e^{r(\tau'-\tau)} P(t, K', \tau),$$

where

$$K' = Ke^{-r(\tau'-\tau)}.$$

See Appendix N for the proof of this result.

2.3.2 Option prices under an ABM emission rate

Here we present the option valuation formula under the ABM emission rate model. For brevity, derivations are presented for an allowance call option with strike K and maturity τ. Given $\mathcal{F}_t$, $0 \leq t < \tau \leq T$, (2.11) gives:

$$S(\tau) = Ge^{-r(T-\tau)}\Phi\left(\frac{\sqrt{\frac{(T-t)^3}{3}}\Phi^{-1}\left(\frac{S_t e^{r(T-t)}}{G}\right) + \lambda(t,\tau)}{\sqrt{\frac{(T-\tau)^3}{3}}}\right). \tag{2.40}$$

The stochastic component $\lambda(t, \tau)$ can be decomposed as:

$$\begin{aligned}\lambda(t,\tau) &= T\left(W(\tau) - W(t)\right) + tW(t) - \tau W(\tau) + \int_t^\tau W(u)du \\ &= -(T-t)W(t) + X(t,\tau),\end{aligned} \tag{2.41}$$

where

$$X(t,\tau) = (T-\tau)W(\tau) + \int_t^\tau W(u)du. \tag{2.42}$$

To further simplify the notation, define:

$$\zeta(t,\tau) = \frac{\sqrt{\frac{(T-t)^3}{3}}\Phi^{-1}\left(\frac{S_t e^{r(T-t)}}{G}\right) - (T-t)W(t)}{\sqrt{\frac{(T-\tau)^3}{3}}}. \tag{2.43}$$

This allows us to rewrite (2.40) as:

$$S(\tau) = Ge^{-r(T-\tau)}\Phi\left(\sqrt{\frac{3}{(T-\tau)^3}}X(t,\tau) + \zeta(t,\tau)\right), \tag{2.44}$$

Given $\mathcal{F}_t$, $\zeta(t,\tau)$ is deterministic, while $X(t,\tau)$ is a normally distributed variable that can be shown to follow:

$$X(t,\tau) \mid \mathcal{F}_t \sim N\left(\mu_x, \sigma_x\right), \tag{2.45}$$

where

$$\mu_x = (T-t)W(t), \tag{2.46}$$

$$\sigma_x^2 = (T-\tau)^2(\tau-t) + (\tau-t)^2(T-\tau) + \frac{(\tau-t)^3}{3}. \tag{2.47}$$

The derivation of these results is given in Appendix O. Proposition 2.5 applies to arrive at the time-t allowance call option price under the ABM emission rate model:

$$C(t,K,\tau) = Ge^{-r(T-t)}\left(\Phi_2\left(\frac{\mu_x - x^*}{\sigma_x}, \frac{\mu_x\sqrt{\frac{3}{(T-\tau)^3}} + \zeta(t,\tau)}{\sqrt{1+\frac{3}{(T-\tau)^3}\sigma_x^2}}, \rho\right) - K^*\Phi\left(\frac{\mu_x - x^*}{\sigma_x}\right)\right), \tag{2.48}$$

where $\zeta(t,\tau)$, μ_x, σ_x, K^*, are defined in (2.43), (2.46), (2.47), and Proposition 2.5 respectively,

$$x^* = \left(\Phi^{-1}\left(K^*\right) - \zeta(t,\tau)\right)\sqrt{\frac{(T-\tau)^3}{3}}, \tag{2.49}$$

and

$$\rho = \frac{\sigma_x\sqrt{\frac{3}{(T-\tau)^3}}}{\sqrt{1+\frac{3\sigma_x^2}{(T-\tau)^3}}}. \tag{2.50}$$

The price of the put option under the model is found using the put-call parity relationship[5]:

$$P(t,K,\tau) = C(t,K,\tau) - S(t) + Ke^{-r(\tau-t)}, \tag{2.51}$$

[5]The put-call parity is built from the arbitrage-free pricing of a synthetic forward consisting of a pair of European call and put options. It is market and model independent so that the result can be directly applied.

where $C(t, K, \tau)$ is the corresponding European call option price given by (2.48).

The allowance futures option value follows by applying Corollary 2.2 to (2.48). Minor modifications to the deterministic quantities are made to simplify the result. Let:

$$R = \frac{e^{r(T-\tau')}}{G} K, \tag{2.52}$$

$$\zeta(t, \tau, \tau') = \frac{\sqrt{\frac{(T-t)^3}{3}} \Phi^{-1}\left(\frac{F_t e^{r(T-\tau')}}{G}\right) - (T-t)W(t)}{\sqrt{\frac{(T-\tau)^3}{3}}}, \tag{2.53}$$

and

$$x^* = \left(\Phi^{-1}(R) - \zeta(t, \tau, \tau')\right) \sqrt{\frac{(T-\tau)^3}{3}}. \tag{2.54}$$

The time-t risk neutral price of a European call option with a strike price K and a maturity time $\tau \in (t, T)$, written on an allowance futures contract to be settled at time $\tau' \in (\tau, T)$, is given by:

$$\begin{aligned} &C_F(t, K, \tau, \tau') \\ &= G e^{-r(T-t+\tau-\tau')} \left(\Phi_2 \left(\frac{\mu_x - x^*}{\sigma_x}, \frac{\mu_x \sqrt{\frac{3}{(T-\tau)^3}} + \zeta(t, \tau, \tau')}{\sqrt{1 + \frac{3}{(T-\tau)^3} \sigma_x^2}}, \rho \right) - R\Phi\left(\frac{\mu_x - x^*}{\sigma_x}\right) \right), \end{aligned}$$

where $\zeta(t, \tau, \tau')$, μ_x, σ_x, R, x^*, and ρ are defined in (2.53), (2.46), (2.47), (2.52), (2.54), and (2.50) respectively. The price of the corresponding put option is derived by using the put-call parity condition as:

$$P_F(t, K, \tau, \tau') = C_F(t, K, \tau, \tau') - F(t, \tau') e^{-r(\tau-t)} + K e^{-r(\tau-t)}, \tag{2.55}$$

As observed, under the ABM emission rate model, the values for both allowance options and allowance futures options conveniently have closed-form expressions. Such a feature facilitates implementation and provides tractability in real world applications such as sensitivity tests, scenario analyses, and other risk management practices.

2.3.3 Option prices under a Vasicek emission

In this section we derive the valuation formula of allowance options under the Vesicek emission model. The same approaches and notation in Section 2.3.2 are adopted. Given $\mathcal{F}_t$, $0 \leq t < \tau \leq T$, (2.40) leads to:

$$S(\tau) = e^{-r(T-\tau)} G\Phi \left(\frac{\sigma e^{\kappa(\tau-t)} \sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}} \Phi^{-1}\left(\frac{S_t e^{r(T-t)}}{G}\right) + \nu(t,\tau)}{\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-\tau)}\right)}{2\kappa}}} \right). \tag{2.56}$$

The stochastic component $\nu(t,\tau)$ can be rewritten as:

$$\begin{aligned} \nu(t,\tau) = &\left[\left(e^{\kappa(\tau-t)} - 1\right)\left(\alpha - (1-e^{-\kappa T})\theta\right) + \sigma e^{-\kappa T}\left(1-e^{\kappa(\tau-t)}\right)\int_0^t e^{\kappa u} dW(u) \right] \\ &+ X(t,\tau), \end{aligned} \tag{2.57}$$

where

$$X(t,\tau) = \sigma e^{-\kappa T} \int_t^\tau e^{\kappa u} dW(u). \tag{2.58}$$

The distribution of $X(t,\tau)$ can be found using the steps leading to (E.2) in Appendix E by mapping the time period (s,t) in the proof to the period (t,τ) along with some simple algebraic manipulations. This gives the result:

$$X(t,\tau) \mid \mathcal{F}_t \sim N(0,\sigma_x), \tag{2.59}$$

where

$$\sigma_x = \sigma e^{-\kappa(T-\tau)} \sqrt{\frac{\left(1-e^{-2\kappa(\tau-t)}\right)}{2\kappa}}. \tag{2.60}$$

To further simplify the notation, define:

$$\begin{aligned} \zeta(t,\tau) = &\frac{\sigma e^{\kappa(\tau-t)} \sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}} \Phi^{-1}\left(\frac{S_t e^{r(T-t)}}{G}\right) + \left(e^{\kappa(\tau-t)} - 1\right)\left(\alpha - (1-e^{-\kappa T})\theta\right)}{\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-\tau)}\right)}{2\kappa}}} \\ &+ \frac{\sigma e^{-\kappa T}\left(1-e^{\kappa(\tau-t)}\right)\int_0^t e^{\kappa u} dW(u)}{\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-\tau)}\right)}{2\kappa}}}. \end{aligned} \tag{2.61}$$

With these results, (2.56) becomes:

$$S(\tau) = Ge^{-r(T-\tau)}\Phi\left(\sqrt{\frac{2\kappa}{\sigma^2\left(1-e^{-2\kappa(T-\tau)}\right)}}X(t,\tau)+\zeta(t,\tau)\right). \tag{2.62}$$

Given $\mathcal{F}_t$, $\zeta(t,\tau)$ is deterministic while $X(t,\tau)$ follows a normal distribution as shown in (2.59). Proposition 2.5 hence applies to arrive at the time-t allowance call option price under the Vasicek emission model:

$$C(t,K,\tau) = Ge^{-r(T-t)}\left(\Phi_2\left(\frac{-x^*}{\sigma_x}, \frac{\zeta(t,\tau)}{\sqrt{1+\frac{2\kappa\sigma_x^2}{\sigma^2\left(1-e^{-2\kappa(T-\tau)}\right)}}}, \rho\right) - K^*\Phi\left(\frac{-x^*}{\sigma_x}\right)\right), \tag{2.63}$$

where $\zeta(t,\tau)$, σ_x, and K^* are defined in (2.61), (2.60), and Proposition 2.5 respectively,

$$x^* = \left(\Phi^{-1}\left(K^*\right) - \zeta(t,\tau)\right)\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-\tau)}\right)}{2\kappa}}, \tag{2.64}$$

and

$$\rho = \frac{\sigma_x\sqrt{\frac{2\kappa}{\sigma^2\left(1-e^{-2\kappa(T-\tau)}\right)}}}{\sqrt{1+\frac{2\kappa\sigma_x^2}{\sigma^2\left(1-e^{-2\kappa(T-\tau)}\right)}}}. \tag{2.65}$$

The risk neutral price of the put option under the model is found using the put-call parity relationship in (2.51), where the expression for $C(t,K,\tau)$ follows (2.63). The allowance futures option value follows by applying Corollary 2.2 to (2.63). Modifications to the deterministic quantities are made to simplify the result. Let:

$$\begin{aligned}\zeta(t,\tau,\tau') = &\frac{\sigma e^{\kappa(\tau-t)}\sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}}\Phi^{-1}\left(\frac{F_t e^{r(T-\tau')}}{G}\right) + \left(e^{\kappa(\tau-t)}-1\right)\left(\alpha-(1-e^{-\kappa T})\theta\right)}{\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-\tau)}\right)}{2\kappa}}} \\ &+ \frac{\sigma e^{-\kappa T}\left(1-e^{\kappa(\tau-t)}\right)\int_0^t e^{\kappa u}dW(u)}{\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-\tau)}\right)}{2\kappa}}},\end{aligned} \tag{2.66}$$

and

$$x^* = \left(\Phi^{-1}(R) - \zeta(t,\tau,\tau')\right)\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-\tau)}\right)}{2\kappa}}. \tag{2.67}$$

The time-t risk neutral price of a European call option with a strike price K and a maturity time $\tau \in (t,T)$, written on an allowance futures contract to be settled at time $\tau' \in (\tau,T)$, is given by:

$$C_F(t,K,\tau,\tau') = Ge^{-r(T-t+\tau-\tau')}\left(\Phi_2\left(\frac{-x^*}{\sigma_x}, \frac{\zeta(t,\tau,\tau')}{\sqrt{1+\frac{2\kappa\sigma_x^2}{\sigma^2\left(1-e^{-2\kappa(T-\tau)}\right)}}}, \rho\right) - R\Phi\left(\frac{-x^*}{\sigma_x}\right)\right), \tag{2.68}$$

where $\zeta(t,\tau,\tau')$, σ_x, R, x^*, and ρ are defined in (2.66), (2.60), (2.52), (2.67), and (2.65) respectively. The value of the corresponding allowance futures put option is derived through put-call parity in (2.55), where the expression for $C_F(t,K,\tau,\tau')$ follows (2.68).

Therefore, similar to those under the ABM emission rate model, the values for allowance options and allowance futures options under the Vasicek emission model also have closed-form expressions, facilitating the implementation of the results for various actuarial and financial applications.

2.3.4 Option prices under a Vasicek emission rate

The option valuation formula under the Vasicek emission rate model is derived following the same logics as the ones used for the previous two models in Sections (2.3.2) and (2.3.3). Conditioning on $\mathcal{F}_t$, $0 \le t < \tau \le T$, (2.35) gives:

$$S(\tau) = \frac{G}{e^{r(T-\tau)}}\Phi\left(\frac{\sqrt{T-t-\frac{2}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-t)}\right)}\Phi^{-1}\left(\frac{S_t e^{r(T-t)}}{G}\right)+\xi(t,\tau)}{\sqrt{T-\tau-\frac{2}{\kappa}\left(1-e^{-\kappa(T-\tau)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-\tau)}\right)}}\right), \tag{2.69}$$

where

$$\xi(t,\tau) = \int_t^\tau \left(1 - e^{-\kappa(T-u)}\right) dW(u). \tag{2.70}$$

The expression for the time-τ allowance price in (2.69) is already in the form required to apply Proposition 2.5. To simplify the notation, define:

$$\zeta(t,\tau) = \frac{\sqrt{T-t-\frac{2}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-t)}\right)}\Phi^{-1}\left(\frac{S_t e^{r(T-t)}}{G}\right)}{\sqrt{T-\tau-\frac{2}{\kappa}\left(1-e^{-\kappa(T-\tau)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-\tau)}\right)}}, \tag{2.71}$$

and

$$c(\tau) = \frac{1}{\sqrt{T-\tau-\frac{2}{\kappa}\left(1-e^{-\kappa(T-\tau)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-\tau)}\right)}}. \tag{2.72}$$

(2.69) can then be written as:

$$S(\tau) = Ge^{-r(T-\tau)}\Phi\left(c(t,\tau)\xi(t,\tau)+\zeta(t,\tau)\right). \tag{2.73}$$

Clearly, given $\mathcal{F}_t$, both $c(\tau)$ and $\zeta(t,\tau)$ are deterministic, while $\xi(t,\tau)$ is a random variable whose distribution is found by using the steps to prove (2.33) in I with trivial adjustments to the integrating time interval. More specifically, it is shown that:

$$\xi(t,\tau) \mid \mathcal{F}_t \sim N\left(0,\sigma_x\right), \tag{2.74}$$

where

$$\sigma_x = \sqrt{\tau - t - \frac{2e^{-\kappa(T-\tau)}}{\kappa}\left(1-e^{-\kappa(\tau-t)}\right)+\frac{e^{-2\kappa(T-\tau)}}{2\kappa}\left(1-e^{-2\kappa(\tau-t)}\right)} \tag{2.75}$$

Proposition 2.5 therefore applies to arrive at the time-t allowance option price under the Vasicek emission rate model:

$$C(t,K,\tau) = Ge^{-r(T-t)}\left(\Phi_2\left(\frac{-x^*}{\sigma_x}, \frac{\zeta(t,\tau)}{\sqrt{1+c(\tau)^2\sigma_x^2}}, \rho\right) - K^*\Phi\left(\frac{-x^*}{\sigma_x}\right)\right), \tag{2.76}$$

where $\zeta(t,\tau)$, $c(\tau)$, σ_x, and K^* are defined in (2.71), (2.72), (2.75), and Proposition 2.5 respectively,

$$x^* = \frac{\Phi^{-1}(K^*) - \zeta(t,\tau)}{c(\tau)}, \tag{2.77}$$

and

$$\rho = \frac{c(\tau)\sigma_x}{\sqrt{1 + c(\tau)^2\sigma_x^2}}. \tag{2.78}$$

The risk neutral price of the put option under the model is found using the put-call parity relationship in (2.51), where $C(t, K, \tau)$ follows (2.76). In addition, observe that the option price expression is much more complex than that under the Vasicek emission model in the previous section. This is a natural result of having a Vasicek emission rate specification, which leads to an emission amount that follows an Integrated Ornstein-Uhlenbeck process in (2.32).

The allowance futures option value follows by applying Corollary 2.2 to (2.76). Modifications to the deterministic quantities are made to simplify the result. Define:

$$\zeta(t, \tau, \tau') = \frac{\sqrt{T - t - \frac{2}{\kappa}\left(1 - e^{-\kappa(T-t)}\right) + \frac{1}{2\kappa}\left(1 - e^{-2\kappa(T-t)}\right)}\Phi^{-1}\left(\frac{F_t e^{r(T-\tau')}}{G}\right)}{\sqrt{T - \tau - \frac{2}{\kappa}\left(1 - e^{-\kappa(T-\tau)}\right) + \frac{1}{2\kappa}\left(1 - e^{-2\kappa(T-\tau)}\right)}}, \tag{2.79}$$

and

$$x^* = \frac{\Phi^{-1}(R) - \zeta(t, \tau, \tau')}{c(\tau)}. \tag{2.80}$$

The time-t risk neutral price of a European call option with a strike price K and a maturity time $\tau \in (t, T)$, written on an allowance futures contract to be settled at time $\tau' \in (\tau, T)$, is given by:

$$C_F(t, K, \tau, \tau') = Ge^{-r(T-t+\tau-\tau')}\left(\Phi_2\left(\frac{-x^*}{\sigma_x}, \frac{\zeta(t, \tau, \tau')}{\sqrt{1 + c(\tau)^2\sigma_x^2}}, \rho\right) - R\Phi\left(\frac{-x^*}{\sigma_x}\right)\right), \tag{2.81}$$

where $\zeta(t, \tau, \tau')$, $c(\tau)$, σ_x, R, x^*, and ρ are defined in (2.79), (2.72), (2.75), (2.52), (2.80) and (2.78) respectively. The value of the corresponding allowance futures put option is derived through the put-call parity relationship given in (2.55) with $C_F(t, K, \tau, \tau')$ defined by (2.81).

Similar to the previous two models, the values for allowance options and allowance futures options under the Vasicek emission rate model have closed-form expressions for the ease of implementations. Numerical implementations and applications of the models will be demonstrated in Section 2.5, where model reliability and fitness will be analyzed.

2.4 Market and information completeness

Market completeness is an ideal property that sets the basis of most derivative valuation models, including the classic Black-Scholes option valuation framework. In fact, the strict definition of market completeness requires that every derivative has a replicating portfolio (or equivalently, all risks are hedgeable), which can rarely (if possible at all) be met in practice. There is a volume of studies presenting empirical evidence of stock market incompleteness as well as solutions to stock option valuation and hedging under incomplete markets, such as Davis (1997)[33], Bellamy and M. Jeanblanc (2000)[5], and Carr et al. (2001)[23]. These findings, however, cannot be easily carried over to emission allowance valuation under a closed trading phase where structural models are necessary. Existing literature on allowance valuation provides limited discussion on market completeness. Instead, market completeness remains an unvalidated assumption underlying most of the frameworks where risk neutral valuations are applied. Two recent pioneering studies on incomplete-market models for allowance valuation are Cetin and Verschure (2008)[25] and Mnif (2012)[75]. In the former study, the allowance price is assumed to follow a continuous-time regime-switching model, where the Markov process is driven by a binary variable representing the market's net emission position (e.g. short or long in allowances). In the latter study, the allowance price dynamics are captured using binomial trees and other diffusion processes. Both approaches adopt reduced-form models. We contribute to this area by providing insights into information and market incompleteness under the context of the structural models presented, supplemented with empirical evidence.

To identify the possible sources of market incompleteness, notice that unlike the stock market where many exotic derivatives are actively traded, the allowance market consists of only fours major segments: the spot allowance market, the allowance futures market, the allowance option market, and the allowance futures option market. While the allowance option market was traded mostly over-the-counter (OTC), the other three are exchange-traded and carry sufficient liquidity for information diffusion and hedging purposes. The option contracts are standardized European style options that can be easily hedged using allowance futures. Hence, aside from transaction frictions, the unheadgeable portion of the risk (and hence market incompleteness) must be based on any price discontinuities or stochastic volatilities. Figure 2.1 below shows the price of December 2007 EUA futures during Phase 1. Indeed, there is a markedly downward price jump at the end of April 2006 from €31 to €14 within 3 trading days, accompanied by a partial recovery around early May of the year. Such a discrete movement clearly falls beyond random motions under diffusion. However, the market's aggregate emission is usually well captured by diffusion

processes where significant jumps are rare, since large emitters are mostly stationary installations in the energy and utility industries that are heavily regulated by operational standards. Discontinuous emission patterns caused by the few small emitters will also likely offset when aggregated at the market level.

It is then reasonable to claim that an allowance price jump is caused by a market factor instead of a structural force from the emission process. In fact, such a phenomenon is purely driven by information incompleteness. Since the true emission information is not known to the public until each compliance filing, for the majority of the trading phase, the values of $y(t)$ and $Q(0,t)$ are estimated by the market participants based on their own experiences and historical allowance prices. Hence, strictly speaking, the emission variables $y(t)$ and $Q(0,t)$ in the models presented in this chapter are replaced by their corresponding market estimates $\widehat{y(t)}$ and $\widehat{Q(0,t)}$ in practice. A public announcement of the true aggregate emission status results in a correction of these estimates, which are in turn reflected in the allowance price. The price jump at the end of April 2006 is not a coincidence: all EU region compliance filings are completed in April, after which the verified emission information is made available to the public. It is an example of an instantaneous price reaction to information release, where knowledge of the market having a generously long position in allowances drove down the allowance futures price significantly as shown in figure 2.1 below.

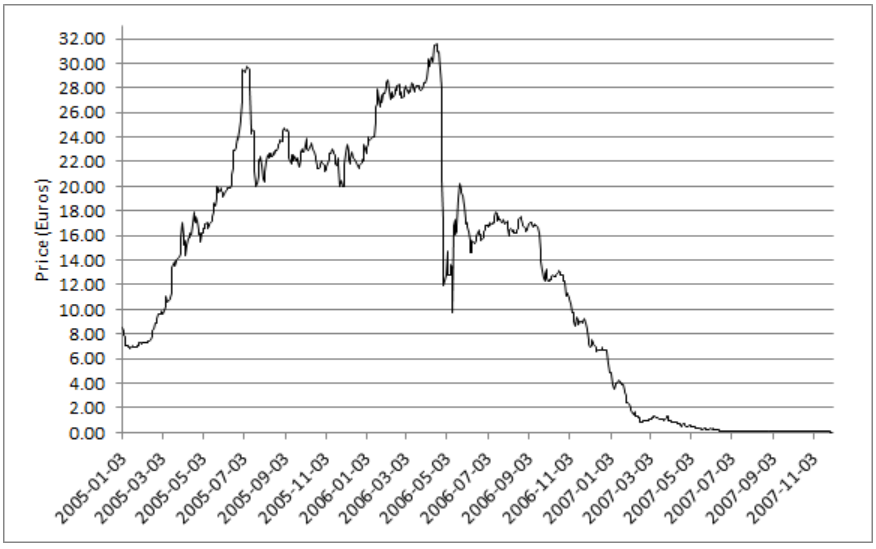


Figure 2.1: Price of December 2007 EUA futures during Phase 1

In addition, notice that reactive price jumps occur only when the released information differs from market expectations. As shown in Figure 2.1, no observable price jump takes place following the April 2007 release as it simply confirms the market's knowledge of an allowance surplus initially conveyed by the April 2006 release. The plummeted price drifts steadily toward zero after the April 2007 information release confirming the inherited allowance surplus status.

Finally, the emission information released annually is a snapshot of the market participants' verified emissions amounts and hence does not reveal the market's emission paths. The lack of data describing the actual aggregate emission dynamics (e.g. the realized values of $y(t)$ and $Q(0,t)$) on an ex-ante as well as ex-post basis is a critical barrier to model calibration. This represents another fundamental difference between the emission allowance market under a closed trading phase and the stock market for which reduced-form models carry risk neutral parameters that can be easily calibrated to the quoted prices in the spot, futures, and options markets. Such an obstacle is expected to be encountered if we pair these two markets and contrast their tiers as contingent claims. Recall that an allowance is essentially an option on the future emission amounts while a stock is often viewed as a fundamental asset. Hence, a calibration to the model allowance price, say for example, in Proposition 2.2, is analogous to an attempt to calibrate a stock option valuation model to market option prices without having an access to the underlying stock price data.

Therefore, strictly speaking, the allowance market in EU ETS Phase 1 is incomplete due to information incompleteness. While such findings inevitably heighten skepticism on the scope of application of the valuation framework, they should be viewed under current and prospective standards of emission trading. Since the end of EU ETS Phase 1 in 2007, there have been rapid technology advancements facilitating reliable and frequent emission monitoring[6], accompanied by proposals of closer audit schedules and stricter supervisory standards by the EU ETS directive board. Emerging emission control schemes are also being developed globally, for which the initial trading phases will most likely be closed to facilitate the optimization of regulatory parameters. Therefore, a closed-phase allowance market with a more optimal information reporting can be expected, under which the models can be confidently applied as presented for an allowance valuation purpose. The applications of the framework under the current scheme are discussed next.

[6]The leading technology in this aspect is Stack Monitoring, which allows for a high-frequency collection and reporting of emission data. The detailed functionality of such a technology is, unfortunately, beyond the scope of this thesis.

2.5 Applications and numerical results

In this section, we discuss the applications of the closed phase allowance valuation framework and models presented in this chapter, followed by numerical demonstrations.

2.5.1 Model applications

Despite the critical implications of information incompleteness, the models presented in Sections 2.2 to 2.4 remain useful in practice. Before proceeding further, we summarize the key characteristics of the three models summarized in Table 2.1 below, where the headings "ABM-ER", "Vas-E" and "Vas-ER" represent the ABM emission rate, Vasicek emission, and Vasicek emission rate models respectively.

Criteria\Models	ABM-ER	Vas-E	Vas-ER
Closed-form allowance price expression	Yes	Yes	Yes
Closed-form option price expression	Yes	Yes	Yes
Price process contains emission variables	No	No	No
Price process contains emission parameters	No	Yes	Yes

Table 2.1: Summary of model characteristics

Since all three models are associated with closed-form allowance price expressions while the corresponding price processes do not contain emission variables, they can be well implemented to serve the following two applications:

1. Deriving aggregate emission information from market prices
2. Performing simulation studies on allowance price with calibrated parameters

We shall demonstrate the derivation of emission information from market prices. Emission information contained in the market prices are valuable inputs for market participants in their forecasts of future allowance positions and adjustments in emission budgets. By establishing estimates of the market's status and trend of emission, firms can make informed decisions that strike an optimal balance among production, emission abatement, and allowance acquisitions. Numerically, the derivation of emission information is equivalent to a model calibration practice as the parameters are estimated at the same time. Hence, the implementation for this application may be a prerequisite for simulation studies that

require the fitted parameters. At the same time, the underlying logic is consistent with the approach of working under an incomplete market as suggested by Kilander (2007)[64], where the issue of selecting the proper risk neutral measure (no longer unique under the incomplete market) for pricing is circumvented by directly specifying the risk neutral emission process that is assumed to reproduce the market prices. The corresponding $\mathcal{Q}$ measure is defined by the market and substantiated via the calibration process. Despite the complex implications of market incompleteness, the presented models are well-suited for such an application of information discovery in practice.

Three major considerations must be noted for this application. First, as a natural consequence of information incompleteness, the emission information carries the same degree of freedom as the data points, and hence some overfitting is inevitable. Caveats are made here so that the readers should be cautious when applying the proposed algorithms while acknowledging such limitations. Second, the derivation of information requires a prior selection of the emission model from those presented in this chapter, which is a subjective input. Our numerical analysis next should provide some insights into model reliability and fitness, which shall facilitate this decision making. Finally, the derived information at points of allowance price jumps are unreliable, since the price jumps usually translate to cascades in emission variables that are unlikely to occur in reality. Fortunately, jumps are caused by information incompleteness that manifest annually upon the releases of compliance data. This makes the issue self-resolving and leads to very few unreliable points in the derived emission information.

On the other hand, simulation studies under the risk neutral scenarios can be generally used for valuation of exotic allowance derivatives, where the numerical implementation involves discretizing the calibrated allowance price processes from the selected model. This is most easily done for the ABM emission rate model, under which the price process does not contain any emission parameters to estimate. We will not demonstrate this application due to the lack of data to formulate the context.

The implementation algorithms and numerical studies for our proposed application constitute a major contribution of our research to the literature. To the best of our knowledge, most studies in the existing literature do not make much effort in calibrating an allowance valuation model to actual price data. Instead, it is often replaced by illustrative simulations with subjectively specified parameters. Moreover, almost no numerical studies are conducted using real ETS data for structural models. The results presented in this chapter represent not only an important step forward in filling this gap, but also provide useful insights in the choice of emission models and exemplify the implementation of structural

models that are easily applicable in many platforms.

2.5.2 Data and methodology

The data used in the implementation of the presented models for an emission information derivation consists of the EU ETS Phase 1 daily closing prices for three different allowance futures contracts with maturities in the December of 2005, 2006, and 2007 respectively[7]. For ease of references, these futures are labeled as EUA 2005, EUA 2006, and EUA 2007 respectively. The data are retrieved from the online database of European Environmental Agency. Aside from the high liquidity in the futures market, the allowance futures prices are used in place of allowance prices due to the scarcity of reliable records for the spot market. Summary statistics of the price data are provided in table 2.2.

Statistic/Futures	EUA 2005	EUA 2006	EUA 2007
Number of points	237	494	749
Mean	17.89	18.36	12.74
Median	19.77	18.12	14.07
Min	6.68	6.75	0.03
Max	29.10	30.50	31.58
Std. Dev.	6.13	6.20	9.86

Table 2.2: Summary statistics of the Phase 1 EUA futures price data

Observe that EUA 2007 has a lower mean and median but a higher standard deviation in closing price. These are largely caused by two factors: the downward price shock in late April 2006 from the information release of the allowance surplus, and the price converging to zero toward the phase end following a confirmation of the continued surplus status in the April 2007 release. In contrast, the EUA 2005 matures in December 2015 and hence its price is not affected by either of the information announcements.

While the algorithms for the derivation of the emission information (and hence the calibration of model parameters) differ among the three models as detailed in Sections 2.6.3

[7]The respective settlements dates for these contracts are 2005/12/01, 2006/12/01, and 2007/12/03.

to 2.6.6 that follow, the common steps below are intended to initialize the implementation of this application:

1. Specify the values of the exogenous variables and the risk free rate using the prevailing economic information for the trading phase.

2. Partition the trading phase into daily intervals following the trading day schedule.

3. Choose the estimation set of data. Without loss of generality, denote the number of data points chosen for the three futures contracts (from longest to shorted maturity) by n_1, n_2, and n_3, where $n_1 > n_2 > n_3$ by construction.

Following the Phase 1 directives, the penalty level G is €40 per excess ton of emission, while the total allowance grant over the three years, α, is 6321 million (tons)[1]. These are universal parameters that apply to all structural models for allowance valuation. The risk free rate parameter r is set to 2%, which is approximately the average European Union Central Bank deposit rate over the three-year period[4]. Such an exogenous input from economic information often performs better as the sample-estimated rate is vulnerable to bias sources, particularly in this case, where the data are unsmoothed market quotes. We also ignore the term structure of interest rates and assume its differential effects across maturities offsets each other overall.

We partition the entire trading phase $(0, 3)$ into daily intervals corresponding to the trading days in the data records. There are 256 days in 2005, 255 days in 2006, and 2007 respectively. Hence, with a total number of points $N = 766$, we obtain the timeline partition (in units of years) below:

$$\mathcal{T} = (t_0, t_1, t_2, ...t_N) := \left(0, \frac{1}{256}, \frac{2}{256}, ..., 1, 1\frac{1}{255}, 1\frac{2}{255}, ..., 2, 2\frac{1}{255}, 2\frac{2}{255}, ..., 3\right).$$

To provide a comprehensive analysis of the results, all price data are used in the derivation of emission information (and hence the calibration of model parameters). Therefore, $n_1 = 766$, $n_2 = 494$, and $n_3 = 237$ in this instance. In practice, the choice of the estimation set should be based on the availability and quality of data, where points of information releases should be ideally avoided.

By substituting in the derived information and the parameter estimates, the fitted allowance prices are calculated and compared to the actual prices, upon which the mean

absolute percentage errors are calculated. This is an in-sample test that assesses the reliability of the derived emission rates and amounts. To assess the accuracy of the derived emission information, the actual emission amount for each of the Phase 1 years is retrieved from the database of the European Environmental Agency (EEA) [1]. An out-of-sample test is subsequently conducted by comparing the derived emission amounts to these actual values, which play no role in the model calibration process. The percentage discrepancies are calculated. Both tests are performed for each of the models, where conclusions are made based on a comparison of the results.

Finally, we will demonstrate an allowance option valuation by showing the model prices of representative allowance futures call options with a series of strike prices on a chosen valuation date. Unfortunately, due to the lack of market option data, we are unable to comparatively assess the fitness of the models for an option valuation purpose.

2.5.3 Deriving emission information under the ABM emission rate model

This section presents the derivation of emission information from market prices under the ABM emission rate model. The proposed algorithm presented applies under the assumption that the market's aggregate emission rate follows the ABM process described in Section 2.2.1. As the derivation of emission information and model calibration occur simultaneously, the parameter estimates are also presented. The data used for the implementation consists of the historical EU ETS Phase 1 futures prices described previously in Section 2.6.2.

Assume that one day is a sufficiently small time differential for all approximations in this section to hold with a reasonable degree of accuracy. Let $F^M(t, \tau')$ denotes the observed time-t market price of an allowance future contract maturing at time τ. In addition, denote the futures contracts' maturity times for EUA 2005, EUA 2006, and EUA 2007 respectively by τ_3', τ_2', and τ_1'. From the data records, we have (in units of years): $\tau_3' = 0.9219$, $\tau_2' = 1.9294$, and $\tau_1' = 2.9294$. (2.16) can be discretized under the partition $\mathcal{T}$ assuming that the following approximation holds:

$$\int_{t_{k-1}}^{t_k} W(s)ds \approx \frac{1}{2}\left(W(t_{k-1}) + W(t_k)\right)\Delta t_k, \quad \forall k \in \{1, 2, ...N\}, \tag{2.82}$$

where

$$\Delta t_k = t_k - t_{k-1}.$$

Using Equation (2.82) and the recursive relationship

$$\int_{t_{k-1}}^{t_k} W(s)ds = \int_0^{t_{k-1}} W(s)ds + \int_{t_{k-1}}^{t_k} W(s)ds,$$

with an initial condition $W(0) = 0$, it can be easily shown that:

$$I(0, t_k) = \int_0^{t_k} W(s)ds \approx \frac{1}{2}W(t_k)\Delta t_k + \frac{1}{2}\sum_{j=1}^{k-1} W(t_j)(\Delta t_j + \Delta t_{j+1}). \tag{2.83}$$

For notational convenience, these approximations are written as equations. Thus, (2.16) is translated to its discretized version so that $\forall t_k < \tau', k \in \{1, 2, ...N\}$:

$$F(t_k, \tau') = e^{-r(T-\tau')}G\Phi\left(\frac{\sqrt{\frac{T^3}{3}}\Phi^{-1}\left(\frac{F_0 e^{r(T-\tau')}}{G}\right) + (T - t_k)\,W(t_k) + I(0, t_k)}{\sqrt{\frac{(T-t_k)^3}{3}}}\right),$$

which is rearranged and simplified to arrive at:

$$W(t_k) = \frac{H(t_k) - \frac{1}{2}\sum_{j=1}^{k-1} W(t_j)(\Delta t_j + \Delta t_{j+1})}{T - t_k + \frac{1}{2}\Delta t_k}, \tag{2.84}$$

where

$$H(t_k) = \sqrt{\frac{(T-t_k)^3}{3}}\Phi^{-1}\left(\frac{F(t_k, \tau')e^{r(T-\tau')}}{G}\right) - \sqrt{\frac{T^3}{3}}\Phi^{-1}\left(\frac{F_0 e^{r(T-\tau')}}{G}\right), \tag{2.85}$$

under the initial condition $W(0) = 0$.

This recursive result allows us to find the Brownian motion values fitted to the market data under the model, which we denote as $\widehat{W(t)}$. The calculation requires replacing the price variable $F(t, \tau')$ in (2.85) by the corresponding market observation $F^M(t, \tau')$ to yield:

$$\widehat{W(t_k)} = \frac{H^M(t_k) - \frac{1}{2}\sum_{j=1}^{k-1} \widehat{W(t_j)}(\Delta t_j + \Delta t_{j+1})}{T - t_k + \frac{1}{2}\Delta t_k}, \quad \forall t_k < \tau', k \in \{1, 2, ...N\}, \tag{2.86}$$

where

$$H^M(t_k) = \sqrt{\frac{(T-t_k)^3}{3}}\Phi^{-1}\left(\frac{F^M(t_k,\tau')e^{r(T-\tau')}}{G}\right) - \sqrt{\frac{T^3}{3}}\Phi^{-1}\left(\frac{F_0^M e^{r(T-\tau')}}{G}\right). \quad (2.87)$$

Once the fitted Brownian motion values are obtained, by (2.6) to (2.8) and (2.83), the emission variables can be expressed as linear functions of the emission parameters:

$$\hat{y(t_k)} = y_0 + \mu t_k + \sigma \hat{W(t_k)} \quad (2.88)$$

$$\widehat{Q(0,t_k)} = y_0 t_k + 0.5\mu {t_k}^2 + \frac{\sigma}{2}\left(\hat{W(t_k)}\Delta t_k + \sum_{j=1}^{k-1} \hat{W(t_j)}(\Delta t_j + \Delta t_{j+1})\right) \quad (2.89)$$

Therefore, after the initialization steps described in Section 2.6.2, the derivation of the emission information follows the algorithm below:

1. Using the selected EUA 2007 data, obtain the fitted Brownian motion values $\widehat{W(t_k)}$ $\forall k \in \{1, 2, ...n_1\}$ following (2.86).

2. Specify a prior parameter space containing plausible values for the unknown parameters y_0, μ, and σ. This may be based on historical experience and knowledge of the aggregate emission process. Add all non-trivial conditions specified to the optimization problem in the next step as additional constraints.

3. Search for the optimal parameter set that minimizes the Mean Absolute Percentage Error (MAPE) between the actual prices and model prices. More formally, the following constrained minimization problem is solved:

$$\begin{aligned}
\underset{y_0,\mu,\sigma}{\text{minimize}} \quad & \sum_{i=1}^{3}\sum_{k=1}^{n_i}\left|\frac{F(t_k,\tau_i') - F^M(t_k,\tau_i')}{F^M(t_k,\tau_i')}\right| \\
\text{subject to} \quad & F(t_k,\tau_i') = e^{-r(T-\tau_i')}G\Phi\left(\frac{\beta(t_k,T)-\alpha(t_k)}{\sigma\sqrt{\frac{(T-t_k)^3}{3}}}\right), \quad i \in \{1,2,3\} \\
& \beta(t_k,T) = \left(y_0 + \mu t_k + \sigma\hat{W(t_k)}\right)(T-t_k) + 0.5\mu(T^2 - {t_k}^2) - \mu t_k(T-t_k), \\
& \alpha(t_k) = \alpha - y_0 t_k - 0.5\mu {t_k}^2 - \frac{\sigma}{2}\left(\hat{W(t_k)}\Delta t_k + \sum_{j=1}^{k-1}\hat{W(t_j)}(\Delta t_j + \Delta t_{j+1})\right), \\
& \sigma > 0, \\
& y_0 > 0.
\end{aligned}$$

4. Calculate the emission values $\widehat{y(t_k)}$ and $\widehat{Q(0,t_k)}$ using Equations (2.88) and (2.89) for $k \in \{1, 2, ...n_1\}$, where the parameters are replaced by their respective estimates in the previous step.

The complexity of the objective function and constraints in the underlying optimization problem may render many optimization algorithms inefficient, particularly when competing local minima are present. In this demonstration, parameter estimates are calculated by performing a global pattern search that guarantees a reasonably accurate solution within a specified plausible parameter space. The performance of particular optimization algorithms on the given allowance futures data set is beyond the scope of our research. The resultant parameter estimates are summarized in Table 2.3.

Parameter	Estimate
y_0	1990.90 Mton/year
μ	48.61 Mton/year
σ	54.62 Mton/year

Table 2.3: Parameter estimates of the ABM emission rate model

The initial emission rate estimate is 1990.90, which is below one-third of the aggregate allowance grant for the three-year trading phase, reflecting a relatively conservative market level emission budget. The volatility estimate is 54.27 million tons/year of CO_2 equivalence, which is a comparatively low value. Hence we expect the derived emission rate to display moderate variations while the cumulative emission increases at a stable rate over time. Figures 2.2 and 2.3 show the derived information from the market prices.

Several key points can be drawn from the observation. First, as emphasized in Section 2.6.1, the derived emission information at price jump points are not reliable values. Although the price plummet in April 2006 is due to the market's reaction to the public information release of an unanticipated allowance surplus, the algorithm is given no input of this exogenous event and interprets the price driver as a structural downward shock in the emission process, leading to the valley observed in $\widehat{y(t)}$ for the end of April in Figure 2.2. This crucial shortcoming of the model and algorithm is in fact common when deriving information parametrically at points where exogenous forces apply. In addition, $\widehat{y(t)}$ increases starting early 2007, which corresponds to a period of decreasing allowance (futures) prices. This interesting phenomenon may be interpreted as a proxy to the market's accelerated response to the confirmed allowance surplus status toward the end of the trading phase, where the firms are incentivized to raise daily emissions (through an increased

production or reduced abatements) following a decrease in allowance prices. Finally, the derived accumulated emission $\widehat{Q(t)}$ rises steadily over time. This is consistent with our previous expectation and resonates with the argument that individual emission noises have an immaterial effect at the aggregate level.

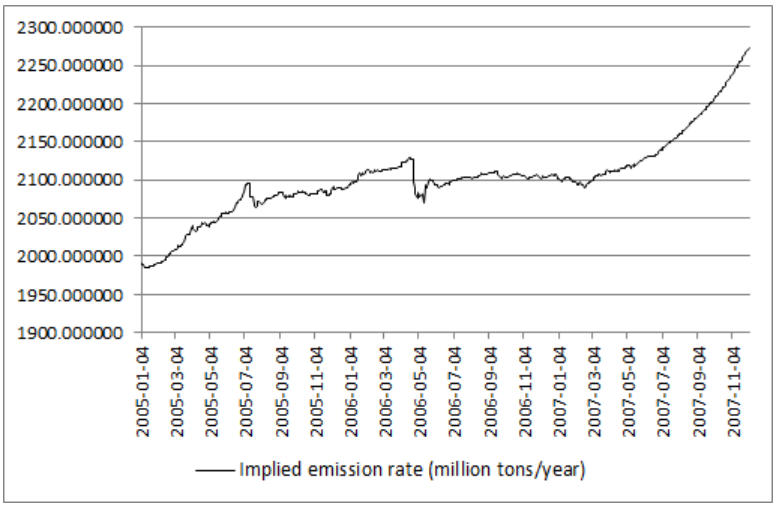


Figure 2.2: ABM emission rate model: derived emission rate $\widehat{y(t)}$

2.5.4 Deriving emission information under the Vasicek emission model

This section demonstrates the derivation of emission information from market prices under the Vasicek emission model. The proposed algorithm applies under the assumption that the market's aggregate emission follows the Vasicek process described in Section 2.2.2. Unfortunately, the algorithm in the previous section for deriving the emission information under the ABM model can not be carried over for an application in this new context. This is because (2.29), the counterpart to (2.16), contains the emission parameters, thus preventing the calculation of the fitted Brownian motion values directly from the price data. Therefore, a few important modifications must be made to the algorithm.

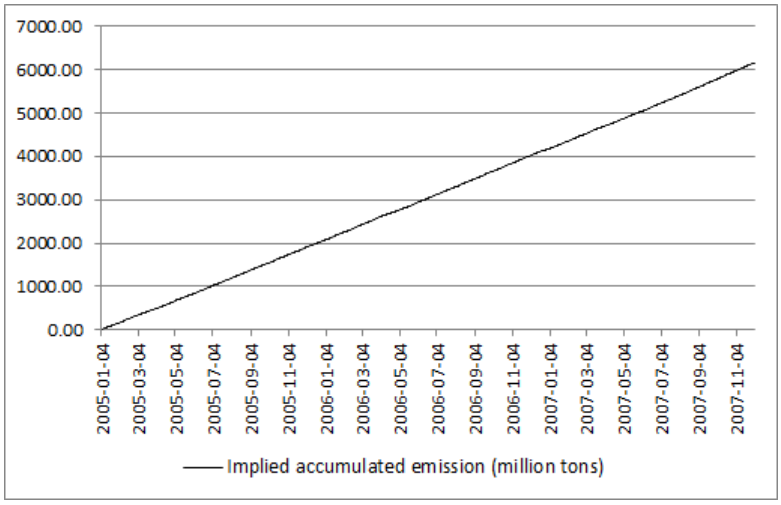


Figure 2.3: ABM emission rate model: derived emission amount $\widehat{Q(0,t)}$

Assume that one day is a sufficiently small time differential for all approximations in this section to hold with a reasonable degree of accuracy. The same set of notations in the previous section are used: $F^M(t,\tau)$ denotes the observed time-t market price of an allowance future contract maturing at time τ'. τ_3', τ_2', and τ_1' denote respectively the settlement times (in units of years) for the EUA 2005, EUA 2006, and EUA 2007 contracts, where $\tau_3' = 0.9219$, $\tau_2' = 1.9294$, and $\tau_1' = 2.9294$ based on the data. We adopt the following approximations $\forall k \in \{0, 1, 2, ...N\}$:

$$dF(t_k,\tau') = \Delta F(t_k,\tau') \approx F(t_{k+1},\tau') - F(t_k,\tau'), \tag{2.90}$$

$$dW(t_k) \approx \Delta W(t_k) = W(t_{k+1}) - W(t_k), \tag{2.91}$$

$$\int_{t_k}^{t_{k+1}} f(s)dW(s) \approx \frac{f(t_k)+f(t_{k+1})}{2}(W(t_{k+1}) - W(t_k)), \tag{2.92}$$

where $f(t)$ can be any continuous deterministic function. For notational convenience, these approximations are treated as if they were equations for the rest of the discussion. (2.28) is hence discretized and rearranged, so that $\forall t_k < \tau'$, $t_k \in \mathcal{T}$.

$$\Delta F(t_k,\tau')\frac{\sqrt{\pi\left(e^{2\kappa(T-t_k)}-1\right)}}{G\sqrt{\kappa}}e^{\left(r(T-\tau')+0.5\Phi^{-1}\left(\frac{F(t_k,\tau')e^{r(T-\tau')}}{G}\right)^2\right)} = \Delta W(t_k). \tag{2.93}$$

The reversion rate parameter κ is estimated by matching the first moments of both sides of (2.93) by using the market futures price data. More specifically, it involves finding the best estimate value of κ, denoted by $\hat{\kappa}$, that satisfies:

$$\sum_{k=1}^{N} \Delta F^M(t_k, \tau') \sqrt{(e^{2\hat{\kappa}(T-t_k)} - 1)} \exp\left(0.5\Phi^{-1}\left(\frac{F_M(t_k, \tau')e^{r(T-\tau')}}{G}\right)^2\right) = 0, \tag{2.94}$$

where $\Delta F^M(t_k, \tau')$ is given in (2.90) with the model prices being replaced by market values.

Once $\hat{\kappa}$ is obtained, the fitted Brownian motion values are calculated $\forall t_k \in \mathcal{T}$:

$$\widehat{W(t_k)} = \sum_{i=0}^{k-1} \widehat{\Delta W(t_i)}, \tag{2.95}$$

where

$$\widehat{\Delta W(t_i)} = \Delta F^M(t_i, \tau') \frac{\sqrt{\pi\left(e^{2\hat{\kappa}(T-t_i)} - 1\right)}}{G\sqrt{\hat{\kappa}}} e^{r(T-\tau') + 0.5\Phi^{-1}\left(\frac{F_M(t_i,\tau')e^{r(T-\tau')}}{G}\right)^2}. \tag{2.96}$$

The expression for the derived emission amount $\widehat{Q(0,t)}$ is found by discretizing (2.32) and substituting into the result all of the estimates obtained in the previous steps:

$$\widehat{Q(0,t_k)} = \theta(1 - e^{-\hat{\kappa}t_k}) + \sigma e^{-\hat{\kappa}t_k} \sum_{i=0}^{k-1} \frac{e^{\hat{\kappa}t_{i+1}} + e^{\hat{\kappa}t_i}}{2} \widehat{\Delta W(t_i)}, \tag{2.97}$$

This leaves two parameters to be estimated, namely θ and σ. Therefore, after the initialization steps described in Section 2.6.2, the emission information is captured by following the algorithm below:

1. Using the selected EUA 2007 data, obtain the reversion rate parameter estimate $\widehat{\kappa}$ and fitted Brownian motion values $\widehat{W(t_k)}$ $\forall k \in \{1, 2, ...n_1\}$ respectively following the moment-matching method described above for (2.94) and the result of (2.95).

2. Specify a prior parameter space containing plausible values for the remaining parameters θ and σ. This may be based on historical experience and knowledge of the aggregate emission process. Add all non-trivial conditions specified to the optimization problem in the next step as additional constraints.

3. Search for the optimal values of θ and σ that minimizes the MAPE between the actual prices and model prices. More formally, the following constrained minimization problem is solved:

$$\begin{aligned}
\underset{\theta,\sigma}{\text{minimize}} \quad & \sum_{i=1}^{3}\sum_{k=1}^{n_i}\left|\frac{F(t_k,\tau_i')-F^M(t_k,\tau_i')}{F^M(t_k,\tau_i')}\right| \\
\text{subject to} \quad & F(t_k,\tau_i') = e^{-r(T-\tau_i')}G\Phi\left(\frac{\beta(t_k,T)-\alpha}{\sigma\sqrt{\frac{\left(1-e^{-2\hat{\kappa}(T-t)}\right)}{2\hat{\kappa}}}}\right), \quad i\in\{1,2,3\} \\
& \beta(t_k,T) = \left(\theta(1-e^{-\hat{\kappa}t_k})+\sigma e^{-\hat{\kappa}t_k}\sum_{i=0}^{k-1}e^{\hat{\kappa}t_i}\widehat{\Delta W(t_i)}\right)e^{-\hat{\kappa}(T-t_k)} \\
& +\theta\left(1-e^{-\hat{\kappa}(T-t_k)}\right), \\
& \sigma>0, \\
& \theta>0.
\end{aligned}$$

4. Calculate the derived emission amount $\widehat{Q(0,t_k)}$ for $k\in\{1,2,...n_1\}$ using (2.97), where the parameters are replaced by their respective estimates.

The optimization problem in step 3 contains more complex, non-linear constraints than that under the ABM emission rate model presented previously. In addition, $\widehat{\kappa}$ in step 1 is usually solved for numerically in practice, whose precision has a material impact on the results of the subsequent steps. Therefore, to guarantee a reasonably accurate solution within a specified plausible parameter space, all parameter estimates in this demonstration are calculated by performing a multi-start pattern search. The performance of particular optimization algorithms on the given allowance futures data set is beyond the scope of our research. The results are summarized in Table 2.4:

Parameter	Estimate
κ	0.5043
θ	7968.34 Mton
σ	142.98 Mton/year

Table 2.4: Parameter estimates of the Vasicek emission model

The reversion rate estimate is around 0.5, representing a relatively slow reversion to the

long term static emission level, which is estimated to be 7968 million tons. The volatility estimate is 142.98 million tons per year, implying moderate fluctuations in the emission amount over time. Hence, by construct, it can be expected that the derived accumulated emission under the calibrated model follows an increasing and concave curve over the trading period. Figure 2.4 shows the fitted $Q(0,t)$ values throughout the period, which confirms such an expectation.

Two other important observations can also be made from Figure 2.4, which are consistent with the key points concluded in the previous section from the derived emission information under the ABM emission rate model. First, the downward price jump at the end of April 2006 is again translated into a structural drop in the underlying emission amount, signaling the unreliability of the derived emission values at information release points where markets may be incomplete. Secondly, the derived emission amount grows at a very slow rate through the trading period but displayed some rapid increase at the very end of the phase. This is a similar phenomenon as in the $\widehat{y(t)}$ plot in Figure 2.2 and may also be interpreted as a proxy to the market's accelerated response to the confirmed allowance surplus status toward the end of the trading phase.

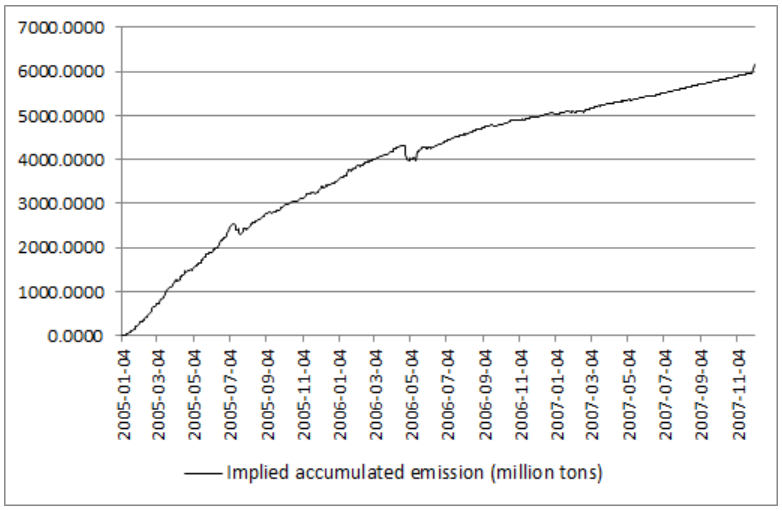


Figure 2.4: Vasicek emission model: derived emission amount $\widehat{Q(0,t)}$

2.5.5 Deriving emission information under the Vasicek emission rate model

This section demonstrates the derivation of emission information from market prices under the Vasicek emission rate model. The proposed algorithm applies under the assumption that the market's aggregate emission rate follows the Vasicek process described in Section 2.2.3. From an implementation perspective, similar to its counterpart under the Vasicek emission model, the futures price expression in (2.37) contains the emissions parameters, preventing the calculation of the fitted Brownian motion values directly from price data, while the futures price dynamics in (2.38) contains only one unknown parameter κ. Therefore, the algorithm for deriving the emission information under the Vasicek emission rate model is methodically the same as that under the Vasicek emission model.

Assume that one day is a sufficiently small time differential for all approximations to hold with a reasonable degree of accuracy. The same set of notations in the previous sections are used: $F^M(t,\tau')$ denotes the observed time-t market price of an allowance future contract maturing at time τ'. τ_3', τ_2', and τ_1' denote respectively the settlement times (in units of years) for the EUA 2005, EUA 2006, and EUA 2007 contracts, where $\tau_3' = 0.9219$, $\tau_2' = 1.9294$, and $\tau_1' = 2.9294$ based on the data. Adopt the same approximations given by (2.90) to (2.92). For notational convenience, these approximations are treated as if they were equations for the rest of the discussion. (2.38) is hence discretized and rearranged so that $\forall t_k < \tau'$, $t_k \in \mathcal{T}$.

$$\frac{\Delta F(t_k,\tau')\sqrt{2\pi\left(T-t-\frac{2}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-t)}\right)\right)}}{G\left(1-e^{-\kappa(T-t)}\right)\exp\left(r-(T-\tau')-0.5\Phi^{-1}\left(\frac{F(t,\tau')e^{r(T-\tau')}}{G}\right)^2\right)} = \Delta W(t_k). \tag{2.98}$$

The reversion rate parameter κ is estimated by matching the first moments of both sides of (2.98) using the market futures price data. More specifically, it involves finding the best estimate value of κ, denoted by $\hat{\kappa}$, that satisfies:

$$\sum_{k=1}^{N}\Delta F(t_k,\tau')\frac{\sqrt{\left(T-t_k-\frac{2}{\hat{\kappa}}\left(1-e^{-\hat{\kappa}(T-t_k)}\right)+\frac{1}{2\hat{\kappa}}\left(1-e^{-2\hat{\kappa}(T-t_k)}\right)\right)}}{\left(1-e^{-\hat{\kappa}(T-t_k)}\right)\exp\left(-0.5\Phi^{-1}\left(\frac{F_M(t_k,\tau')e^{r(T-\tau')}}{G}\right)^2\right)} = 0, \tag{2.99}$$

where $\Delta F^M(t_k, \tau')$ is given in (2.90) with the model prices being replaced by market values.

Once $\hat{\kappa}$ is obtained, the fitted Brownian motion values are calculated $\forall t_k \in \mathcal{T}$ as:

$$\widehat{W(t_k)} = \sum_{i=0}^{k-1} \widehat{\Delta W(t_i)}, \tag{2.100}$$

where

$$\widehat{\Delta W(t_i)} = \frac{\Delta F^M(t_i, \tau') \sqrt{2\pi \left(T - t_i - \frac{2}{\hat{\kappa}} \left(1 - e^{-\hat{\kappa}(T-t_i)}\right) + \frac{1}{2\hat{\kappa}} \left(1 - e^{-2\hat{\kappa}(T-t_i)}\right) \right)}}{G \left(1 - e^{-\hat{\kappa}(T-t_i)}\right) \exp\left(-r(T-\tau') - 0.5\Phi^{-1}\left(\frac{F(t_i,\tau')e^{r(T-\tau')}}{G} \right)^2 \right)} \tag{2.101}$$

The expressions for the derived emission rate $\widehat{y(t)}$ and amount $\widehat{Q(0,t)}$ are found by discretizing (2.31) and (2.32) respectively, and substituting into the results all of the estimates obtained in the previous steps:

$$\widehat{y(t_k)} = y_0 e^{-\hat{\kappa} t_k} + \theta(1 - e^{-\hat{\kappa} t_k}) + \sigma e^{-\hat{\kappa} t_k} \sum_{i=0}^{k-1} \frac{e^{\hat{\kappa} t_{i+1}} + e^{\hat{\kappa} t_i}}{2} \widehat{\Delta W(t_i)}, \tag{2.102}$$

$$\widehat{Q(0,t_k)} = \beta(0,t_k) + \frac{\sigma}{\hat{\kappa}} \sum_{i=0}^{k-1} \frac{2 - e^{-\hat{\kappa}(t_k - t_i)} - e^{-\hat{\kappa}(t_k - t_{i+1})}}{2} \widehat{\Delta W(t_i)}, \tag{2.103}$$

where

$$\beta(0,t_k) = \frac{y_0}{\hat{\kappa}} \left(1 - e^{-\hat{\kappa}(t_k)}\right) + \theta \left(t_k - \frac{1}{\hat{\kappa}} \left(1 - e^{-\hat{\kappa} t_k}\right) \right).$$

This leaves three parameters to be estimated, namely y_0, θ and σ. Therefore, after the initialization steps described in Section 2.6.2, derivation of the emission information follows the algorithm below:

1. Using the selected EUA 2007 data, obtain the reversion rate parameter estimate $\widehat{\kappa}$ and fitted Brownian motion values $\widehat{W(t_k)}$ $\forall k \in \{1, 2, ...n_1\}$ respectively following the moment-matching method described above for (2.98) and the result of (2.99).

2. Specify a prior parameter space containing plausible values for the remaining parameters θ and σ. This may be based on historical experience and knowledge of the aggregate emission process. Add all non-trivial conditions specified to the optimization problem in the next step as additional constraints.

3. Search for the optimal values of θ and σ that minimizes the MAPE between the actual prices and model prices. More formally, the following constrained minimization problem is solved:

$$
\begin{aligned}
\underset{y_0,\theta,\sigma}{\text{minimize}} \quad & \sum_{i=1}^{3}\sum_{k=1}^{n_i}\left|\frac{F(t_k,\tau_i')-F^M(t_k,\tau_i')}{F^M(t_k,\tau_i')}\right| \\
\text{subject to} \quad & F(t_k,\tau_i') = Ge^{-r(T-\tau_i')}\Phi\left(\frac{\beta(t_k,T)-\alpha(t_k)}{\gamma(t_k,T)}\right), \\
& \beta(t_k,T) = \frac{\widehat{y(t_k)}}{\hat{\kappa}}\left(1-e^{-\hat{\kappa}(T-t_k)}\right)+\theta\left(T-t_k-\frac{1}{\hat{\kappa}}\left(1-e^{-\hat{\kappa}(T-t_k)}\right)\right), \\
& \alpha(t_k) = \alpha - \widehat{Q(0,t_k)}, \\
& \gamma(t_k,T) = \frac{\sigma}{\hat{\kappa}}\sqrt{T-t_k-\frac{2}{\hat{\kappa}}\left(1-e^{-\hat{\kappa}(T-t_k)}\right)+\frac{1}{2\hat{\kappa}}\left(1-e^{-2\hat{\kappa}(T-t_k)}\right)}, \\
& \widehat{y(t_k)} \quad \text{follows} \quad (2.102), \\
& \widehat{Q(0,t_k)} \quad \text{follows} \quad (2.103), \\
& \theta > 0, \\
& \sigma > 0, \\
& y_0 > 0.
\end{aligned}
$$

4. Calculate the derived emission rate $\widehat{y(t_k)}$ and amount $\widehat{Q(0,t_k)}$ for $k \in \{1,2,...n_1\}$ using (2.102) and (2.103), where the parameters are replaced by their respective estimates.

The optimization problem in step 3 is non-linear and more complicated than the ones in the algorithms for the ABM emission rate and Vasicek emission models introduced previously. This may contribute to a larger in-sample discrepancy, as presented in the next section. To guarantee a reasonably accurate solution within a specified plausible parameter space, all parameter estimates in this demonstration are calculated by performing a multi-start pattern search. The performance of particular optimization algorithms on the given allowance futures data set is beyond the scope of our research. The results are summarized in Table 2.5.

The reversion rate estimate is around 1.31, which represents relatively rapid reversions to the target rate. The target emission rate is estimated to be 2013.08 million tons/year,

Parameter	Estimate
κ	1.3113
θ	2013.08 Mton/year
y_0	2140.40 Mton/year
σ	236.43

Table 2.5: Parameter estimates of the Vasicek emission model

which is a reasonable budget that is slightly less than one-third of the total emission grant. The volatility estimate is over four times that from the ABM emission rate model. By such a construct, it is expected that $\widehat{y(t)}$ exhibits comparatively frequent fluctuations around the target without clearly identifiable trends. Figures 2.5 and 2.6 show the fitted emission values throughout the period, which confirms our expectation.

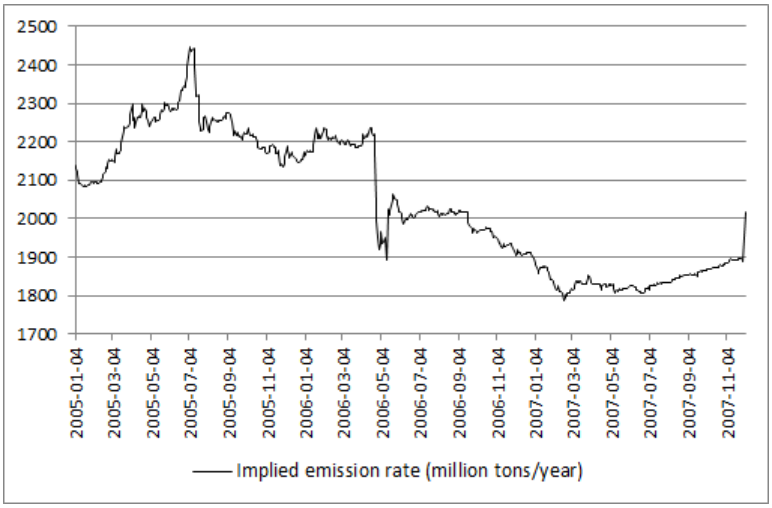


Figure 2.5: Vasicek emission model: implied emission amount $\widehat{y(t)}$

Consistent with the findings for the other two models, the downward price jump at the end of April 2006 is translated into a markedly structural drop in the underlying emission rate, exemplifying again the unreliability of the derived emission values at information release points. The magnitude of the drop is also much larger than that observed under ABM emission rate. In addition, $\widehat{y(t)}$ experiences a sharp increase at the end of the trading phase. Though this may be again interpreted as the market's response to the confirmed allowance surplus as for the result under ABM emission rate, the increase is much more

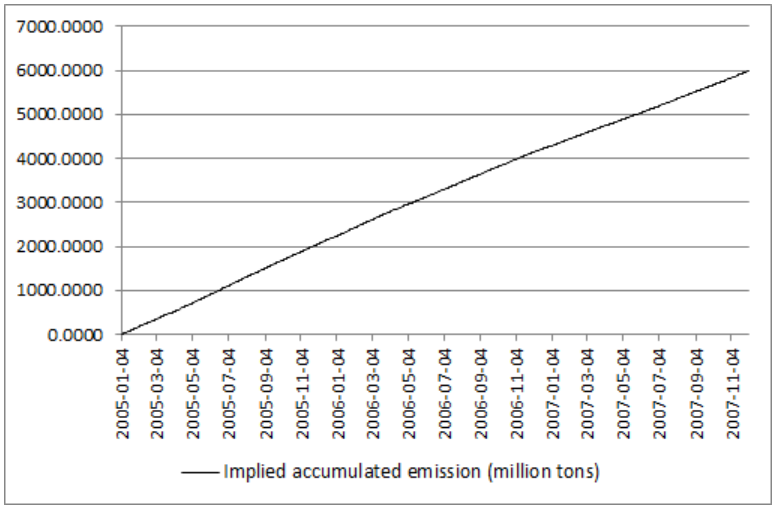


Figure 2.6: Vasicek emission model: implied emission amount $\widehat{y(t)}$

rapid. Therefore, compared to the others, structural changes in the price data translate to more pronounced shifts in the derived emission information under the Vasicek emission rate model.

2.5.6 Comparative analysis

Both in-sample and out-of-sample tests are conducted to compare the three models in their fitness for the demonstrated application of deriving emission information from market data. The in-sample test is conducted by examining the fitness of model prices against the actual using the implementations in Sections 2.5.3 to 2.5.5. The result of this analysis should provide an assessment of the reliability of the derived emission rates and amounts from the models.

Figures 2.7 and 2.8 present the fitting results for EUA 2005 and EUA 2006 under the implementation of the ABM emission rate model. For EUA 2007, there are no graphically observable differences between the actual and model prices and hence a plot is not presented here. We observe that the actual price movements are extremely closely tracked by the model price series. Numerically, the Mean Absolute Percentage Errors (MAPE)

between the actual and model prices is 0.87% and the corresponding Mean Square Errors (MSE) is 0.08. Although being shadowed by the inevitable overfitting effect mentioned previously in our discussion for model implementations, these appealing results are indicative of a high reliability of the derived emission information.

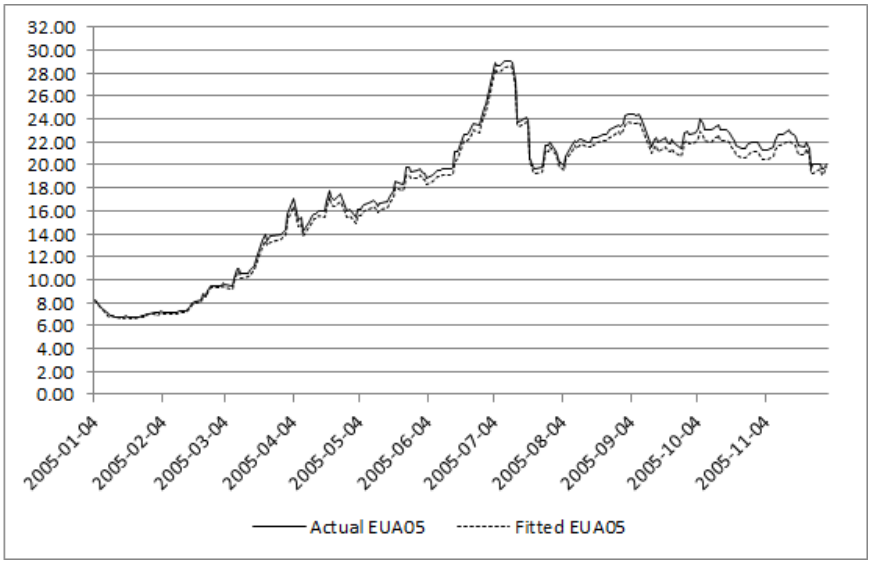


Figure 2.7: ABM emission rate model: actual vs. fitted prices for EUA 2005 futures

Figures 2.9 to 2.11 present the fitting results for the three allowance futures under the implementation of the Vasicek emission model. The actual price movements are reasonably well tracked by the model price series. Numerically, the MAPE between the actual and model prices is 9.84% and the corresponding MSE is 0.13. Both discrepancy metrics are substantially higher than those for the ABM emission rate model, indicating inferior reliability of the derived emission values.

Figures 2.12 to 2.14 present the fitting result for the three allowance futures under the implementation of the Vasicek emission rate model. Unlike the observations for the previous two models, the graphs show noticeable differences between the actual and model prices, which are particularly pronounced for EUA 2006 and EUA 2007 after the April 2006 price drop. Numerically, the MAPE between the actual and model prices is 16.23% and the corresponding MSE is 0.43. These large discrepancy metrics well exceed the ones under the previous two models and can be mostly attributed to the fitting errors after April.

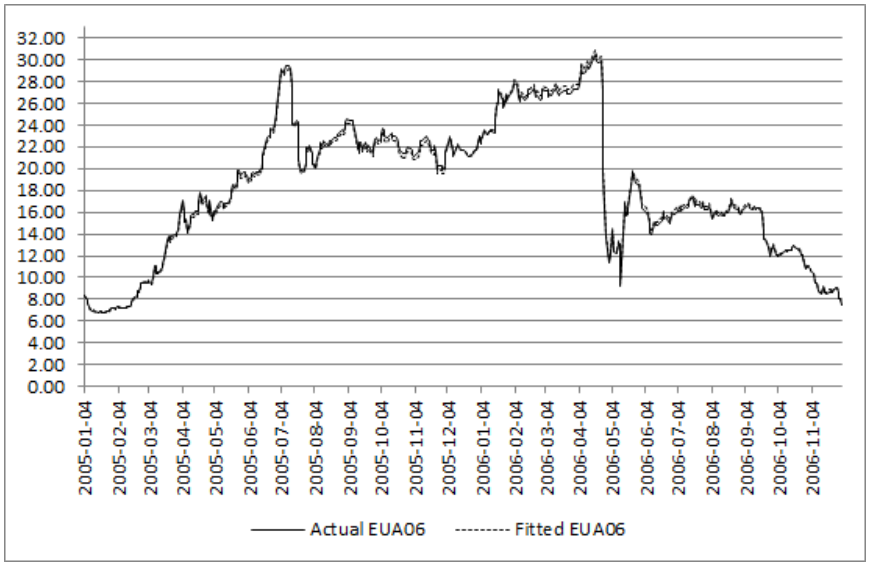


Figure 2.8: ABM emission rate model: actual vs. fitted prices for EUA 2006 futures

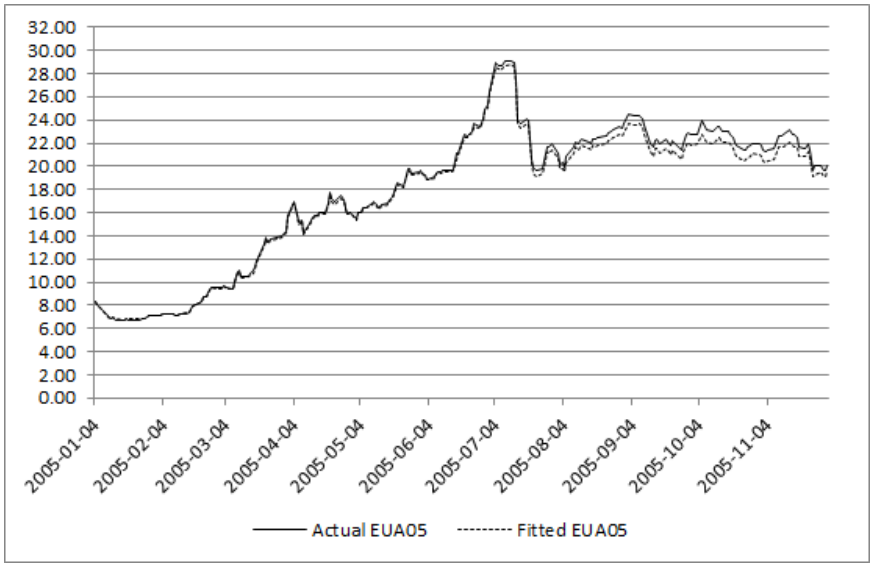


Figure 2.9: Vasicek emission model: actual vs. fitted prices for EUA 2005 futures

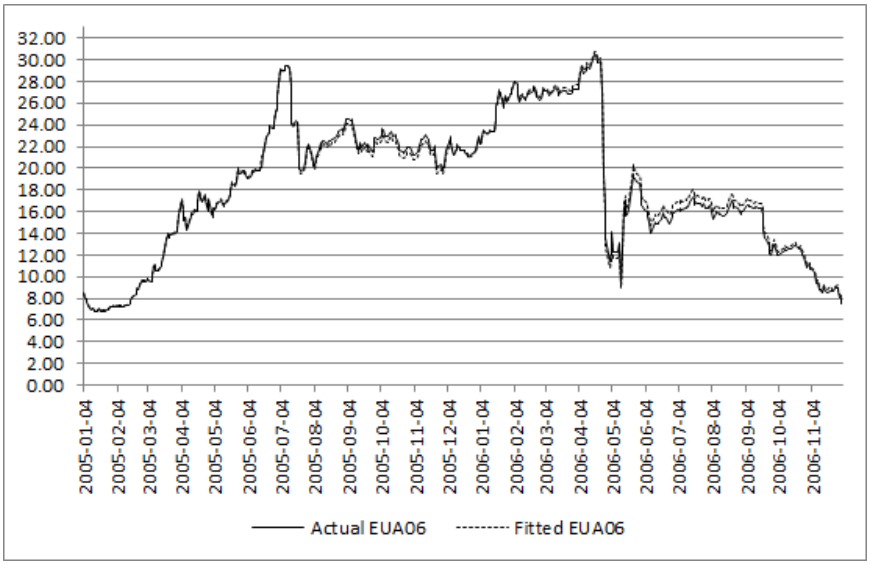


Figure 2.10: Vasicek emission model: actual vs. fitted prices for EUA 2006 futures

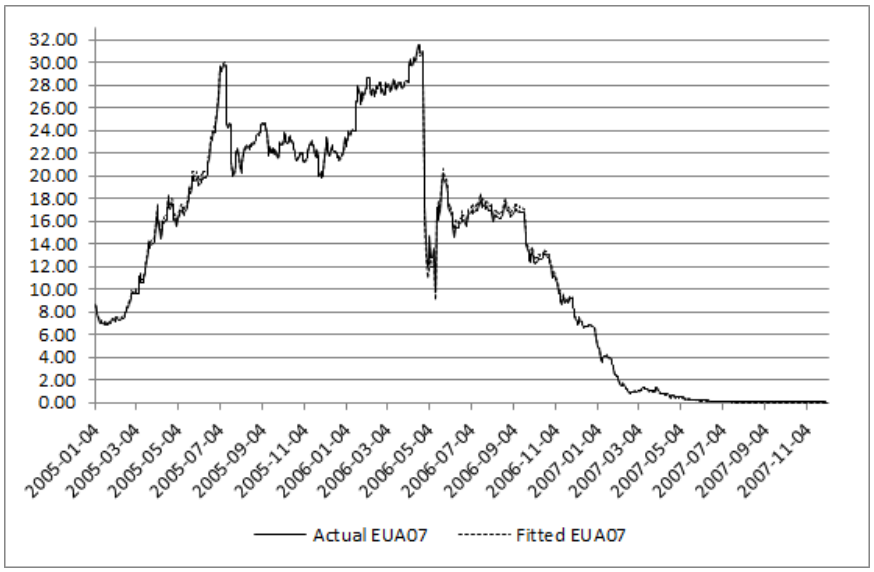


Figure 2.11: Vasicek emission model: actual vs. fitted prices for EUA 2007 futures

Without a robust method of validation, such a phenomenon may be caused by an undue influence of the price jump on the numerically derived emission information and parameter estimates under the complex structural specification of a mean-reverting emission rate. Nevertheless, the derived emission information is likely to be least reliable based on these results.

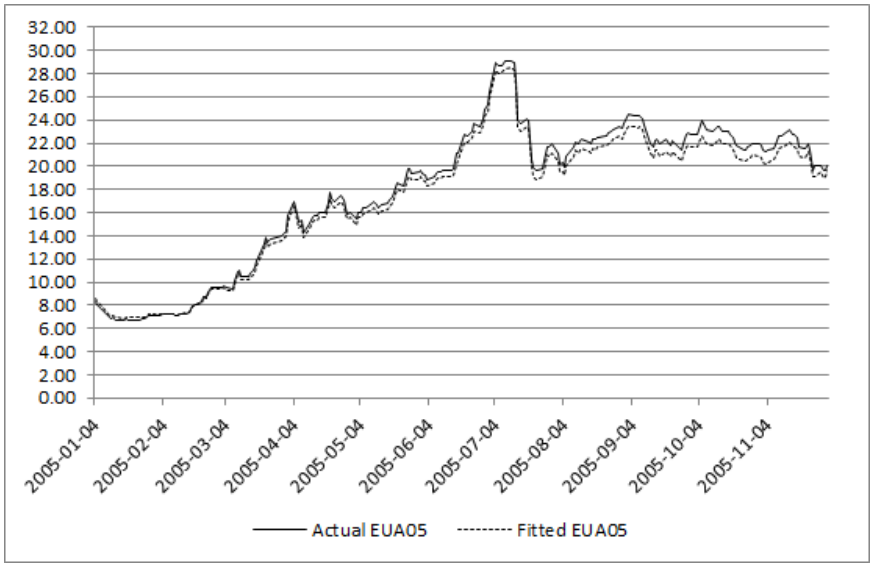


Figure 2.12: Vasicek emission rate model: actual vs. fitted prices for EUA 2005 futures

The out-of-sample test is conducted to assess the accuracy of the derived emission information under the proposed implementations. As described in the Data and Methodology section, this is performed by comparing the emission amounts derived from the demonstrated implementations in Section 2.5.3 to 2.5.5 with the actual reported values. Results are summarized in Table 2.6 for year-over-year emissions and Table 2.7 for accumulated emissions. All values are in million tons of carbon dioxide equivalent emissions. Notice that the emission reporting mechanism leads to only three annual data points for this analysis, which are used in an effort to provide a more comprehensive understanding of the models presented.

The derived emission amounts under the ABM emission rate model are closest to the actual values on both an annual and cumulative basis, which supplements the findings from the in-sample test. Hence, we tentatively conclude that the ABM emission rate model leads

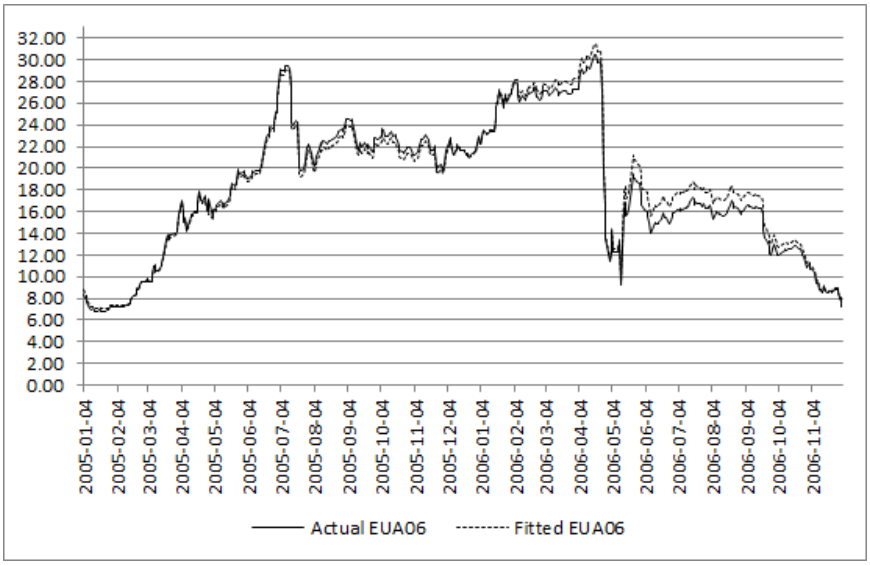


Figure 2.13: Vasicek emission rate model: actual vs. fitted prices for EUA 2006 futures

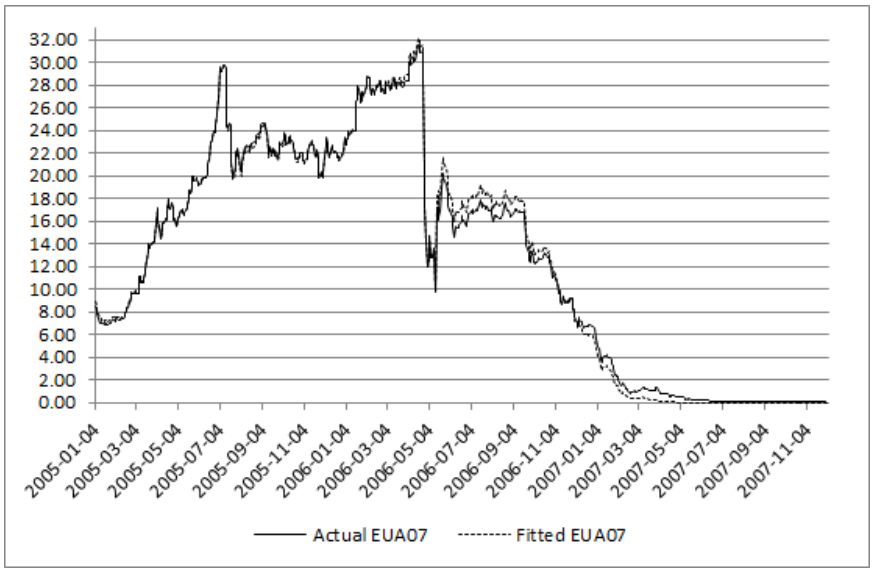


Figure 2.14: Vasicek emission rate model: actual vs. fitted prices for EUA 2007 futures

		ABM-ER		Vas-E		Vas-ER	
Year	Actual	Fitted	Error	Fitted	Error	Fitted	Error
2005	2014.08	2054.54	2.01%	3584.45	77.97%	2218.97	10.17%
2006	2035.79	2104.71	3.39%	1557.94	-23.47%	2051.84	0.79%
2007	2164.73	1999.26	-7.64%	1093.30	-49.49%	1713.08	-20.86%

Table 2.6: Actual vs. derived annual emission amounts

		ABM-ER		Vas-E		Vas-ER	
Year	Actual	Fitted	Error	Fitted	Error	Fitted	Error
2005	2014.08	2054.54	2.01%	3498.20	73.69%	2218.97	10.17%
2006	4049.87	4159.26	2.70%	5056.15	24.85%	4270.81	5.46%
2007	6214.60	6158.51	-0.90%	6149.45	-1.05%	5983.89	-3.71%

Table 2.7: Actual vs. derived accumulated emission amounts

to the most reliable and accurate derived emission information upon implementation using the algorithm in Section 2.5.3. Interestingly, the Vasicek emission rate model substantially outperforms the Vasicek emission model here despite its inferior result from the in-sample test. This is likely to be caused by a structural imposition of the mean-reverting emission amount process under the Vasicek emission model, as evidenced by having an unreasonably large positive error in annual emissions for 2005, followed by substantial negative errors in 2006 and 2007. It also implies that the prior belief in the Vasicek emission model may not be fully justifiable.

Deriving emission information from market data is one application of the models presented in this chapter under the imperfect real world setting. As discussed in Section 2.5.1, when the algorithms are applied in practice at a time point of the trading phase, the parameter estimates obtained can be used to perform simulation studies, which may be used for valuation and risk-management of more complex or customarily-structured emission derivatives that do not carry active market quotes. Finally, we remind the readers that the implementations of the models for the demonstrated application are subject to the caveats identified in Section 2.5.2. as well as the specification of a plausible parameter space a priori.

2.5.7 Option valuation results

Due to the lack of actual option price data for EU ETS Phase 1, we demonstrate option valuation under the three models using a representative call on the EUA 2006 allowance futures. December 1, 2005 is arbitrarily chosen as the valuation date, where the corresponding EUA futures price is €20.35. The option's expiry date is set to October 17, 2006. Using the parameter estimates obtained in Sections 2.5.3 to 2.5.5, the model prices of the option under different strikes are calculated and displayed in Figure 2.15. As expected, the

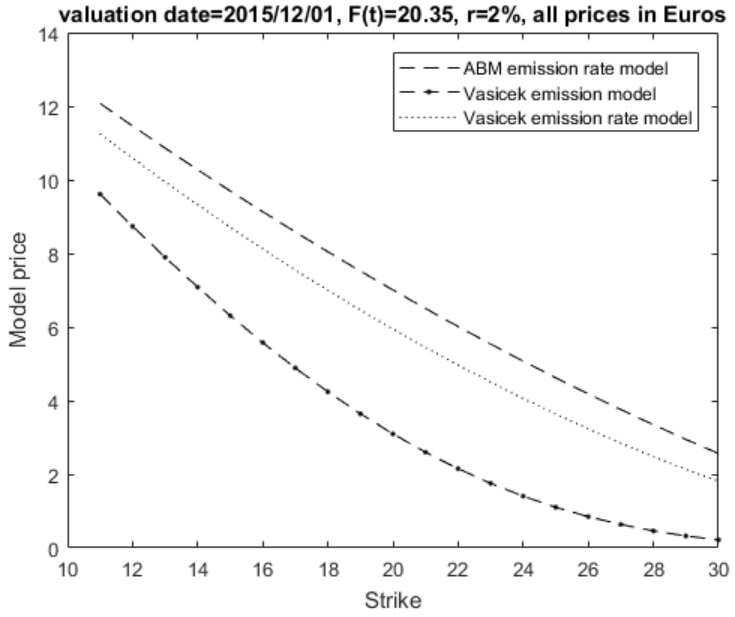


Figure 2.15: Value of a representative EUA 2006 futures call option with varying strikes

call option value is a decreasing convex function of the strike price under all models, which is a necessary condition of an arbitrage-free option market. Specifically, under the ABM emission rate model, the magnitudes of the convexity range from 0.012 to 0.014 across different strikes. Such low values imply the possibility of applying linear approximations and interpolations in the valuation of allowance futures options for different strikes, which can be used in the pricing of options with illiquid strikes. In contrast, the option value under the Vasicek emission model is much more convex with respect to the strike price, which is likely to be the result of the underlying specification of a mean-reverting emission amount. In addition, for this selected series of strike prices, the option values from the

ABM emission rate model remain well above those from the Vasicek, though the differences in convexity may result in cross-overs at very high or low strikes. Finally, the option values under the Vasicek emission rate model lie between those under the other two models, where the shape of the plot is similar to that from the ABM emission rate model but slightly more convex around the at-the-money region.

2.6 Allowance valuation for individual firms

The allowance valuation models introduced previously are based on different specifications of the aggregate emission dynamics and are sensitive to the effect of information incompleteness at the market level. Fortunately, the valuation framework can be readily extended to apply at the firm level, which should provide more accurate results since each firm possesses full knowledge of its own emission process on a business-as-usual basis. Such an extension may be of interest to practitioners specializing in emission advisory.

To motivate this concept, take the position of a representative business or stationary installation governed under the ETS regime, such as a power-plant. The business does not speculatively profit from the price movements in the allowance market, but is more concerned about emission compliance. Hence, the marginal value of each allowance contract equals the expected value of its compliance payoff given the firm's current allowance positions and emission status, discounted at a proper rate. The market price then becomes an exogenous benchmark based on which trading and abatement decisions are made. This setting allows us to eliminate most of the restrictive assumptions in the valuation framework presented previously for implementation in the real world. Without loss of generality, for the rest of the discussion in this section, we work under the real world measure $\mathcal{P}$ for an incomplete market unless otherwise specified.

Consider a representative firm i in a closed trading phase of the governing ETS scheme for emission compliance. At any arbitrary time point $t \in (0, T)$, the value of an allowance to firm i, denoted by $V^i(t)$, may differ from the market price $S(t)$, providing an incentive to trade. This perspective is discussed previously in Section 2.1 in the context of a market price formation. Since the allowance contract is essentially a binary call option over the terminal accumulated emission amount, the marginal value of an allowance to the firm is

given by:

$$V^i(t) = e^{-r^i(T-t)} GPr\left(Q^i(t,T) > A^i(t) - Q^i(0,t)\right), \tag{2.104}$$

where:

1. r^i is the firm's risk-adjusted cost of capital for allowance valuation purposes.
2. $A^i(t)$ is the firm's allowance balance at time t including the future years' grants, which is known at the beginning of the trading phase.
3. $Q^i(0,t)$ denotes the firm's accumulated emission during (s,t) net of emission reduction credits.

This framework can be implemented in three steps. First, the firm's risk-adjusted cost of capital r_i is estimated. Then, based on the firm's historical emission data and future production plans, the appropriate model for the emission process is specified and calibrated. This involves devising, implementing, testing, and selecting the best alternative among various candidate models. Lastly, the value of the allowance is calculated using (2.104) under the chosen model. Each firm may have a unique set of emission characteristics and hence a generalization cannot be made here.

This adapted valuation framework for individual firms leads to incomplete market models as the value of the allowance becomes position-dependent. More specifically, the model estimates the marginal value of each allowance to the firm, which changes immediately as the allowance balance changes. Results of such an application can then be used to identify optimal periods and price ranges for trading the emission allowances in the market.

An illustrative example is provided below using a North American power plant as the representative subject. Data on the plant's daily emission amounts are collected for the period from January 1, 2005 to April 30, 2006, which appear to best fit the ABM emission rate model presented in Section 2.2.2. Using all 2005 emission data as the training set, Maximum Likelihood Estimation (MLE) are performed to obtain the parameter estimates. For illustration, the total allowance grant for the firm over the 3-year trading period and the risk-adjusted discount rate r^i are assumed to take the values of 1300 thousand tons and 2% respectively. The marginal allowance values are calculated for the first four months in 2006 assuming no trading is undertaken by the firm. Results are plotted in conjunction with the spot market prices in Figure 2.16.

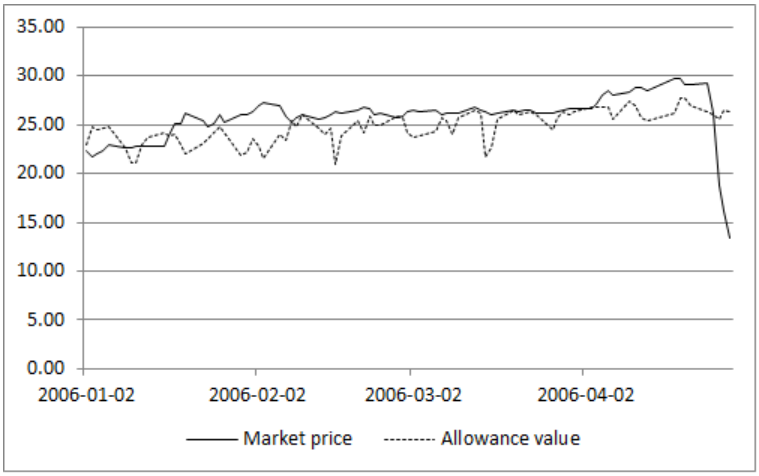


Figure 2.16: Values and market prices of allowance from January 1, 2016 to April 30, 2016

With the hypothetical parameters, several crossovers between the market prices and firm-specific allowance values are observed to occur at the beginning of 2016, which may represent trading opportunities. Subsequently, the allowance value remains below the market price for the majority of the period until the end of April, when the emission surplus status was announced to causes the price plummet. Notice that the allowance value does not follow suit since the firm's emission process shows no structural breaks. This again highlights the issue of information incompleteness discussed previously. Depending on the future development in the emission status, the firm may be advised to acquire additional allowance at low prices for compliance purposes.

In addition to allowance valuation for individual firms, this incomplete-market framework can also be adapted to serve simulation purposes and create firm-specific emission scenarios under the real world measure. While risk neutral scenarios, which may be generated through the simulation studies introduced in Section 2.5.1, are often used for the valuation and pricing of more exotic emission derivatives, the real world scenarios are best deployed to facilitate risk management practices by identifying emission-driven exposures on both the asset and liability side.

In concluding this chapter, we review the scope and relevance of the valuation framework introduced. The structural models presented apply to a closed trading phase where unused allowances at the end of the phase are lapsed, as exemplified by the EU ETS Phase

1 from 2005 to 2007. Despite the historical nature of this referenced phase in our model development, the closed-trading phase remains relevant to date due to its continued deployment and potential adoption by ETS with planned implementations, where our framework, models, and analysis will hopefully provide valuable insights to both businesses and policy makers.

Chapter 3

Models under an open trading phase

In this chapter, we present valuation models of allowance options under the setting of an open trading phase. Contrary to the closed trading phase in the previous chapter, an allowance trading phase is defined to be open if the allowances issued during the phase can be banked without any limit into the subsequent trading phases. More precisely speaking, allowances not surrendered by the end of a phase are automatically converted into the new allowances issued for the subsequent phase via a predetermined mechanism set by regulators. On the other hand, intra-phase banking and borrowing are also permitted so that unused allowances at the end of a compliance year are carried over to future years. Therefore, in an open trading phase, an emission allowance is a non-expiring contract with a positive surrender value, dissociating it from the terminal conditions described in Proposition 2.1 for closed trading phases. Such a lapse-free feature also leads to higher trading volumes and liquidity for open phases. Most of the contemporary emission markets are deploying the open phase specifications, which includes the current EU ETS, Korean ETS, and the Chinese national ETS launched by the world's largest emission producing economy. The analyses in this Chapter are based on the Phase 2 of EU ETS spanning the period of 2008 to 2012, and Phase 3 of EU ETS covering the period of 2013 to 2020 according to the EU ETS directives, both of which are open. Figure 3.1 below shows the daily settlement prices of the December 2012 EUA futures over Phase 2. Clearly, the price fluctuate between 6 to 8 Euros per contract toward the end of the phase, which are far from zero as in the case of Phase 1.

The non-maturing nature of allowances also theoretically eliminates the upper bound of the allowance price, which is otherwise capped at the penalty level set at the beginning of the phase. The allowance price in an open trading phase reflects the market's view of its

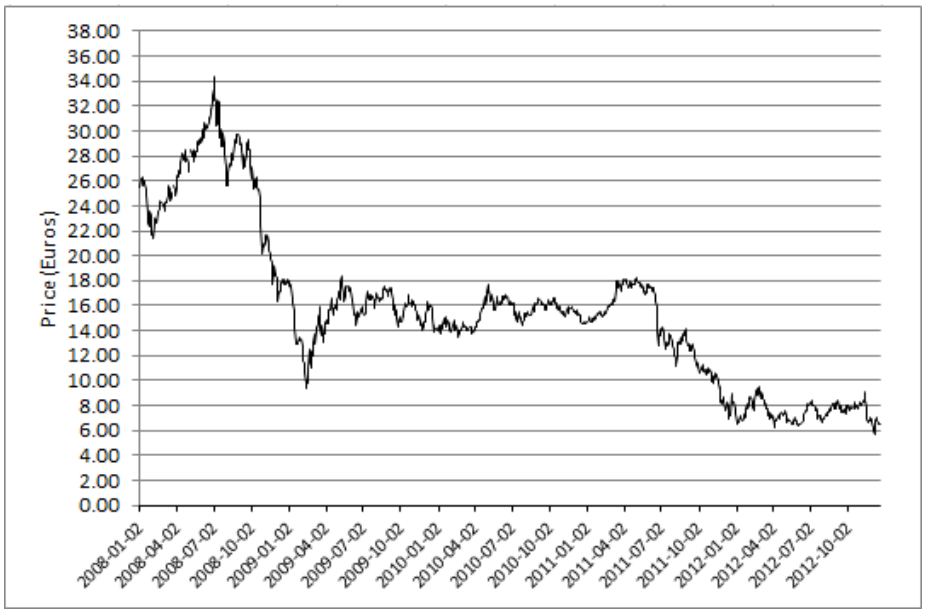


Figure 3.1: Price of December 2012 EUA futures during Phase 2

position in allowance (long or short, as defined in the proof of Proposition 2.1 previously) as well as the expectation of penalty levels applicable in subsequent trading phases, which is set to €100 per metric tonne of excess emission for both Phase 2 and Phase 3.

Finally, an important feature of EU ETS starting from trading Phase 2 (i.e. 2008) is the migration toward an auction market allocation instead of a free distribution of allowance contracts at the beginning of each compliance year. Such a feature is not a defining characteristic of an open trading phase, but is implicitly assumed to apply in this section unless otherwise stated, since the model and analysis are constructed based on the EU ETS directives. The auction revenues are used by the regulators to fund various climate and energy purposes. The migration takes place gradually. Over trading phase 3, the share of allowances auctioned will be higher: it is estimated that up to half of the allowances may be auctioned[1].

As a result, for an open trading phase, the valuation of allowance derivatives can be based

[1]See the European Commission webpage at `https://ec.europa.eu/clima/policies/ets/auctioning_en`

on continuous-time reduced-form models that specify the allowance price process directly as opposed to structural models for the closed trading phase. This avoids the complexity of capturing the auction market price formation. On the other hand, an analogy can be drawn between EU ETS open trading phases and the stock market. The stocks and emission allowances can now be viewed as perpetual instruments with similar market tiers: the auction-based allocation of allowances can be viewed as a primary market dealing or seasoned offering, while the subsequent emission trading activities constitute the secondary market, which is the subject of our study.

We will analyze several valuation models for allowance options and allowance futures options based on continuous-time reduced-form allowance price dynamics. Reduced-form models for emission allowance price have been used in few existing works. However, those models usually contain some structural emission component such as Cetin and Verscheure (2009)[25], or rely on discrete-time approximations such as Mnif(2012) [75]. Moreover, they lack a formalization on the scope of application, where model implementations are mostly based on EU ETS phase 1, for which structural models are appropriate as discussed in the previous chapter. To the best of our knowledge, the formalized use of continuous-time reduced-form models for allowance derivative valuation under open trading phases is rare in the literature. This chapter aims to fill this gap. In summary, our contribution in this chapter consists of three components. First, we perform an empirical study of the stylized facts of allowance prices and allowance option prices, using recent market data from EU ETS Phase 3 that are more relevant to the current economy. Second, we propose several valuation models for allowance options based on continuous-time reduced-form allowance price processes. The proposed models have different characteristics and trade-offs in complexity and fitness as well as the ability to capture stylized facts in market prices, which will be discussed in detail. Finally, we calibrate the models to real market option data and perform comparative analyses on model fitness.

3.1 Stylized facts of allowance prices under an open trading phase

To analyze the stylized facts of emission allowance prices and related derivatives, we use data from EU ETS Phase 3. There are several reasons for this. First, Phase 3 is the most recent trading phase spanning the period of 2012 to 2020, and hence is more relevant to

reach forward-looking conclusions that are meaningful for both academia and practitioners. Second, the models to be presented as well as their numerical demonstrations are based on EU ETS Phase 3 directives and data. Hence, this empirical analysis should cover the same period. Thirdly, compared to the previous two trading phases, the phase 3 market is more mature and stable as evidenced from elevated trading volumes and a harmonized set of market disciplines[2], which are expected to lead to more reliable conclusions. Lastly, the market behavior of trading Phase 3 is rarely studied in the literature, to which our study represents a fresh contribution.

Figure 3.2 shows the daily closing EUA spot prices during the period from January 2014 to July 2017, which comprises a sufficiently large sample covering the majority of the trading phase 3 period to date. The 2013 prices are excluded in the analysis due to data availability and reliability concerns. The daily prices during the period are relatively stable and fluctuate between 4 and 9 Euros per contract. Although rapid price decreases took place during the period from January to March 2016, the overall price movements exhibit no clearly observable jumps, indicating the appropriateness of continuous-time reduced-form models for the price dynamics.

Figure 3.3 shows the daily log returns of the EUA prices for the period, where the corresponding ACF plot is given by Figure 3.4 (a). Similar to a typical stock price return series, the EUA daily returns display both volatility clustering and different volatility regimes, which calls for consideration of more advanced models capturing these volatility variations. On the other hand, the lack of autocorrelations encourages us to start from simple price processes, such as the Geometric Brownian Motion, as a benchmark model for a comparative analysis. Starting from the basic benchmark model allows us to better assess model fitness and hence the tradeoff against model complexity and implementability.

Key statistics of the price returns are summarized in Table 3.1. Notice that though the return data displays zero mean and median, it exhibits very small negative skewness and substantial excess kurtosis as expected from the observed return volatility clustering behavior over the study period. The resultant heavy tail is also shown by the histogram of the returns in Figure 3.4 (b), which nevertheless does not clearly display the asymmetry implied by the negative skewness.

[2]Specifically, starting Phase 3, a single EU-wide emission cap applies in place of the previous national caps, where more sectors and emission gas types are covered by the scheme.

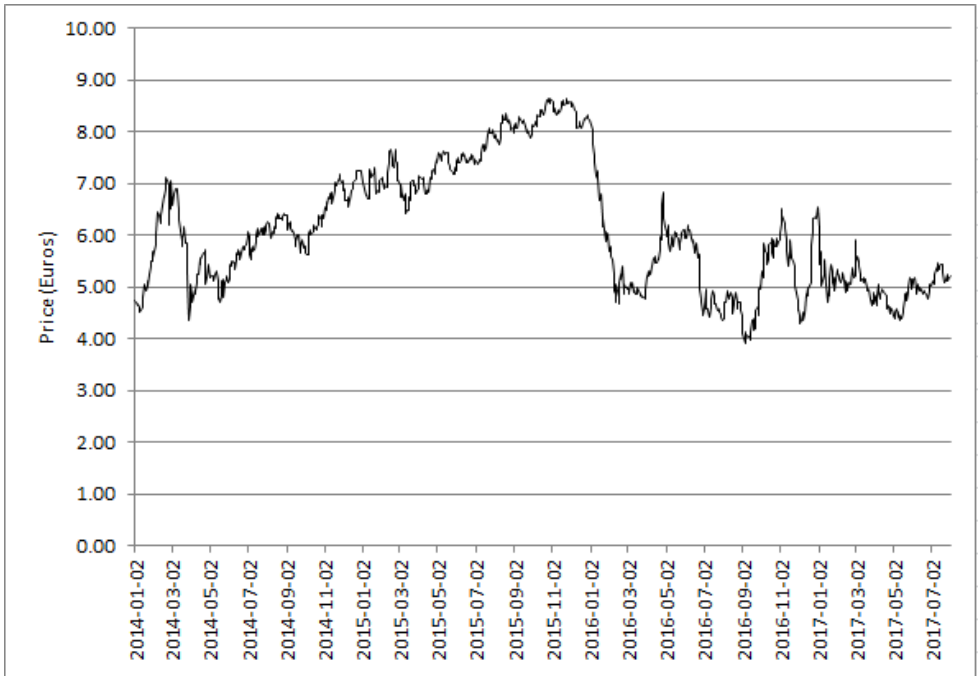


Figure 3.2: Daily closing EUA spot price from January 2014 to July 2017

These empirical stylized facts of daily allowance returns discussed in this section provide

Min	Max	Mean	Median	Variance	Skewness	Kurtosis
-0.1861	0.1263	0.0001	0.0000	0.0008	-0.4597	7.3686

Table 3.1: Summary statistics of EUA returns

insights in our specification of the plausible price process assumptions. They will also be used as qualitative criteria when assessing each model presented in this section, in addition to the fitness of the resultant option valuation models to real market data. Stylized facts of allowance option prices, particularly the volatility smiles, will be defined and discussed in later sections as necessary.

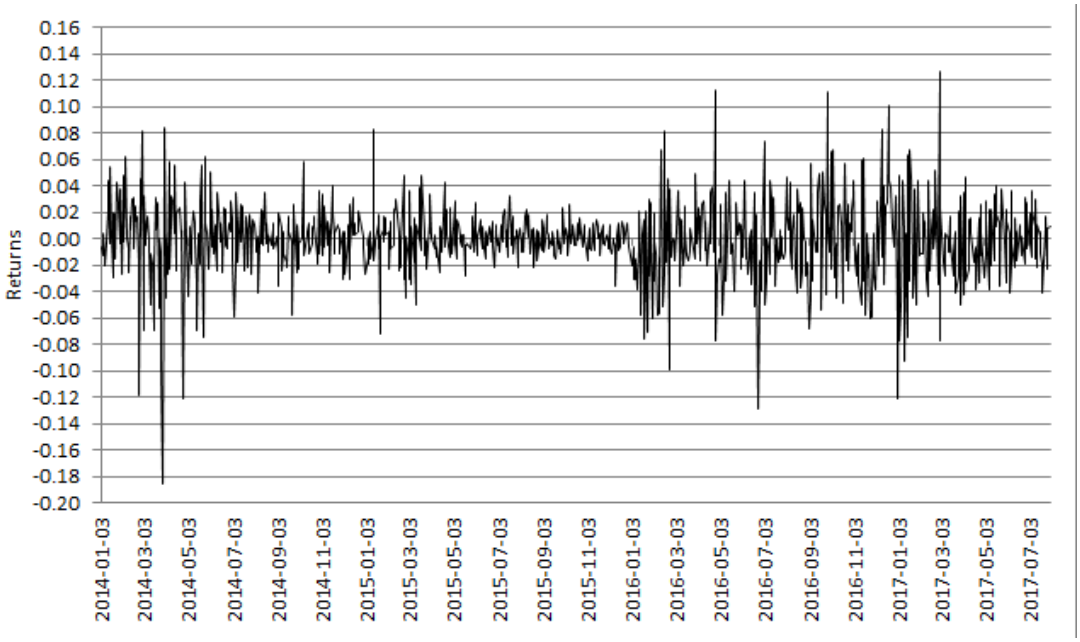


Figure 3.3: Daily EUA returns from January 2014 to July 2017

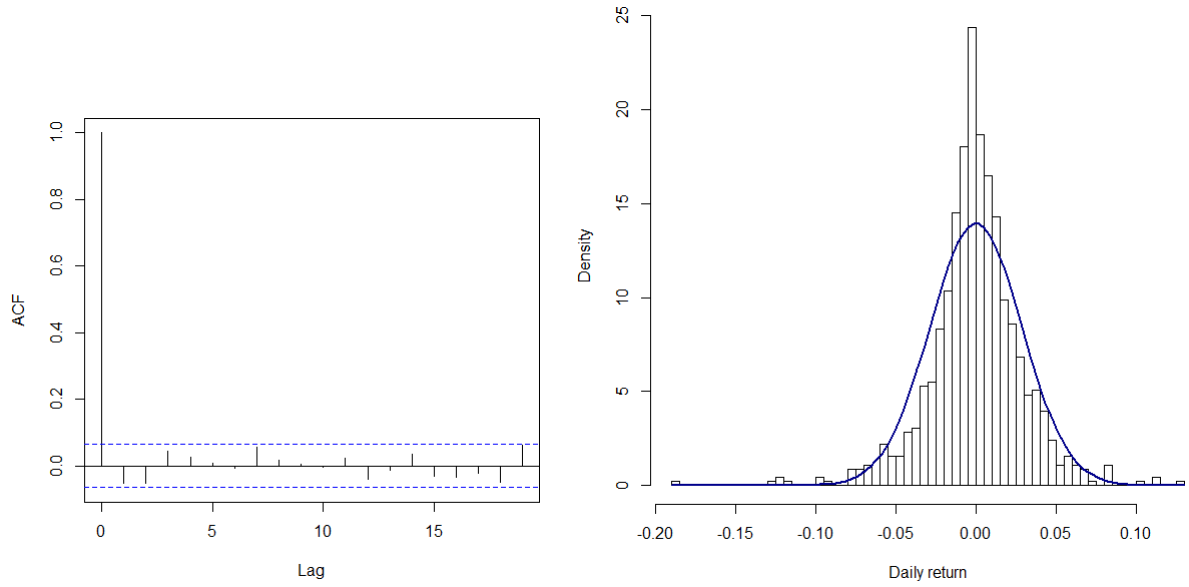
3.2 Model specifications and assumptions

As before, we denote the filtered probability space by $(\Omega, \{\mathcal{F}_t\}, \mathcal{P})$, where $t \in [0, T]$ in units of years is the time index and T is the ending time of the current trading phase. We shall later see that the phase end time T plays a limited role in allowance option valuation for open trading phases, but is included for completeness.

The following assumptions are made, which apply to all models presented in this chapter unless otherwise stated:

1. The market is complete for the trading phase(s) of interest.
2. Borrowing and lending take place at the constant continuously compounded risk free rate r, which is the same as the setting for closed-trading phase in Chapter 2.
3. All trading can be made continuously over time with negligible transaction costs.

Externalities such as the undue influence of political forces on the ETS directives are not considered in our models.



(a) ACF plot (b) Histogram

Figure 3.4: ACF plot and histogram and EUA daily returns

Unlike the discussion in Section 2.4 for the closed trading phase, the validity of a complete market assumption differs heavily in scope under the open trading phase and can be assessed in a number of ways. On one hand, under the open trading phase, the emission information has a reduced effect on allowance prices due to the non-lapsing nature of the allowance contracts. Releases of the emission status may induce corrective shifts in market sentiments toward the compliance value of the allowances, which in turn lead to price movements. However, they do not necessarily lead to price jumps, the absence of which is evidenced in Figure 3.2. The effect of an emission status information release under an open trading phase is thus analogous to that of a public company's annual earnings release on its stock price. As a result, the incompleteness of emission information is now of limited implications on market completeness. On the other hand, absent of the price jump risk, replicating portfolios can be readily constructed for traded contingent claims on allowance and allowance futures contracts, including the options market which is of particular interest in this chapter. The improved market stability and volume in the open trading phase exemplified by EU ETS Phase 2 and 3 has substantially incentivized the participation of global market makers, leading to cost-efficient hedging for allowance derivatives. Therefore, the complete market assumption is considered to be reasonable under our open phase

setting.

3.3 Lognormal allowance price model

In the first reduced-form model we present in this chapter, we assume that the allowance price follows a Geometric Brownian Motion (GBM) with constant drift and volatility parameters. This specification coincides with that in the Black-Scholes framework for a stock option valuation proposed in the seminal work of Black and Scholes (1973)[9], which has been later extended to the valuation of various financial derivatives and structured securities. For brevity, we adapt and reproduce only the important results here in the context of an emission allowance option valuation.

Assume that under real world probability, the spot allowance price $S(t)$ follows a time-homogeneous Geometric Brownian Motion given by:

$$dS(t) = \mu S(t)dt + \sigma S(t)dZ(t) \tag{3.1}$$

with the initial condition $S(0) = S_0$, where

1. μ is a drift parameter representing the expected allowance price return.
2. σ is the volatility of the allowance price return.
3. $Z(t)$ is a standard Brownian Motion under the real world measure $\mathcal{P}$.

The risk neutral allowance price process is obtained by the application of Girsanov's Theorem; so that under the risk neutral measure $\mathcal{Q}$, we have:

$$dS(t) = rS(t)dt + \sigma S(t)dW(t), \tag{3.2}$$

with the initial condition $S(0) = S_0$, where $W(t)$ is a standard Brownian Motion under $\mathcal{Q}$ with a likelihood ratio:

$$\frac{d\mathcal{Q}}{d\mathcal{P}} = \exp\left(\frac{\mu - r}{\sigma}Z(T^*) - \frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2 T^*\right), \quad \forall T^* > 0.$$

From now on, unless otherwise specified, we work under the risk neutral measure and hence the probability space of interest becomes $(\Omega, \{\mathcal{F}_t\}, \mathcal{Q})$. The solution to (3.2) is:

$$S(t) = S_0 e^{\left(r-\frac{1}{2}\sigma^2\right)t+\sigma W(t)}, \tag{3.3}$$

where S_0 represents the initial allowance price at time 0. More generally, given the time-u price information $S(u) = S_u$, we have:

$$S(t) = S_u e^{\left(r-\frac{1}{2}\sigma^2\right)(t-u)+\sigma W(t-u)} \quad \forall t > u. \tag{3.4}$$

Due to market completeness, arbitrage-free pricing applies, so that the allowance futures price is given by:

$$F(t, \tau') = e^{r(\tau'-t)} S(t). \tag{3.5}$$

By construction, an allowance futures contract always settles before the end of current trading phase, or equivalently, $\tau' < T$. This represents an important difference between the emission allowance market and the stock market by imposing a natural upper bound to the life of a futures contract. In addition, this reveals the necessity of rolling futures portfolios over trading phases for both investment and hedging purposes, for which prudent strategies must be devised.

The allowance futures price is a martingale whose dynamics is given by:

$$dF(t, \tau') = \sigma F(t, \tau') dW(t) \tag{3.6}$$

with the initial condition $F(0, \tau) = F_0$ and terminal condition $\mathrm{F}(\tau, \tau) = S(\tau)$, as well as the solution:

$$F(t, \tau') = F_0 e^{-\frac{1}{2}\sigma^2 t+\sigma W(t)}. \tag{3.7}$$

More generally, given the time-u futures price information $F(u, \tau') = F_u$, we have:

$$F(t, \tau') = F_u e^{-\frac{1}{2}\sigma^2(t-u)+\sigma W(t-u)} \quad \forall t \in (u, \tau'). \tag{3.8}$$

These results above translate to conditional Lognormality of the prices in the allowance spot and futures markets, which can be expressed in general as:

$$S(t) \mid S_u \sim LN(\ln(S_u) + (r - 0.5\sigma^2)(t-u), \sigma\sqrt{t-u}) \quad \forall u < t, \tag{3.9}$$

$$F(t, \tau') \mid F_u \sim LN(\ln(F_u) - 0.5\sigma^2(t-u), \sigma\sqrt{t-u}), \quad \forall t \in (u, \tau'), \tag{3.10}$$

where $LN(.,.)$ denotes the Lognormal distribution with the two arguments respectively being the mean and standard deviation parameters of the underlying normal variable.

3.3.1 Option valuation under a Lognormal allowance price

This section presents the valuation formula for allowance options under the Lognormal allowance price model (abbreviated as the LN model). As mentioned previously in Chapter 2, all allowance (futures) options traded under the EU ETS framework are European style. Consequently, option valuation in this setting is an adaptation of the results from the Black-Scholes framework. For consistency, we use the same notation as in Chapter 1, with superscripts added to indicate the model under which valuation is performed. For example:

1. $C^{LN}(t, K, \tau)$ and $P^{LN}(t, K, \tau)$ denote respectively the time t values of a call and put option on the allowance contract with strike price K and maturity τ under the Lognormal allowance price model.

2. $C_F^{LN}(t, K, \tau, \tau')$ and $P_F^{LN}(t, K, \tau, \tau')$ denote respectively the time t values of a call and put option with maturity τ and strike price K written on the allowance futures to be settled at τ' under the Lognormal allowance price model, where $0 \leq t < \tau \leq \tau' \leq T$.

In addition, whenever the context is clear, the time index argument may be moved from the bracket to the subscript to emphasize that the information is known as of the valuation point. For example, the time-t allowance price $S(t)$ may be denoted by S_t when considering the time-t option values. The value of an allowance option under the Lognormal allowance price model follows Proposition 3.1.

Proposition 3.1: *Consider an open trading phase spanning the period* $[0, T]$. *Assume that the spot allowance price follows the GBM given by (3.2). The risk neutral prices at time* t *of the European allowance call and put option with strike price* K *maturing at time* $\tau \in (t, T)$ *are respectively given by:*

$$
\begin{aligned}
C^{LN}(t, K, \tau) &= S_t \Phi(d_1) + K e^{-r(\tau - t)} \Phi(d_2), \\
P^{LN}(t, K, \tau) &= K e^{-r(\tau - t)} \Phi(-d_2) - S_t \Phi(-d_1),
\end{aligned}
$$

where

$$
\begin{aligned}
d_1 &= \frac{\ln\left(\frac{S_t}{K}\right) + (r + 0.5\sigma^2)(\tau - t)}{\sigma\sqrt{\tau - t}}, \\
d_2 &= d_1 - \sigma\sqrt{\tau - t}.
\end{aligned}
$$

The results in the proposition are reached using the risk neutral valuation as well as equations (3.9) to (3.10).

For an allowance futures option, the risk neutral price can either be derived using the result of (3.10), or through a simpler approach using the arbitrage-free relationship between the spot and futures allowance price given by (3.5) as well as the result of Proposition 3.1. To see this, notice that the expected payoff of the futures call option is:

$$E\left[(F(\tau,\tau')-K)^{+}\right] = e^{r(\tau'-\tau)}E\left[\left(S(\tau)-Ke^{-r(\tau'-\tau)}\right)^{+}\right]. \tag{3.11}$$

From this equation, we see that an allowance futures call option is equal in expected payoff to a scalar multiple of an allowance call option. A similar result can be derived for an allowance futures put option. Using this relationship as the basis, simple algebraic manipulations of the results in Proposition 3.1 leads to Proposition 3.2 below.

Proposition 3.2: *Consider an open trading phase spanning the period* $[0,T]$*. Assume that the spot allowance price follows the GBM given by (3.2). The risk neutral prices at time* t *of the European allowance futures call and put option with strike price* K *maturing at time* τ*, written on an allowance futures contract to be settled at* τ' *where* $0 \le t < \tau \le \tau' \le T$*, are respectively given by:*

$$\begin{aligned} C_F^{LN}(t,K,\tau,\tau') &= e^{-r(\tau-t)}\left(F(t,\tau')\Phi(d_1')+K\Phi(d_2')\right), \\ P_F^{LN}(t,K,\tau,\tau') &= e^{-r(\tau-t)}\left(K\Phi(-d_2')-F(t,\tau')\Phi(-d_1')\right), \end{aligned}$$

where

$$\begin{aligned} d_1' &= \frac{\ln\left(\dfrac{F(t,\tau')}{K}\right)+0.5\sigma^2(\tau-t)}{\sigma\sqrt{\tau-t}}, \\ d_2' &= d_1-\sigma\sqrt{\tau-t}. \end{aligned}$$

Notice that τ', the allowance futures settlement time, does not enter into the allowance futures option value expression. This is because the option payoff is fully realized at maturity and any resultant marked-to-market cashflows from the futures position is not attributable to the option.

The Lognormal allowance price model serves as the benchmark model in our analysis, based on which more complex models are built as presented in subsequent sections. Numerical

results from the implementation of allowance futures option valuations under alternative models are also compared to those from Proposition 3.1 to assess model fitness. Due to the heavy tails observed in the actual allowance price returns in Section 3.1 that are not captured by the GBM price process, it is expected that models capturing the excess kurtosis shall bring some recognizable improvements in model fitness, which will be verified by calibrating them to market option data.

3.3.2 Modification for skewness and kurtosis

One direct modification to the Lognormal allowance price model is inspired by Corrado and Su (1996)[28] who introduced skewness and kurtosis corrections to the stock return model in analyzing the observed deviations from Normality in S&P500 index returns. In theory, similar corrections can be applied to the Lognormal allowance price model to achieve improved fitness by accounting for the skewness and excess kurtosis in allowance price returns as discussed in Section 3.1, where the key results are described below.

First, consider the allowance return over an arbitrary period $[u, t]$ given all price information up to u. Based on the result of (3.9), this return variable follows a Normal Distribution, which can be expressed as:

$$\ln\left(\frac{S(t)}{S_u}\right) \mid \mathcal{F}_u \sim N(r - 0.5\sigma^2, t - u), \quad \forall u < t. \tag{3.12}$$

Hence, the standardized return variable $R(u, t)$ follows the standard Normal distribution with a density function:

$$f_R(x) = \phi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-0.5x^2}, \tag{3.13}$$

where

$$R(u,t) = \frac{\ln\left(\frac{S(t)}{S_u}\right) - (r - 0.5\sigma^2)(t-u)}{\sigma\sqrt{t-u}} \qquad \forall u < t. \tag{3.14}$$

To introduce skewness and excess kurtosis to the return distribution, Corrado and Su (1996) applies a Gram-Charlier expansion to the density function with truncated Hermite polynomial terms, so that the standardized return $R(u, t)$ now has the following conditional

density function:

$$g_R(x) = \phi(x)\left[1 + \frac{\mu_3(x^3 - 3x)}{3!} + \frac{(\mu_4 - 3)(x^4 - 6x^2 + 3)}{4!}\right], \tag{3.15}$$

where μ_3 and μ_4 are the explicit parameters for the skewness and kurtosis of the allowance price returns. Under this transformed density function, we have:

$$\begin{aligned} E(R) &= 0 \qquad E(R^2) = 1 \\ E(R^3) &= \mu_3 \qquad E(R^4) = \mu_4 \end{aligned}$$

In the original Lognormal allowance price model, we would have $\mu_3 = 0$ and $\mu_4 = 3$, which corresponds to the standard Normal specification for the standardized allowance price return $R(u,t)$. Therefore, this modified model can be viewed as a generalized version of the Lognormal allowance price model, allowing for nonzero skewness and excess kurtosis.

Now consider the time-t risk neutral price of an allowance call option with strike price K and maturity τ under this skewness-kurtosis-modified Lognormal allowance price model (abbreviated as the SKM model), which we denote by $C^{SKM}(t,K,\tau)$. Under $\mathcal{Q}$, we have:

$$\begin{aligned} C^{SKM}(t,K,\tau) &= e^{-r(\tau-t)}[(S(\tau) - K)^+] \\ &= e^{-r(\tau-t)} \int_K^\infty (S(\tau) - K) f_{S(\tau)|\mathcal{F}_t}(s) ds \\ &= e^{-r(\tau-t)} \int_k^\infty \left(S_t \exp(x\sigma\sqrt{\tau - t} + (r - 0.5\sigma^2)(\tau - t)) - K\right) g_R(x) dx, \end{aligned} \tag{3.16}$$

where

$$k = \frac{\ln\left(\frac{K}{S_t}\right) - (r - 0.5\sigma^2)(\tau - t)}{\sigma\sqrt{\tau - t}}, \tag{3.17}$$

and $g_R(x)$ is given in (3.23). This expression is evaluated to yield the following expression:

$$C^{SKM}(t,K,\tau) = C^{LN}(t,K,\tau) + \mu_3 Y_3 + (\mu_4 - 3)Y_4, \tag{3.18}$$

where $C^{LN}(t,K,\tau)$ is the corresponding call option price under the Lognormal allowance price model given in Proposition 3.1, and

$$Y_3 = \frac{S_t\sigma\sqrt{\tau - t}}{3!}\left((2\sigma\sqrt{(\tau - t)} - d_1)\phi(d_1) - \sigma^2(\tau - t)\Phi(d_1)\right), \tag{3.19}$$

$$Y_4 = \frac{S_t \sigma \sqrt{\tau - t}}{4!} \left((d_1^2 - 1 - 3\sigma\sqrt{(\tau - t)} d_2)\phi(d_1) + \sigma^3 (\tau - t)^{1.5} \Phi(d_1) \right), \tag{3.20}$$

with d_1 and d_2 also defined in Proposition 3.1.

The Y_3 and Y_4 components in this option price expression represent, respectively, the marginal effect of non-normal skewness and kurtosis in allowance returns to the call option value under the risk neutral measure. Negative skewness, as shown by Corrado and Su (1996), causes the Lognormal model to overprice out-of-the-money options and underprice in-the-money options, while the opposite case holds for positive skewness.

Based on the results of (3.11) and (3.18), we derive the valuation expression of allowance futures options. The time-t risk neutral price of a European allowance futures call option with strike price K and maturity $\tau \in (t, T)$ written on an allowance futures contract to be settled at τ', denoted by $C_F^{SKM}(t, K, \tau, \tau')$, $0 \leq t < \tau \leq \tau' \leq T$, follows:

$$C_F^{SKM}(t, K, \tau, \tau') = C_F^{LN}(t, K, \tau, \tau') + \mu_3 Y_3' + (\mu_4 - 3) Y_4', \tag{3.21}$$

where $C_F^{LN}(t, K, \tau, \tau')$ is the corresponding allowance futures call option price under the Lognormal allowance price model given in Proposition 3.2, and

$$Y_3' = \frac{F(t, \tau')\sigma\sqrt{\tau - t}}{3! e^{r(\tau - t)}} \left(\left(2\sigma\sqrt{(\tau - t)} - d_1' \right) \phi(d_1') - \sigma^2 (\tau - t) \Phi(d_1') \right), \tag{3.22}$$

$$Y_4' = \frac{F(t, \tau')\sigma\sqrt{\tau - t}}{4! e^{r(\tau - t)}} \left(\left({d_1'}^2 - 1 - 3\sigma\sqrt{(\tau - t)} d_2' \right) \phi(d_1') + \sigma^3 (\tau - t)^{1.5} \Phi(d_1') \right). \tag{3.23}$$

with d_1' and d_2' defined in Proposition 3.2.

Similarly, the Y_3' and Y_4' components in this option price expression represent, respectively, the marginal effect of non-normal skewness and kurtosis in allowance returns to the allowance futures call option value under $\mathcal{Q}$. The futures contract settlement time does not enter into the option price expression for the same reason pointed out in the previous section for the LN model.

The put option prices under the skewness-kurtosis-modified Lognormal model can be found using the put-call parity relationships described in (2.51) and (2.55), whose validity are not affected by the choice of model and the type of trading phase. For brevity, the results are not presented here.

3.3.3 Volatility smile in allowance options

While the option valuation formula under the Lognormal allowance price model is appealing in practice for its simplicity and ease of implementation, it shares several shortcomings of the Black-Scholes framework for a stock option valuation that puts the fitness and appropriateness of this benchmark model under scrutiny, one of which is the phenomenon of the volatility smile. In this section, we study this stylized feature of option values for the allowance market, which shall motivate the presentation of the subsequent models in this Chapter, while contributing to the literature as similar studies on the subject is rare.

Volatility smile refers to the phenomenon that the option-implied volatility varies with respect to the strike price, where the relationship graphically displays a convex curve. Heuristically, implied volatility is derived from the spot option price by solving for the volatility value that equates the market and model price of an option. However, in reality, allowance options are traded over-the-counter, forming a relatively small market with comparatively low liquidity and trading volumes. Therefore, as a preferred alternative, the implied volatility should be derived from the price of allowance futures options. In contrast to allowance options, allowance futures options are standardized contracts actively traded in major commodity exchanges such as the ICE Europe. A formal definition for implied volatility is thus given below.

Definition 3.1: *The implied volatility of an allowance futures option, denoted by σ_{im}, is the volatility that equates the option's price under the Lognormal allowance price model to its market price. For a European allowance futures call with strike price K and maturity τ written on an allowance futures contract to be settled at τ', its time-t implied volatility is defined by the following relationship*:

$$C_F^m(t, K, \tau, \tau') = C_F^{LN}(t, K, \tau, \tau'), \tag{3.24}$$

or, in an expanded form, by

$$C^m(t,K,\tau,\tau') = F(t,\tau')\Phi\left(\frac{\ln\left(\frac{F(t,\tau')}{K}\right) + 0.5{\sigma_{im}}^2(\tau-t)}{\sigma_{im}\sqrt{\tau-t}}\right) - Ke^{-r(\tau-t)}\Phi\left(\frac{\ln\left(\frac{F(t,\tau')}{K}\right) - 0.5\sigma_{im}^2(\tau-t)}{\sigma_{im}\sqrt{\tau-t}}\right) \tag{3.25}$$

where $C^m(t,K,\tau,\tau')$ is the time-t market price of the allowance futures call option.

For the rest of this chapter, unless otherwise specified, the term"implied volatility" and the notation σ_{im} refers to the volatility implied by an allowance futures option price as in Definition 3.1, which is different from the implied volatility of stock options. Equation (3.25) shows that solving for the implied volatility from market prices requires numerical procedures except for the ideal case where the futures option is at-the-money (ATM). For this case, equation (3.25) is substantially simplified so that an analytical expression for the implied volatility can be obtained.

According to the specifications of the Lognormal allowance price model, the constant volatility parameter σ is invariant to the option parameters. However, our empirical analysis reveals the existence of volatility smiles: the phenomenon that occurs when holding all else constant, the implied volatility varies with the strike price, which forms the shape of a"smile". Figure 3.5 illustrates a typical volatility smile derived using the April 11, 2017 daily closing quotes of a call option written on the December 2017 EUA futures.

Notice that the smile is asymmetric and skewed toward the in-the-money region, which is common for allowance futures call options. Volatility smiles are very well studied in the stock option market in the literature, both on the causes of the phenomenon and the solutions in the form of more complex models aiming to capture the smile. Leveraging the various plausible explanations and theories, a major cause for the observed volatility smile from the allowance futures options is the crash-o-phobia effect, reflected as a density cluster of a series of substantial losses. This effect also contributes to the fat tails and hence excess kurtosis observed in the allowance price returns identified in Section 3.2, which is not properly captured by the Lognormal allowance price model. Prior studies on the crash-o-phobia effect include Melick and Thomas (1997)[72], who reported similar

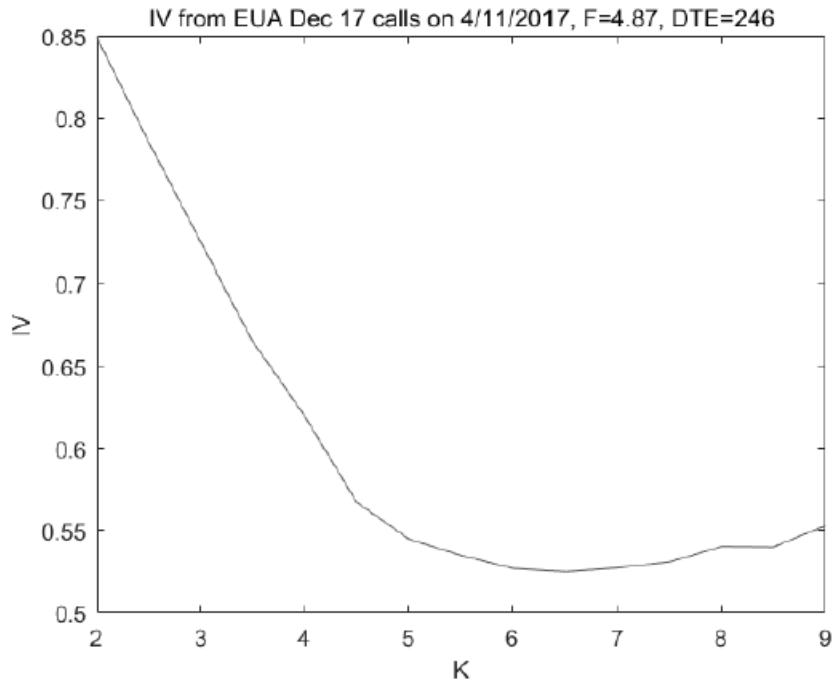


Figure 3.5: Illustrative volatility smile from allowance futures option

problems using the American oil futures options during the Gulf War crisis. Since emission allowance can also be viewed as a production input for industries governed by the emission control scheme, it may reasonably share some characteristics of the commodity market. Hence crash-o-phobia and the resultant volatility smiles are expected in our case.

On the solution side, well-known efforts aimed at capturing stock option volatility smiles include the stochastic volatility models by Hull and White (1987)[56] and Heston (1993)[54], the local volatility pricing method by Dupire (1994)[37], and econometric models incorporating GARCH-type conditional volatility processes by Engel and Mustafa (1992)[39]. Given that no studies have been published on the volatility smiles observed in the allowance option markets, those results inspire us to present the models in the subsequent section that improve over the benchmark model by accounting for this stylized fact in allowance option values.

Finally, for stock options, volatility smile is also observed across the maturity dimen-

sion. Unfortunately, a study of implied volatility across maturities is by construct difficult for allowance futures options, since the contract is standardized with a limited number of maturity points and hence provides insufficient data for us to draw any reliable conclusion.

3.4 Mixture Lognormal allowance price model

In this section, we present the Mixture Lognormal model for allowance prices (abbreviated as the MLN model) that should deliver improved fitness over the benchmark model by capturing the excess kurtosis and volatility smile.

Assume that under the real world probability space $(\Omega, \{\mathcal{F}_t\}, \mathcal{P})$, the spot allowance price $S(t)$ follows a mixture of N diffusion processes:

$$S(t) = \sum_{i=1}^{N} 1_{C=i} S_i(t), \tag{3.26}$$

$$dS_i(t) = S_i(t)\mu_i dt + \sigma_i S_i(t) dZ(t), \quad i \in \{1, 2...N\}, \tag{3.27}$$

with the common initial condition $S_i(0) = S_0$, where $Z(t)$ is a standard Brownian Motion under $\mathcal{P}$, and

1. $1_{C=i}$ is the binary indicator variable that equals 1 when $C = i$ and 0 otherwise.
2. C is the unobservable underlying market state random variable taking values on $1, 2, 3...N$ with $Pr(C = i) = \lambda_i$ subject to $\sum_{i=1}^{N} \lambda_i = 1$.
3. μ_i and σ_i are respectively the drift and volatility of the allowance price return given that the market is in state i.

From a mathematical perspective, this mixture diffusion specification of allowance price is a superposition of N components price processes denoted by triplets (S_i, μ_i, σ_i) with corresponding mixing weights $\lambda_i, i \in \{1, 2, ...N\}$. These weights also serve as the probability masses of the underlying market state variable C that determines the component process in force, and hence make up the joint distribution of the parameter pairs (μ_i, σ_i). This extra source of randomness introduces flexibility to the model in capturing possible excess

kurtosis in allowance price returns. Notice that in the most general case, the state variable C should model a continuum of market state scenarios reflecting the aggregate effect of various economic and non-economic factors, in which case a continuous bivariate joint distribution of the parameter pairs must be specified. Here we have implicitly discretized the possible market state into N cases to better serve our valuation purpose in actual implementation.

In terms of its interpretation, the mixed diffusion specification emphasizes that the allowance price dynamics is subject to uncertainty due to the agents' different behaviors and reactions to information under different market states, which is consistent with modern economics theory as pointed out by Neumann (2002)[78]. More specifically, market participants trade in response to the arrival of new information. Hence each component diffusion process can be perceived as random results of information events under the corresponding market state. However, the effect of a specific information event on the allowance price differs by the market state. For instance, news of a large emitter downsizing its business should not result in much price reactions in an optimistic and confident market expecting steady economic growth, compared to a pessimistic market with high information asymmetry. The market state hence represents the aggregate effect of economic outlooks and market sentiments, which respectively dictates the drift and volatility terms in the component price processes in (3.27). Since the market state is not observable in the market, it is modeled by a categorical variable C with probability masses over $1, 2...N$. Note that the state remains time-invariant in the specification. Thus, inter-temporal changes in market state are not explicitly captured. Instead, one chooses among the component diffusion processes to capture the market dynamics by calibrating the model to derivative prices, and consequently the parameters change over time.

We now formally introduce the Mixture Lognormal distribution in Definition 3.2 below:

Definition 3.2: *Under a filtered probability space* $(\Omega, \{\mathcal{F}_t\}, \mathcal{P})$*, let* C *denote the underlying discrete state variable taking values on* $\{i \mid i \in 1, 2, ...N\}$ *with* $P(C = i) = \lambda_i$*. Also, let* $1_{C=i}$ *denote a binary indicator variable that takes the value of* 1 *if* $C = i$ *and* 0 *otherwise. A random variable* Y *follows a Mixture Lognormal distribution iff*

$$Y = \sum_{i=1}^{N} 1_{C=i} X_i,$$

where the X_i's are independent Lognormal random variables with density functions:

$$f_{X_i}(x; m_i, \sigma_i) = \frac{1}{\sqrt{2\pi} x \sigma_i} \exp\left\{ -\frac{(\ln x - m_i)^2}{2\sigma_i^2} \right\}, \quad i \in \{1, 2...N\}.$$

In addition, let Λ, M, and $\Sigma \in R^N$ be vectors of the weight parameters λ_i, the mean parameters m_i, and the standard deviation parameters σ_i, $i \in 1, 2...N$ respectively. Then the parametric distributional relationship of Y is written as:

$$Y \sim MLN(\Lambda, M, \Sigma)$$

with a density function:

$$f_Y(y) = \sum_{i=1}^{N} \lambda_i f_{X_i}(y).$$

Each component allowance price follows a Geometric Brownian Motion in (3.27), which translates to conditional Lognormality by the results in Section 3.3 on the LN model. Hence, it follows that:

$$S_t \mid S_u \sim MLN\left(\Lambda, 1^N ln S_u + U(t-u) - \frac{1}{2}\Sigma^2(t-u), \Sigma\sqrt{t-u} \right), \quad \forall u < t \qquad (3.28)$$

where 1^N denotes a vector of ones of size N and $U \in R^N$ is the vector containing the component drift parameters (i.e the μ_i's).

Finally, it is of interest to discuss the relationship between the MLN model and the Regime Switching Lognormal (RSLN) model proposed by Hardy (2001)[49]. Both models assume the existence of different return dynamics described by pairs of drift and volatility parameters. Given N possible states, the RSLN model explicitly captures the migrations between dynamics over time steps using a Markov Chain process, in which a transition probability matrix of size N by N is used. In contrast, no transition between states is explicitly modeled in the MLN framework. The allowance return is simply assumed to follow one of the N dynamics corresponding to the N states, the choice of which is unknown and hence modeled discretely with a fixed probability mass distribution.

From a numerical prospective, the closed-form option valuation expression under the RSLN model requires the discretized distribution of the total sojourns for each regime[3], which

[3]Given a fixed time frame, the sojourn of a regime is a variable representing the total number of periods the process spends in the regime. See Hardy (2001)[49]

must be estimated recursively. On the other hand, closed-form option valuation formula naturally exists under the MLN framework, where all parameters can be easily calibrated to market option prices using error-minimizing principles. Overall, the MLN model is more parsimonious and numerically efficient in practice, while capturing both the volatility smile and heavy tails in the allowance return.

3.4.1 Option valuation under the Mixture Lognormal Model

In this section we present the valuation of allowance options under the Mixture Lognormal allowance price model. While it is tempting to apply a risk neutral valuation directly, the MLN model is by construction an incomplete-market model due to the extra source of randomness introduced by the market state variable, and hence the risk neutral measure is not unique. Approaches to valuation of contingent claims under incomplete-market models are abundant in the literature, which includes:

1. Assuming a particular form of the market's utility function. This is equivalent in effect to specifying a market price of risk to be used for an application of the change of measure. Examples include Pellizzari and Gamba (2002)[82] and Zhang and Han (2013)[92], both of which are under the stock market settings.

2. Assuming that the risk neutral measure is chosen to be the one closest to the real world measure under some metric. Examples include He and Zhu (2015)[52] on European stock option valuation using the minimal entropy martingale measure, as well as Cetin and Verscheure (2009)[25] on allowance valuation under closed trading phases using the minimal martingale measure.

3. Directly specifying the risk neutral dynamics of the allowance price process and calibrate the resultant model option prices to market data. Examples include Leison (2004)[67] on stock option valuation using mixture models.

From our perspective, the last approach is appealing for its objectivity and is hence implemented in our analysis. There is no convincing evidence on any particular utility functions being followed by the market agents, which challenges the validity and robustness of the first approach. A similar argument can be made for the second approach that also lacks a comprehensive theoretical foundation. There is no economic rationale on why the market must price a derivative under a risk neutral measure close to $\mathcal{P}$. In contrast, the last approach does not require any assumptions of utility functions or a particular risk measure

being used by the market. Instead, The risk neutral price dynamics is directly specified and the model option prices are calibrated to their market quotes, which reflect all of the information of the true risk neutral measure $\mathcal{Q}$ chosen by the market.

From now on, we work under the risk neutral measure unless otherwise specified and hence the probability space of interest becomes $(\Omega, \{\mathcal{F}_t\}, \mathcal{Q})$. Assume that the spot allowance price under $\mathcal{Q}$ follows a mixture of diffusions given by (3.26) and component processes:

$$dS_i(t) = rS_i(t)dt + \sigma_i S_i(t)dW(t), \quad i \in \{1, 2...N\}, \tag{3.29}$$

with common initial condition $S_i(0) = S_0$, where $W(t)$ is a standard Brownian Motion under $\mathcal{Q}$ and r is the continuously compounded risk free rate. Notice that compared with the $\mathcal{P}$-measure allowance price process in (3.26) and (3.27), the only change made under $\mathcal{Q}$ is that all component drift parameters are set to the risk free rate. Since each component process follows a Geometric Brownian Motion that translates to conditional Lognormality, based on Definition 3.2, the allowance price follows the conditional Mixture Lognormal distribution that can be written as:

$$S_t \mid S_u \sim MLN\left(\Lambda, 1^N lnS_u + 1^N r(t-u) - \frac{1}{2}\Sigma^2(t-u), \Sigma\sqrt{t-u}\right), \quad \forall u < t. \tag{3.30}$$

The results above are sufficient for the valuation of allowance options. The same set of notations as in Section 3.3 are adopted with an updated superscript to indicate the model used, so that $C^{MLN}(t, K, \tau)$ and $P^{MLN}(t, K, \tau)$ denote respectively the time t values of a call and put option on the allowance contract with strike price K and maturity τ under the Mixture Lognormal allowance price model. The best property of the MLN model is that, for a given option, its value is simply a linear combination of the N option values under the Lognormal allowance price model calculated with the component drift and volatility parameters, with the mixture weights being the coefficients. Proposition 3.3 below formally summarizes this relationship:

Proposition 3.3: *Consider an open trading phase spanning the period* $[0, T]$. *Assume that the spot allowance price follows the mixture diffusion process given by (3.26) and (3.29) under* $\mathcal{Q}$. *The time-t price of a European allowance option (either call or put) with strike price* K *maturing at time* τ *is equal to the linear combination of those under the Lognormal allowance price model calculated using component volatilities* σ_i *weighed by the mixing weights* λ_i, $i \in \{1, 2, ...N\}$. *That is:*

$$C^{MLN}(t, K, \tau) = \sum_{i=1}^{N} \lambda_i C^{LN}(t, K, \tau, \sigma_i),$$

$$P^{MLN}(t,K,\tau) = \sum_{i=1}^{N} \lambda_i P^{LN}(t,K,\tau,\sigma_i),$$

where $C^{LN}(t,K,\tau,\sigma_i)$ *and* $P^{LN}(t,K,\tau,\sigma_i)$ *denote respectively the European call and put allowance option values under the Lognormal allowance price model given by Proposition 3.1 calculated with volatility* σ_i.

Proof: See Appendix P.

Assume that the market is arbitrage-free in spite of the MLN allowance price specification being by construction an incomplete-market model, so that the relationship in (3.5) holds.[4]. The allowance futures price is a martingale with dynamics as given by a mixture of diffusion processes below:

$$F(t,\tau') = \sum_{i=1}^{N} 1_{C=i} F_i(t,\tau'), \tag{3.31}$$

$$dF_i(t,\tau') = \sigma_i F_i(t,\tau') dW(t), \quad i \in \{1,2...N\}, \tag{3.32}$$

with the common initial condition $F_i(0,\tau') = F_0$ and terminal conditions $F_i(\tau',\tau') = S_i(\tau')$. These results lead us to the expressions for the risk neutral price of allowance futures options, which are also linear combinations of the corresponding prices under the Lognormal model as given by Proposition 3.4 below.

Proposition 3.4: *Consider an open trading phase spanning the period* $[0,T]$. *Assume that the spot allowance price follows the mixture diffusion process given by (3.26) and (3.29) under* $\mathcal{Q}$. *The time-t price of a European allowance option (either call or put) with strike price* K *maturing at time* τ, *written on an allowance futures contract to be settled at* τ' *where* $0 \leq t < \tau \leq \tau' \leq T$, *is equal to the linear combination of those under the Lognormal allowance price model calculated using component volatilities* σ_i *weighed by the mixing weights* λ_i, $i \in \{1,2,...N\}$. *That is:*

$$C_F^{MLN}(t,K,\tau,\tau') = \sum_{i=1}^{N} \lambda_i C_F^{LN}(t,K,\tau,\tau',\sigma_i),$$

$$P_F^{MLN}(t,K,\tau,\tau') = \sum_{i=1}^{N} \lambda_i P_F^{LN}(t,K,\tau,\tau',\sigma_i),$$

[4]Detailed discussions on the relationship between arbitrage-free and complete markets can be found in Bjork (2004)[6].

where $C_F^{LN}(t, K, \tau, \sigma_i)$ and $P_F^{LN}(t, K, \tau, \sigma_i)$ denote respectively the European call and put allowance futures option values under the Lognormal allowance price model given by Proposition 3.2 calculated with volatility σ_i.

The proof for Proposition 3.4 follows exactly the same logic as the one for Proposition 3.3 in Appendix P, which is omitted for brevity. In fact, by generalization of the arguments presented, it is quite clear that the risk neutral price of any contingent claim under the MLN model is a λ_i-weighted linear combination of those under the LN model calculated with the mixture component parameters. Such feature makes the model an appealing choice for the valuation of complex allowance derivatives when a balance is desired between the adequacy of the results and the ease of implementation.

3.4.2 The Two-component Lognormal Mixture

For a practical implementation of the MLN model, the total number of mixture components must be specified. Introducing additional components leads to improved fitness to market option prices at the cost of numerical efficiency, since a pair of parameters (i.e the mixture weight and the component volatility) must be estimated under $\mathcal{Q}$ for each component. In this section, we present a few important results for the two-component Mixture Lognormal allowance price model (abbreviated as the MLN-2 model), which is the simplest instance of the MLN framework. Incorporating the crash-o-phobia effect and volatility smiles[5], the model relies on a mixture of two component variables with interesting interpretations that are not transparent for the MLN model in general.

The specification of the MLN-2 model is a trivial instance of those of the MLN model by letting $N = 2$. Under $\mathcal{P}$, assume that the spot allowance price follows a mixture of two diffusion processes:

$$S(t) = \sum_{i=1}^{2} 1_{C=i} S_i(t), \tag{3.33}$$

$$dS_i(t) = \mu_i S_i(t)dt + \sigma_i S_i(t)dZ(t), \quad i \in \{1, 2\} \tag{3.34}$$

[5]Fang (2012)[41] demonstrates the capability of a simple two-component Lognormal Mixture model in producing the volatility smiles in stock index options under a semi-hypothetical setting. Such capability is derived from the structural flexibility of the mixture specification and hence is market independent.

with the common initial condition $S_1(0) = S_2(0) = S_0$, where $Pr(C = 1) = \lambda \in [0, 1]$ and $Z(t)$ is a standard Brownian Motion under $\mathcal{P}$. Due to the conditional Lognormality implied by the component price processes, the allowance price follows the conditional Mixture Lognormal distribution that can be written as:

$$S(t) \mid S_u \sim MLN\left(\begin{bmatrix}\lambda \\ 1-\lambda\end{bmatrix}, \begin{bmatrix}\ln(S_u) + (\mu_1 - 0.5\sigma_1^2)(t-u) \\ \ln(S_u) + (\mu_2 - 0.5\sigma_2^2)(t-u)\end{bmatrix}, \begin{bmatrix}\sigma_1\sqrt{t-u} \\ \sigma_2\sqrt{t-u}\end{bmatrix}\right). \tag{3.35}$$

For an option valuation, assume that the allowance price under $\mathcal{Q}$ follows a mixture of diffusions given by (3.33) and component processes:

$$dS_i(t) = rS_i(t)dt + \sigma_i S_i(t)dW(t), \quad i \in \{1, 2\}, \tag{3.36}$$

with the common initial condition $S_1(0) = S_2(0) = S_0$, where $Pr(C = 1) = \lambda \in [0, 1]$ and $W(t)$ is a standard Brownian Motion under $\mathcal{Q}$. The time-t risk neutral option prices are given by:

$$C^{MLN2}(t, K, \tau, \sigma_1, \sigma_2) = \lambda C^{LN}(t, K, \tau, \sigma_1) + (1-\lambda)C^{LN}(t, K, \tau, \sigma_2), \tag{3.37}$$

$$C_F^{MLN2}(t, K, \tau, \sigma_1, \sigma_2) = \lambda C_F^{LN}(t, K, \tau, \tau', \sigma_1) + (1-\lambda)C_F^{LN}(t, K, \tau, \tau', \sigma_2), \tag{3.38}$$

where the superscript MLN-2 indicates that valuation is performed under the two-component Mixture Lognormal model. The same results apply to put options.

The simplicity of the two-component mixture facilitates the study of the component prices directly, which is clumsy to perform as the number of components increases. First, under $\mathcal{P}$, consider the expected terminal price of the allowance given $\mathcal{F}_t$ and a time frame of interest $[t, \tau]$:

$$\begin{aligned} E[S(\tau) \mid \mathcal{F}_t] &= E\left[E\left[S(\tau) \mid \mathcal{F}_t, C\right]\right] \\ &= \lambda E[S_1(\tau) \mid \mathcal{F}_t] + (1-\lambda)E[S_2(\tau) \mid \mathcal{F}_t]. \end{aligned} \tag{3.39}$$

It is easy to see that the expected terminal price is a linear interpolation of the expected component prices. Without loss of generality, assume that $E[S_1(\tau) \mid \mathcal{F}_t] < E[S_2(\tau) \mid \mathcal{F}_t]$, which directly implies:

$$E[S_1(\tau) \mid \mathcal{F}_t] < E[S(\tau) \mid \mathcal{F}_t] < E[S_2(\tau) \mid \mathcal{F}_t].$$

Then λ is the probability that the pessimistic low-price market state dominates (i.e. $C = 1$), under which the allowance price is expected to decrease. The associated component price variable $S_1(t)$ is referred to as the downside component that captures the

crash-o-phobia effect in the allowance market. While the level of crash-o-phobia depends on λ and $E[S_1(\tau) \mid \mathcal{F}_t]$, the possibility of large price drops is incorporated. Opposite arguments apply to the upside component $S_2(t)$ associated with the optimistic high-price market state (i.e. $C = 2$), which dominates with a probability of $1 - \lambda$.

From an economic perspective, the MLN-2 specification is a simplified discretization of the continuum of possible market scenarios into a dichotomy: a pessimistic state established by bearish factors such as allowance surplus and production downturns, and an optimistic state established by bullish factors including allowance deficiency and tightening regulatory standards. The corresponding downside and upside component dynamics describe the market's reactions to the random arrival of information under each state. This is consistent with our previous interpretations for the MLN model as well as the modern finance theory, where asset price movements are explained by a set of fundamental factors and a random component.

Nevertheless, the above interpretations of the downside and upside components fails to apply to the model under $\mathcal{Q}$, where both expected terminal component prices are equal to the current price accumulated at the risk free rate as required by the martingale restriction. This is easily verified by examining the specification in (3.36). Upon the calibration of the model to market option prices, the contrasting implications of the two components are reflected in their volatility parameter estimates. Therefore, to properly differentiate between the two components under the risk neutral measure, we identify the downside component to be the one associated with a higher volatility parameter. This follows from that the component price process with a higher volatility has a smaller drift term in the corresponding return process based on the property of the Geometric Brownian Motion. Therefore, we define the Share of Downside Return Risk (SDRR) over $[t, \tau]$, denoted by $\Xi(t, \tau)$, as:

$$\Xi(t,\tau) = \begin{cases} \dfrac{\mathrm{Var}\left(\ln\left(\dfrac{S(\tau)}{S_t}\right)\right) - \lambda \mathrm{Var}\left(\ln\left(\dfrac{S(\tau)}{S_t}\right) \mid C = 1\right)}{\mathrm{Var}\left(\ln\left(\dfrac{S(\tau)}{S_t}\right)\right)}, & \sigma_1 < \sigma_2 \\ \dfrac{\mathrm{Var}\left(\ln\left(\dfrac{S(\tau)}{S_t}\right)\right) - (1-\lambda)\mathrm{Var}\left(\ln\left(\dfrac{S(\tau)}{S_t}\right) \mid C = 2\right)}{\mathrm{Var}\left(\ln\left(\dfrac{S(\tau)}{S_t}\right)\right)}, & \sigma_1 > \sigma_2. \end{cases} \tag{3.40}$$

By working out the expression for each term and simplifying the result, we arrive at:

$$\Xi(t,\tau) = \begin{cases} \dfrac{(1-\lambda)\sigma_2^2 + \dfrac{1}{4}\lambda(1-\lambda)\left(\sigma_1^2-\sigma_2^2\right)^2}{\lambda\sigma_1^2 + (1-\lambda)\sigma_2^2 + \dfrac{1}{4}\lambda(1-\lambda)\left(\sigma_1^2-\sigma_2^2\right)^2}, & \sigma_1 < \sigma_2 \\ \dfrac{\lambda\sigma_1^2 + \dfrac{1}{4}\lambda(1-\lambda)\left(\sigma_1^2-\sigma_2^2\right)^2}{\lambda\sigma_1^2 + (1-\lambda)\sigma_2^2 + \dfrac{1}{4}\lambda(1-\lambda)\left(\sigma_1^2-\sigma_2^2\right)^2}, & \sigma_1 > \sigma_2. \end{cases} \tag{3.41}$$

The SDRR measures the contribution to the risk in stock return by the downside component that is implied by option prices. Conversely, it also shows the impact of the downside component on option values, which is determined by both the downside variance as well as the difference between the two component variances.

In our analysis, we will implement the MLN-2 model as a representative of the MLN class for its numerical simplicity and its ease of interpretation. Model fitness for option valuation will be comparatively analyzed. In theory, the MLN-2 model should outperform the benchmark model, since the latter is an instance of the former recoverable by pushing the parameter λ to either of its boundaries. The degree of such outperformance in practice is of interest to investigate.

3.5 Applications and numerical results

This section presents the applications of the open phase allowance price models in this chapter, followed by numerical implementations of allowance option valuations using real market data.

3.5.1 Model applications

Emission allowance under an open trading phase is fundamentally different from that under a closed trading phase due to its non-lapsing feature. In the current global emission economy, the majority of the established ETS are in open-trading phases where the markets

are relatively stable and liquid. As a result, it is reasonable to view emission allowance as an alternative asset class contributing to the diversification of investment portfolios. This significantly broadens the scope of application of the models, which can be implemented to serve the followings tasks under open trading phase settings:

1. Pricing, valuation, and risk management of allowance options and related derivatives.
2. Generation of allowance price scenarios using calibrated parameters.
3. Pricing, valuation, and hedging of guarantee riders in variable annuities backed by allowance-based funds.

In addition to option valuation runs that the models are intended to serve, the calibrated model of choice can be used to price options that have limited active quotes. This often happens in the OTC market where the strike-maturity combination of interest serves particular risk management purposes. Numerical efficiency is warranted by having closed-form valuation expressions, from which the option Greeks are easily derived for the design and implementation of proper hedging strategies in order to manage the market risk exposures in option positions.

Moreover, the allowance price processes in the models can be easily discretized to simulate price scenarios that serve a wide range of actuarial calculations. Risk neutral allowance price scenarios are generated using parameters calibrated to market option prices, which can be subsequently used to price exotic contracts that lack a closed form valuation expression. On the other hand, real world allowance price scenarios are generated using parameters estimated from the spot market data, which is heuristically achieved using methodologies such as the conditional Maximum Likelihood Estimation. These scenarios are then used by insurers exposed to the allowance market for statutory capital calculations, cash flow testing, and risk management purposes. Note that the exposure may reside in both the asset side (e.g. having allowances in the asset portfolio) and the liability side (e.g. crediting a contractual rate indexed to allowance prices).

Finally, the models can be used for the pricing and valuation of common benefit guarantee riders in the variable annuity (VA) products backed by allowance-based funds[6]. Joint specifications of the allowance price model and other actuarial assumptions such as mortality,

[6]While details of these riders will be introduced in the next Chapter, at this point, it suffices to know that they are effectively put options sold to the policyholders to guarantee a minimum level of future benefits regardless of the actual performance of the underlying funds that the premium is invested in.

lapse, utilization, and withdrawal rates are required. In general, since many assumptions are dependent on the allowance price (e.g. dynamic lapse), the pricing and valuation of the guarantee liabilities are performed using scenario simulations as described above. However, under certain simplifications to the decrement and policyholder behavior assumptions, the guarantee values can be expressed in closed-form as functions of put option values from the selected allowance price model. By extension, this also leads to numerical efficiency in both liability hedging and business projection runs, where costly inner loop scenarios are otherwise required through a nested stochastic framework. Note that the actual degree of simplifications required vary by guarantee riders.

In the following sections, we provide numerical demonstrations of allowance option valuation for the three models presented in this chapter, which has not been performed in the existing literature. Additional discussions on the other two applications of the models are provided in the subsequent chapter.

3.5.2 Data and methodology

Due to various limitations, intra-day allowance option quotes are not available for our study. As an alternative approach, we manually collected and organized the EU ETS Phase 3 weekly closing call option prices for two allowance futures contracts settling in the Decembers of 2018 and 2019, which are labeled as EUA 2018 and EUA 2019 respectively for convenient reference. Details of these selected futures and option contracts are summarized in Table 3.2 below.

Futures	ICE Ticker	Settlement	Option Maturity
EUA 2018	CKZ18	Dec 17, 2018	Dec 12, 2018
EUA 2019	CKZ19	Dec 16, 2019	Dec 11, 2019

Table 3.2: Contractual features of the selected futures and option contracts

For a systematic analysis, the information is gathered and aggregated over a period of five months.

The raw data are collected at Friday market closings from online databases reflecting ICE price quotes for a study period of November 17, 2017 to April 6, 2018, with the exclusion of November 25, 2017 and December 22, 2017 due to missing data points. To construct our

sample, we identify a strike price range associated with the fewest number of noises, which defines the option data selected for the implementation. Having a common set of strikes for the option data also facilitates the analysis of the numerical results across the strike price dimension. The entries are further cross-validated to ensure accuracy and reliability. We do not exclude any low-liquidity options from the sample set as their prices still carry fair information from the market. The option prices are then matched with the corresponding weekly closing futures price to create our final sample, where each row of data (i.e. a record) is a combination of an allowance future option price, the underlying allowance futures price, the option strike, the option maturity day, and the futures settlement day. Note that the futures settlement is for an identification purpose only, as it does not enter into the option valuation expressions for any of the three models presented. The strike prices range from €3 to €16 in increments of €0.5, resulting in 54 data records for each observation day. Finally, the risk free rate r is specified to be 0%, which is the European Union Central Bank rate[7] over the study period[4].

There are two advantages of using allowance futures option data as our sample of choice. First, both options as well as the underlying futures are standardized contracts actively traded in global marketplaces including the European Climate Exchange (ECX) and the Intercontinental Exchange (ICE), where the high liquidity guarantees both the availability and quality of the price data for model calibration. Second, it minimizes the undue influence of subjectively specifying the risk free rate, which as one can verify only affects the futures option values from a discounting perspective for all three models. On the other hand, the disadvantage of this choice is that it results in a limited number of strike-maturity combinations for a comprehensive investigation. As an exchange-traded instrument, each futures contract is associated with only one available option maturity (i.e. 5 days before the futures settlement). This precludes a static analysis of model performance over the maturity dimension.

The sample is divided into a training and testing data set. The training set consists of sample data collected two weeks apart starting on November 17, 2017 (i.e. data from November 17, 2017, December 1, 2017... April 4, 2018), while the testing set consists of the rest. Notice that both days of missing data fall under the testing set to avoid biasing the model calibration results. A summary statistic of the sample data is given in Table 3.3 below, where the moneyness of an option is defined as the logarithm of the ratio between

[7]One may also consider using the average short-term AAA-rated bond yield for the EU region for the study period as the risk free rate, which nonetheless would lead to a negative value that distorts the model settings.

Data Set	Training		Testing	
Futures Contract	EUA 2018	EUA 2019	EUA 2018	EUA 2019
No. of Data Points	297	297	189	189
Total Sample Points	594		378	
Option Prices				
Average	2.48	3.08	2.19	2.79
Min/Max	0.18/10.07	0.56/10.19	0.16/8.25	0.56/8.34
Moneyness				
Average	0.0689	0.0763	0.0429	0.0496
Min/Max	-0.7971/1.4656	-0.7875/1.4724	-0.8027/1.3146	-0.7943/1.3209

Table 3.3: Summary statistics of the sample used for the implementation of open-phase allowance valuation models

the underlying allowance futures price and the strike price:

$$\text{moneyness} = \ln\left(\frac{F(t,\tau')}{K}\right). \tag{3.42}$$

In our implementation, the models are calibrated to the training set by minimizing the sum of squared differences between the options' model prices and the corresponding market observations, where the search of the optimal parameter set within the prescribed domains are performed using a grid pattern search in Matlab. The parameter estimates obtained are then used to calculate the model prices for the options in the testing set. Both in-sample and out-out-sample fitness are comparatively analyzed between the models.

In addition, the estimates of the models-specific parameters and metrics are examined. For the SKM model, these include the skewness and kurtosis coefficients (i.e. μ_3 and μ_4), the magnitudes of which are indicative of the deviations in option-implied return distributions from the LN model specifications. Similarly, for the MLN-2 model, we focus on the estimates of the mixture weight λ, the differences between the component volatilities, and the Share of Downside Return Risk. Unfortunately, the lack of intra-day data precludes the analysis on the statistical significance of these results.

Finally, note that in theory, both the SKM and MLN-2 models should perform at least as well as the benchmark model, which is essentially a special instance of the two. This numerical exercise aims at developing a better understanding of the level of outperformance in model fitness (if any) for an allowance option valuation.

3.5.3 Implementation and analysis

For a model calibration, index the records in the training set by $i \in \{1, 2, 3..., 594\}$. Let C_i^m be the actual market price of the allowance futures option in record i with a corresponding model price of:

- C_i^{LN} for the the Lognormal allowance price model, or
- C_i^{SKM} for the Skewness-Kurtosis Modified Lognormal allowance price model, or
- C_i^{MLN2} for the Two-component Mixture Lognormal allowance price model,

where C_i^{LN}, C_i^{SKM}, and C_i^{MLN2} follow Propositions 3.2, (3.21), and (3.38) respectively under the information in record i. The models are subsequently calibrated to the training set by solving the following minimization problems:

- For the LN model:

$$\begin{aligned} \underset{\sigma}{\text{minimize}} \quad & \sum_{i=1}^{594} \left(C_i^{LN} - C_i^m\right)^2 \\ \text{s.t.} \quad & \sigma > 0. \end{aligned}$$

- For the SKM model:

$$\begin{aligned} \underset{\sigma,\mu_3,\mu_4}{\text{minimize}} \quad & \sum_{i=1}^{594} \left(C_i^{SKN} - C_i^m\right)^2 \\ \text{s.t.} \quad & \sigma > 0. \end{aligned}$$

- For the MLN-2 model:

$$\begin{aligned} \underset{\lambda,\sigma_1,\sigma_2}{\text{minimize}} \quad & \sum_{i=1}^{594} \left(C_i^{SKN} - C_i^m\right)^2 \\ \text{s.t.} \quad & \sigma_1 > 0, \\ & \sigma_2 > 0, \\ & \lambda \in [0, 1]. \end{aligned}$$

Model	Parameter Estimates		
Lognormal allowance price (LN)	σ		
	0.5094		
Skewness-Kurtosis Modified Lognormal (SKM)	σ	μ_3	μ_4
	0.5057	-0.0804	3.4564
Two-component Mixture Lognormal (MLN-2)	λ	σ_1	σ_2
	0.7278	0.3848	0.8261

Table 3.4: Parameter estimates of the open-phase models from futures option prices

The LN model is clearly the simplest to calibrate, since it is associated with only one unknown parameter, making the search uni-dimensional. Both the SKM and MLN-2 models have three parameters that must be calibrated to the training data. The resultant parameters estimates from this exercise are given in Table 3.4. The volatility estimate from the LN model is 0.5094, indicating a relatively high level of risk in allowance returns implied by the option price. For the SKM model, the volatility parameter estimate is very close to that under the LN model, which is expected since the SKM model adjusts for excess skewness and kurtosis only. The skewness parameter μ_3 takes a slightly negative value, which is consistent with our previous comments on the open-phase allowance return characteristics hypothesizing an immaterial skewness effect based on Table 3.1. However, to our surprise, the kurtosis parameter μ_4 takes a value just over 3, which is far from that in Table 3.1 to conclude a significant effect of excess kurtosis. This may be attributed to the option-implied return distribution having a thinner tail over our study period.

Finally, for the MLN-2 model, the two component volatilities differ drastically, characterizing a distinct pair of return regimes that the allowance price may follow. λ is estimated to be 0.7282, which is far from either of its boundary values to support the validity of the mixture model specification. The corresponding Share of Downside Return Risk (SDRR) is 0.6498 calculated using (3.41). This indicates that over the study period, the downside component contributes to about 65% of the return variance on average as implied by the option prices. Accordingly, the crash-o-phobia effect is found to be material[8].

The in-sample performances of the models are summarized in Table 3.5. To assess model fitness, three discrepancy metrics are calculated, namely the Mean Square Error (MSE),

[8]Note that ideally, the model is best calibrated to the intra-day high-frequency price data to obtain a time series of parameter estimates and the SDRR metric, which will deliver a more comprehensive insight on how the option market's view varied over the study period.

Model	LN	SKM	MLN-2
MSE	0.006249	0.005337	0.004604
MAE	0.06761	0.06124	0.05771
MAPE	4.562%	4.244%	3.793%

Table 3.5: In-sample performance metrics for the open-phase models

Model	LN	SKM	MLN-2
MSE	0.005169	0.004405	0.004022
MAE	0.05766	0.05111	0.04785
MAPE	3.901%	3.682%	4.282%

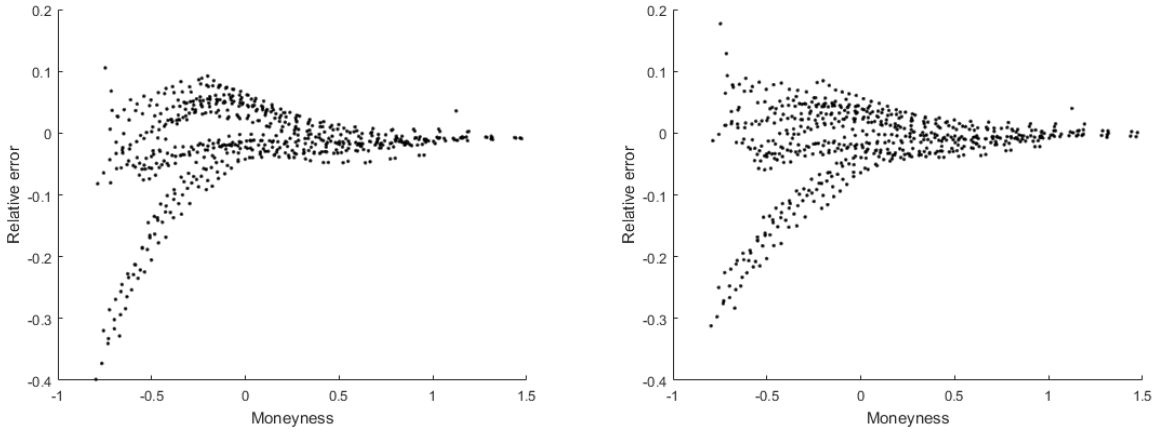
Table 3.6: Out-of-sample performance metrics for the open-phase models

the Mean Absolute Error (MAE), and the Mean Absolute Percentage Error (MAPE). The results show an overall satisfactory in-sample fit for all three models. As expected, both the SKM and the MLN-2 model outperform the LN model. At the same time, the MLN-2 model delivers recognizably more accurate prices than the SKM model. This conclusion is consistent across the metrics, which can be attributed to the ability of the MLN-2 model to explicitly capture the crash-o-phobia effect in the option market.

Figure 3.6 shows the relative errors of the model prices against the option moneyness for the calibration set. Observe that the LN and SKM models show a tendency to underprice the out-of-the-money options, which does not apply to the MLN-2 model. In fact, a mild tendency of overpricing toward the out-of-the-money region can be argued for the MLN-2 model based on the plot in 3.6(c). At-the-money and in-the-money options are quite accurately priced by all three models.

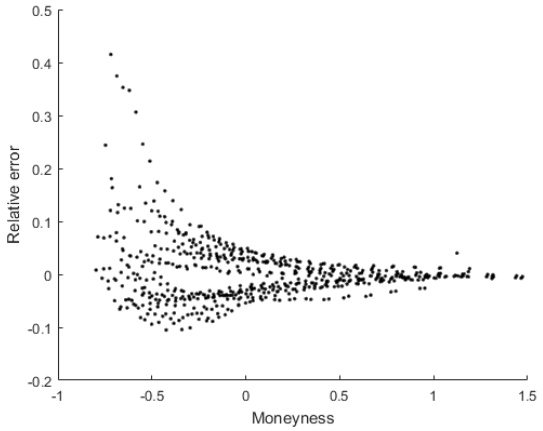
To assess the out-of-sample performances of the models, we index the records in the testing set by $j \in \{1, 2, 3..., 594\}$. For each record j, the futures option price is calculated under the models using the parameter estimates in Table 3.4 and compared to the corresponding actual market value. Fitness is quantified using the same discrepancy metrics as those in the in-sample assessment. The results are summarized in Table 3.6.

We see that all three models show good out-of-sample performance. The relative rankings in model fitness is consistent with our in-sample results based on the MSE and MAE metrics. An exception resides in the MAPE metrics, based on which the MLN-2 model underperforms the SKM and the LN. This is caused by the model's overpricing of a series of



(a) LN

(b) SKM



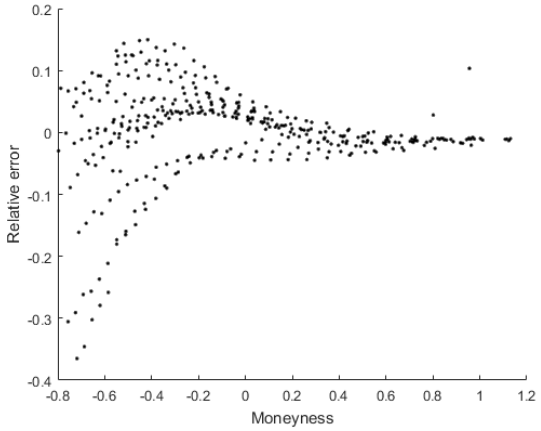
(c) MLN-2

Figure 3.6: In-sample relative error vs. Option moneyness

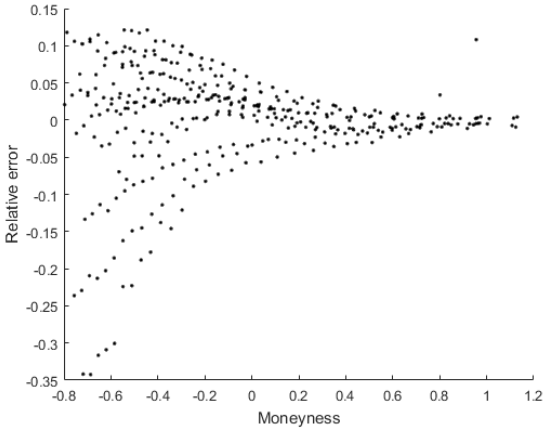
out-of-the-money options on the EUA 2018 futures on January 8, 2018, where the relative errors are inflated by extremely low market option prices.

Figure 3.7 shows the relative errors of the model prices against the option moneyness for the testing set. There is more dispersion in the relative errors for the out-of-the-money region, while the overall patterns observed are consistent with the in-sample results. In addition, the MLN-2 model's tendency to overprice out-of-the-money options is clearly re-

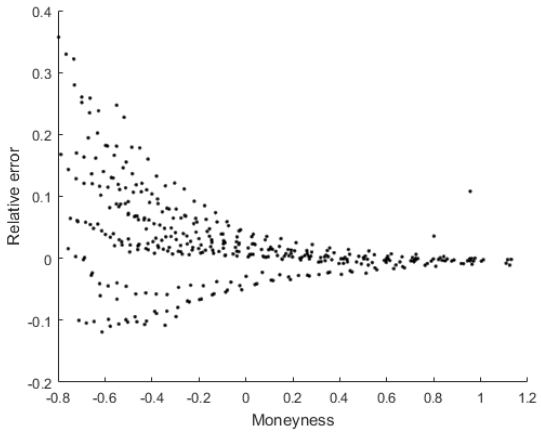
vealed here, reinforcing our previous argument made based on the in-sample information.



(a) LN



(b) SKM



(c) MLN-2

Figure 3.7: Out-of-sample relative error vs. Option moneyness

We conclude this chapter with a few remarks on the models in connection with the insights obtained from this study. First, the presented open-phase allowance price models offer closed-form option valuation expressions for superior numerical efficiency in the various actuarial applications they may serve. This is often an under-emphasized criterion in academic researches in assessing models. In addition, the selection of the model should

be based on the particular application it is intended to serve. For example, though the MLN-2 model has the best overall fitness, it also has a drawback of overpricing the out-of-the-money options. This is undesirable if the valuation results are used to set transaction upper bounds in establishing long positions in the options. However, it may be an acceptable solution for an insurer in calculating the required capital of an allowance-indexed liability with option-like payoff to the counterparty[9]. Similar arguments can be made for the SKM and LN models. There is not a perfect-for-all choice. The conclusions from this study should offer useful insights to possible stakeholders, such as insurers and pension funds, on the tradeoffs involved. Finally, for insurance regulators, the theoretical and numerical findings on the models should hopefully shed light on the development of guidelines for allowance-linked insurance and annuity products, which is not explicitly covered in scope by existing policies.

[9]This may include, for example, rate credited to policyholders of allowance-indexed fixed annuities.

Chapter 4

Allowance-linked annuities: Design and modeling

In this chapter, we present the design and modeling of annuity products with benefit payoffs that are linked to the performance of emission allowance markets. To be consistent with the rest of the thesis, the emission allowances in this chapter refer to the EUA contracts in the prevailing EU ETS Phase 3 directives, which encompasses the key characteristics of most contemporary ETS in open trading phases. Accordingly, the allowances are viewed as an alternative investment class that can be considered for long term asset allocation. In addition, unless otherwise specified, the term"annuities" refers to deferred annuities only. Immediate annuities, which are also commonly referred to as payout annuities, are not of interest in this analysis.

Annuity products often comprise substantial portions of the business for major insurance companies. Aside from providing natural hedges to the insurance policies underwritten by the insurers[1], their popularity largely resides in the demand side by providing a range of tax-deferring investment vehicles for retirement planning, catering to the market's diversity in risk appetites. However, to the best of our knowledge, annuities linked to the emission allowance market have not been launched to date. One possible reason may be the concerns around the political stability underlying the ETS schemes from a long run perspective. With the prevailing coordinated international effort in emission control, evi-

[1]Annuities are living benefit products whose payoffs are contingent on the survival of the policyholder. They introduce longevity risk that offsets in nature the mortality risk embedded in the insurance products sold.

denced by the ratification of the Paris Agreement[2] and the rapid growth of several leading ETS markets, an in-depth analysis of such annuities is called for, as they attract policyholders seeking further portfolio diversification via thematic investments aligned with the global emission control incentives. This chapter contributes to the literature by providing forward-looking insights to insurers, pensioners, and regulators on the conceptual design, viability, and modeling approaches of these relatively innovative products.

To facilitate the discussion, we first review and standardize the fundamental features of a fixed annuity. At the policy inception, the policyholder pays a premium to the insurer. The premium adds to the general account asset on the insurer's balance sheet, which follows a specified investment and reinvestment strategy. Simultaneously, an account is set up for the policyholder with an initial value equal to the premium. Each year, the account is credited a rate equal to the insurer's general account asset return less a pricing spread, subject to a contractual minimum[3]. The policyholder account value grows over an accumulation phase, at the end of which the policyholder may elect to annuitize the account value under the benefit forms offered by the product or take a lump-sum payment. Lapses are allowed during the accumulation phase with applicable surrender charges and market value adjustments specified in the contract.

Due to insurance regulations and relevant risk-based capital requirements, the insurer's general account assets often comprises fixed income instruments and similar derivatives, with small exposures in alternative investments and commodities. The associated returns are comparatively predictable with the major risks being interest and credit risks. In order to participate in the returns of other markets (i.e. stock market), policyholders need to resort to variable annuities and index annuities, which are the subjects of discussion in this chapter.

Before proceeding further, we define the notations that will be used in this chapter:

- $[0, T]$: the accumulation phase of the policy.
- $A(t)$: the time-t value of the policyholder account given that it's still in force.

[2]The Paris Agreement is an agreement within the United Nations Framework Convention of Climate Change (UNFCCC), dealing with the greenhouse gas emission mitigation, adaptation, and finance, signed in 2016. As of March 2019, 195 UNFCC members have signed it, confirming their coordinated commitments to emission control stipulated within the Agreement. See Paris Agreement (2016)[84] for more details.

[3]The credited rate also differs between existing and new premium deposits, in which the latter usually earns the new money rate less a spread. As a simplification, we do not make this differentiation in this chapter.

- m: the pricing spread/fees payable to the insurer as a percentage of account value.
- K_x: the curtate future life time for a life aged x.
- ${}_tp_x^{(\tau)}$: the double-decrement survival probability for a life aged x.
- ${}_tq_x^{(l)}$ and ${}_tq_x^{(d)}$: the double-decrement exit probabilities for a life aged x, where l denotes lapse and d denotes death.

The time index t is measured in the time step of interest that will be identified in the context. Notice that mortality and lapse rates are treated as deterministic assumptions in this analysis. This approach is consistent with that in Hardy (2003)[50], which is based on the principle that relative errors in these assumptions diminish with a sufficient diversification in product sales. Nevertheless, common actuarial practices in setting these assumptions for different products will be briefly mentioned to facilitate the readers' understanding. Finally, notations related to the allowance market prices in the previous chapters are inherited for use in this chapter.

4.1 Allowance-based variable annuities

A variable annuity (VA) differs fundamentally from a fixed annuity contract in that the premium does not become part of the general account asset of the insurer. Instead, it is deposited into a separate account of the policyholder, which is then invested in a series of managed funds offered by the variable annuity following the policyholder's discretion. These funds often represent holdings in the various equity and bond markets, and hence the account value is subject to market risk. Similar to fixed annuities, at the end of an accumulation phase, the policyholder chooses to annuitize the account or receive a lump-sum. However, the existence of various benefit guarantee riders makes such a choice interesting. For the insurer issuing the variable annuity contracts, a fee income is received periodically as a percentage of the account value. This fee is the pricing spread that covers the expenses incurred by the insurer in administering the product, the cost of the guarantee rider, and a profit margin.

Without the guarantee riders, the market risk underlying the VA contract is born by the policyholder. There are little modeling needs for risk management purposes, as the liability simply equals the separate account value. It is only necessary to model the account

value process for pricing and budgeting purposes. To improve the marketability of VA products, a variety of benefit guarantee riders have been introduced to transfer portions of the risk to the insurer, which add an actuarial sophistication to the models. Commonly offered riders include:

- Guaranteed Minimum Death Benefits (GMDB): guarantees a minimum level of death benefit payable upon the policyholder's death during the accumulation phase, without which the entitlement is confined to the account value at the time of death.
- Guaranteed Minimum Maturity Benefit (GMMB): guarantees a minimum account value for the policyholder at the end of the accumulation phase.
- Guaranteed Minimum Income Benefits (GMIB): guarantees a minimum level of periodic income for the policyholder should she choose to annuitize the account value at the end of the accumulation phase, regardless of the actual account value at that point.
- Guaranteed Minimum Withdrawal Benefits (GMWB) guarantees a level of allowable periodic withdrawals for the policyholder starting at the end of the accumulation phase should she choose not to annuitize, regardless of the actual account value.

We are interested in a hypothetical VA policy backed by a managed fund in emission allowances. To make the scope of our research manageable, the following assumptions apply to this section:

1. The account value remains fully invested in the allowance-based fund. In the existing actuarial practice, assumptions of investment allocation changes are deterministic and derived from experience studies of historical policyholder behaviors, which still lacks a mathematically rigorous foundation to date.
2. The fund management expense ratio (MER) is negligible. This is reasonable since the allowance-based fund requires no active management. This assumption can also be relaxed with small adjustments to the fund value calculations in the framework.

As allowance contracts under open trading phases share many characteristics with equity investments, features of the existing VA products can be applied, with particular considerations to the modeling and design discussed below.

4.1.1 Generating account value scenarios

One critical procedure for modeling variable annuities is the generation of future scenarios for the separate account value. In practice, most actuarial calculations for VA policies are performed via numerical simulations of the account values. This is necessary in capturing the dynamic lapse rates that vary with the moneyness of the associated guarantee riders as well as for pricing and valuations that lack analytical solutions.

The value of the separate account depends on the performance of the underlying funds. Existing products usually offer actively-managed stock market funds, which are impractical to model directly at the fund level. Instead, the current best practice groups the funds into classes based on their investment themes and strategies. The return of each fund class is then related via a simple linear model to the returns of several key indices and benchmarks, which are either simulated or prescribed. This approach is consistent with the CAPM framework and is generally accepted by regulators. Unfortunately, it does not apply to an allowance based fund, which has relatively less correlation in return to the stock market. Table 4.1 below shows the linear and rank correlations between the returns of the December 2017 EUA futures and those of selected benchmark stock indices for the period from September 1, 2010 to July 31, 2017, which covers parts of Phase 2 and Phase 3 of EU ETS for a comprehensive analysis.

Index	Correlation	Spearman's Rho
Dow Jones	0.1316	0.1276
SP500	0.1321	0.1359
NASDAQ Composite	0.1275	0.1193
Euronext 100	0.1409	0.1326
SPTSX composite	0.1388	0.1551
FTSE	0.1243	0.1195
DAX	0.1320	0.1262

Table 4.1: Return correlations between EUA17 futures and selected indices

Therefore, the allowance-based fund value scenarios should be generated separately, which is also made possible by the fund being a single-asset investment free of active management. The reduced-form allowance price models introduced in Chapter 3 can be applied to serve this purpose. The procedure is straightforward once the relevant parameter estimates are available.

First consider the generation of real world scenarios. Index the set of pivot points for

the projection horizon by $t \in \{1, 2, ...n\}$, where n denotes the end of the horizon, and assume that a total of N scenarios are needed. To simulate the allowance price path under the LN price model specified by (3.1) with parameters μ and σ per time unit:

1. Generate a standard random normal variable z_1.
2. Allowance price at $t = 1$ is calculated as $S(1) = S_0 \exp(\mu + \sigma z_1)$.
3. Repeat Steps 1 and 2 for $t = 1, 2, ...n$, which completes a price scenario.
4. Repeat Steps 1 to 3 to obtain the N scenarios.

The same procedures can be followed to simulate allowance prices under the SKM model given the parameters μ, σ, μ_3 and μ_4, except that the random normal variable in Step 1 needs to be drawn from the skewness-kurtosis adjusted normal density in (3.15). No generator is naturally available for this purpose. To do so, one needs to apply the inversion method by first generating a uniform variable $u_1 \in [0, 1]$, then numerically solving for z_1 such that:

$$\int_{-\infty}^{z_1} \phi(x) \left[1 + \frac{\mu_3(x^3 - 3x)}{3!} + \frac{(\mu_4 - 3)(x^4 - 6x^2 + 3)}{4!}\right] dx = u_1. \tag{4.1}$$

Finally, to simulate the allowance price path under the MLN-2 model specified by (3.34) with parameters μ_1, μ_2, σ_1, σ_2 per unit of time and mixing weight λ:

1. Generate a uniform variable $u_1 \in [0, 1]$.
2. Generate a standard random normal variable z_1.
3. Allowance price at $t = 1$ is calculated as:

$$S(1) = \begin{cases} S_0 \exp(\mu_1 + \sigma_1 z_1), & u_1 \leq \lambda \\ S_0 \exp(\mu_2 + \sigma_2 z_1), & u_1 > \lambda. \end{cases}$$

4. Repeat Steps 2 and 3 for $t = 1, 2, ...n$, which completes a price scenario.
5. Repeat Steps 1 to 4 to obtain the N scenarios.

The above procedures generate the allowance price scenarios under $\mathcal{P}$, which are mostly used for budgeting and cashflow testing purposes. To generate scenarios under $\mathcal{Q}$, simply replace all parameters by their risk neutral versions, where all drift parameters (i.e. μ, μ_1 and μ_2) are substituted by the risk free rate r.

Fees for VA contracts are collected at the beginning of periods in practice. Hence, we have:

$$\begin{aligned} A(t) &= A(t-1)\frac{S(t)}{S(t-1)}(1-m) \\ &= A_0\frac{S(t)}{S_0}(1-m)^t, \quad \forall t \in \{1,2,...n\}. \end{aligned} \tag{4.2}$$

Conditioning on the policy remaining in-force at each step, the account value scenarios are then derived from the simulated allowance price scenarios using (4.2).

4.1.2 Design, pricing, and valuation of guarantees

In this section, we examine the benefit guarantee riders that can be offered to the VA policyholders, which have been listed previously. Pricing in this case refers to the calculation of a rider charge δ that is drawn from the fee m to finance the future liability under the guarantee. On the other hand, valuation refers to the calculation of the actuarial present value of the payoffs provided by the guarantee rider. If the same assumption basis is used for the two calculations, then, as we shall see, valuation is just a step in the pricing process.

Despite the different protections offered by the riders, a common contractual feature dictating the guarantee payoff upon its exercise is the guaranteed benefit base, denoted by $B(t)$ without loss of generality. Absent of a ratchet feature, $B(t)$ is deterministic and may include the following forms:

- Level guarantee at the initial deposit: $B(t) = A_0$
- Level guarantee at a fixed amount: $B(t) = G$.
- Initial deposit accumulated at a specific rate, for example: $B(t) = A_0(1+r)^t$.

Complexity arises with a ratchet feature, where $B(t)$ is set to the high watermark of the fund value in $[0, t]$. Details of the benefit base will be more naturally revealed in the subsequent discussions. Finally, in VA rider designs, a partial withdrawal from the account value prior to the exercise of the rider leads to a proportional decrease in the benefit base. If an amount of $D(t)$ is withdrawn at time t, where $D(t) < A(t)$, the benefit base is reduced as following:

$$B(t^+) = B(t)\frac{A(t) - D(t)}{A(t)},$$

where the superscript + indicates the time instance immediately after. To avoid such behavioral sophistication in modeling, we assume no partial withdrawals during the accumulation phase.

A **GMMB** is the simplest rider among the ones introduced. By purchasing a GMMB, the policyholder is entitled to an account value of $\max(A(T), B(T))$ at T given the policy remains in-force. The advantage of GMMB is that the policyholder has a full discretion on the treatment of this terminal benefit (annuitize, lumpsum, transfer, etc.), which is not bound by the rider. Denote by $V_{AB}(t)$ the time t value of the GMMB rider. The terminal payoff of the rider conditioning on the policy remaining in force is given by:

$$V_{MB}(T) = (B(T) - A(T))^+ \tag{4.3}$$

Hence, without a ratchet feature, the GMMB is a European put option written on the account value with strike $B(T)$ maturing at T, whose risk neutral value at $t \in [0, T]$ is given by:

$$\begin{aligned} V_{MB}(t) &= {}_{T-t}p^{(\tau)}_{x+t} e^{-r(T-t)} E\left[(B(T) - A(T))^+\right] \\ &= {}_{T-t}p^{(\tau)}_{x+t} M(T) P(t, B^*(T), T), \end{aligned} \tag{4.4}$$

where

$$M(T) = \frac{A_0}{S_0}(1-m)^T, \tag{4.5}$$

$$B^*(T) = \frac{B(T)}{M(T)}. \tag{4.6}$$

The last line of (4.4) follows from (4.2), the result of which shows that a GMMB is equivalent to a scalar multiple of an allowance put option maturing at T. Pricing is performed at $t = 0$ by setting the actuarial present value (APV) of the insurer's rider charge income

(denoted by $I(0)$) to the GMMB value under $\mathcal{Q}$. We have:

$$\begin{aligned} I(0) &= E\left(\sum_{t=0}^{K_x} e^{-rt}\delta A(t)\right) \\ &= \delta A_0 \sum_{t=0}^{T-1} {}_tp_x^{\tau}(1-m)^t \\ &= \delta A_0 \ddot{a}_{x:\overline{T}|m'}, \end{aligned} \tag{4.7}$$

where $\ddot{a}_{x:\overline{T}|m'}$ is the APV of a T-year annuity due for age x and interest rate m' defined by

$$m' = \frac{m}{1-m}.$$

This leads to the pricing expression:

$$\delta = \frac{V_{MB}(0)}{A_0 \ddot{a}_{x:\overline{T}|m'}}. \tag{4.8}$$

A **GMDB** is slightly more difficult to value and price than GMMB in that the rider payoff point is variable. For a VA policy with GMDB, a policyholder aged x is entitled to a death benefit of $\max\left(A(K_x+1), B(K_x+1)\right)$ at point $K_x+1 \in [0,T]$. The corresponding rider payoff is given by

$$V_{DB}(K_x+1) = \left(B(K_x+1) - A(K_x+1)\right)^+, \tag{4.9}$$

where $V_{DB}(t)$ denotes the time-t value of the GMDB rider. Without a ratchet feature, the GMDB is a European put option written on the account value with strike $B(K_x+1)$ maturing at K_x+1 given $K_x < T$. Its risk neutral value at $t \in [0,T]$ is:

$$\begin{aligned} V_{DB}(t) &= E\left(e^{-r(K_x+1-t)}\left[\left(B(K_x+1) - A(K_x+1)\right)^+\right]\right) \\ &= \sum_{k=0}^{T-t-1} {}_{k|}q_{x+t}^{(d)} M(k+1) P(t, B^*(k+1), k+1), \end{aligned} \tag{4.10}$$

where $M(\cdot)$ and $B^*(\cdot)$ follow (4.5) to (4.6). This result shows that a GMDB is equivalent to a linear combination of allowance put options with sequentially increasing strikes and maturities. Pricing of the rider is performed using the same logic as that for GMMB.

Note that the APV of the rider fee income is the same across riders, leading to the pricing expression:

$$\delta = \frac{V_{DB}(0)}{A_0 \ddot{a}_{x:\overline{T}|m'}}, \tag{4.11}$$

where $V_{DB}(0)$ is calculated from (4.10).

A **GMIB** differs from the previous two riders by guaranteeing the policyholder a minimum level of periodic income upon annuitization of the account value at T, subject to a maximum payout term n. The guaranteed payment equals a fixed percentage c of the benefit base $B(T)$. Assume that annuity payments without GMIB utilization are determined using the equivalence principle, the payoff of a GMIB rider at T is:

$$V_{IB}(T) = \left(cB(T)\ddot{a}_{x+T:\overline{n}|} - A(T)\right)^+, \tag{4.12}$$

where $V_{IB}(t)$ denotes the time-t value of the GMIB rider. Without a ratchet feature, the GMIB is a European put option written on the account value with strike $cB(T)\ddot{a}_{x+T:\overline{n}|}$ maturing at T conditioning on the policy remaining in-force, whose risk neutral value at $t \in [0, T]$ is:

$$\begin{aligned} V_{IB}(t) &= {}_{T-t}p^{(\tau)}_{x+t} e^{-r(T-t)} E\left[\left(cB(T)\ddot{a}_{x+T:\overline{n}|} - A(T)\right)^+\right] \\ &= {}_{T-t}p^{(\tau)}_{x+t} M(T) P(t, cB^*(T)\ddot{a}_{x+T:\overline{n}|}, T), \end{aligned} \tag{4.13}$$

where $M(T)$ and $B^*(T)$ follow (4.5) to (4.6). Hence, a GMIB is also equivalent to a scalar multiple of an allowance put option maturing at T with the strike price being a product of the benefit percentage, benefit base, and an annuity factor. For pricing, the APV of the rider fee income is the same as that in the previous two cases, which gives:

$$\delta = \frac{V_{IB}(0)}{A_0 \ddot{a}_{x:\overline{T}|m'}}, \tag{4.14}$$

where $V_{IB}(0)$ is calculated from (4.13). Note that in reality, a GMIB provides protection over both market and basis change risks for the policyholder. This is because under the standard annuitization, annuity payments are calculated based on the mortality assumption at T, which typically differs from the time-0 assumption based on which the GMIB is priced.

All results above are straightforward due to the absence of the ratchet feature, an attractive design that further increases the marketability of the VA policy and riders. Under

a **Ratchet** feature, the contractual benefit base is set to the maximum account value reached over its paths, which is expressed as:

$$B(t) = \max\left\{A(t_i) \mid t_i \in \mathcal{T}, t_i \leq t\right\}, \tag{4.15}$$

where $\mathcal{T}$ is the set of ratcheting points specified in the rider, making them look-back put options written on the account values. Specifically, let $LP(t, \mathcal{T}, \kappa, \tau)$ denote the time-t value of a floating strike look-back put option maturing at τ with terminal payoff $(\kappa \max\{A(t_i) \mid t_i \in \mathcal{T}\} - A(\tau))^+$, we have:

$$V_{MB}(t) = {}_{T-t}p_{x+t}^{(\tau)} LP(t, \mathcal{T}, 1, T),$$

$$V_{DB}(t) = \sum_{k=0}^{T-t-1} {}_{k|}q_{x+t}^{(d)} LP(t, \mathcal{T}, 1, k+1),$$

$$V_{IB}(t) = {}_{T-t}p_{x+t}^{(\tau)} LP(t, \mathcal{T}, c\ddot{a}_{x+T:\overline{n}|}, T).$$

In these cases, the rider values must be calculated numerically through a simulation of the account value paths as described in Section 4.1.1, though the APV of the rider fee income retains the analytical expression in (4.7) for pricing purposes.

To summarize, without the ratchet feature, values of GMMB, GMDB, and GMIB riders can be expressed as linear combinations of allowance put option values, for which closed-form expressions under three alternative models derived in Chapter 3 may be conveniently applied. Unfortunately, no analytical solutions are available for the pricing and valuation of a GMWB, which requires a numerical procedure regardless of whether a ratchet feature is present. For brevity, we only present the product design for a GMWB rider along with the key modeling considerations.

A **GMWB** guarantees a minimum level of periodic withdrawal that a non-annuitizing policyholder may take against its account value starting at T up to a maximum term of n conditioning on the policy remaining in-force. The guaranteed withdrawal level equals a fixed percentage c_w of the benefit base $B(t)$ and may be drawn regardless of the actual account balance. Compared to the other riders, one distinguishing feature of a GMWB is that upon its exercise, the separate account remains active through the payout phase and hence the account fee continues to be collected. In addition, the policyholder has discretion over the withdrawals, where an amount over the guaranteed level is allowed up to the remaining account value or a contractual maximum, whichever is less. Such withdrawals reduce the benefit base proportionally and thus subsequent guaranteed withdrawal

levels. For instance, if an amount of $D(t) > c_w B(t)$ is withdrawn at t, the benefit base is immediately adjusted:

$$B(t^+) = B(t) \max\left\{\left(1 - \frac{D(t)}{A(t)}\right), 0\right\}, \quad \forall t \in [T, T+n]. \tag{4.16}$$

Two modeling complexities hence arise for the valuation and pricing of a GMWB:

1. A need to model the policyholder withdrawal patterns, which depends on a variety of behavioral factors.

2. A need to model the account value in both the accumulation and payout phase.

These two components should be implemented in the scenario-based numerical framework. Simulations can be performed to estimate the APV of the GMWB, which start to pay off once the account value is depleted. Pricing follows by setting the result to the APV of rider incomes. In practice, the withdrawal behaviors can be modeled using a discrete-time Markov process calibrated to experience data, with states representing withdrawals below, at, and above the guaranteed level. Finally, for a GMWB policy, the lapse rates typically vary before and after the exercise of the rider, where lower surrenders apply after the first GMWB withdrawal. The lapse also exhibits a dynamic behavior by varying on the moneyness of the rider, adding complexity to the modeling requirements.

To conclude this section, we presented the design, valuation, and pricing expressions of various guarantee riders for an allowance-based VA separate account. Depending on the type of rider and the presence of the ratchet feature, closed-form pricing and valuation results may be available under the allowance price models presented in Chapter 3. However, we remind the reader that the pricing equation derived in this section assumes that the rider charge is drawn from the account fee without increasing it. If this assumption is relaxed, a numerical procedure will be required to solve for the rider fee rate that equates the actuarial present values of the fee income and the rider payoff.

4.1.3 Hedging and volatility control fund

Hedging portfolios for the guarantee riders can be built by taking positions in the allowance and bond market, where the strategies vary by rider designs. In an ideal allowance market

where all possible option specifications are traded at fair values under negligible transaction costs, a complete hedge is possible for GMMB, GMDB, and GMIB designs without the ratchet feature, which are equivalent to linear combinations of allowance put options. Exact replicating portfolios can be built by longing put options, where the required positions and option parameters are exactly as given in (4.4), (4.10), and (4.13) respectively for the three riders. However, such a desirable setting rarely, if ever at all, holds in reality, as the allowance option market is relatively lacking in size and liquidity. As a result, a dynamic hedging portfolio can be built following the delta-hedge approach with a time-t value:

$$H(t) = \Psi(t)S(t) + \Upsilon(t), \tag{4.17}$$

where

$$\Psi(t) = \frac{\partial V(t)}{\partial S(t)} \tag{4.18}$$

represents the time-t delta of the guarantee liability, and

$$\Upsilon(t) = H(t) - \Psi(t)S(t) \tag{4.19}$$

is the bond portion of the hedge. Under continuous-time, the hedging portfolio exactly replicates the rider liability. In practice, the hedge is rebalanced in discrete time and hence hedge errors exist. In this chapter, hedge rebalancing follows the same time step as t to be defined in the context. The time-t hedging error before the rider is exercised, denoted by $HE(t)$, is given by the difference between the required hedge at t and the hedge at $t-1$ rolled forward, which has the unconditional expression:

$$HE(t) = {}_tp_x^{(\tau)} \left(\Psi(t)S(t) - \Upsilon(t)\right) + {}_{t-1}p_x^{(\tau)} \left(\Psi(t-1)S(t) - \Upsilon(t-1)e^{r\Delta t}\right), \tag{4.20}$$

where Δt is the length of the time step expressed in years. For GMMB, GMDB, and GMIB without the ratchet features, the liability deltas in (4.18) can be expressed as functions of allowance put option deltas as below, conditioning on the policy being in-force at t:

$$\Psi_{MB}(t) = {}_{T-t}p_{x+t}^{(\tau)} M(T)\Delta P(t, B^*(T), T), \tag{4.21}$$

$$\Psi_{DB}(t) = \sum_{k=0}^{T-t-1} {}_{k|}q_{x+t}^{(d)} M(k+1)\Delta P(t, B^*(k+1), k+1), \tag{4.22}$$

$$\Psi_{IB}(t) = {}_{T-t}p_{x+t}^{(\tau)} M(T)\Delta P(t, cB^*(T)\ddot{a}_{x+T:\overline{n}|}, T), \tag{4.23}$$

where the subscript denotes the rider type. An advantage of this exercise is that allowance

option deltas have analytical expressions under the LN and MLN allowance price models presented in Chapter 3, which are detailed in Appendix R. Hedge and rebalancing can be efficiently implemented without numerical procedures.

Unfortunately, for a GMWB or riders with the ratchet feature, the required positions in the hedge need to be determined numerically at each rebalancing point t following the three steps below:

1. Determine a proper shock size, denoted by ΔS to be applied to the allowance price in opposite directions.
2. Evaluate via simulations the APV of the rider liability for each shocked time-t allowance price. Denote the results by $V^u(t)$ and $V^d(t)$ for the upward and downward shocks respectively.
3. $\Psi(t)$ is estimated as:
$$\Psi(t) \approx \frac{V^u(t) - V^d(t)}{2\Delta S}. \tag{4.24}$$

Apparently, such an actuarial practice leads to substantial computational resource requirements for the inner loop runs should a full cash flow projection be needed.

We end the discussion on allowance-based VA products by introducing the Volatility Control Fund (VCF). A VCF is an actively managed fund aiming to achieve a target range of return volatility over time. It leverages the view that the allowance return volatility changes over time, as we previously pointed out under the empirical evidence in Figure 3.3. Since volatility models are out of the scope of this thesis, we present only the basic concepts and designs of the fund.

Effectiveness of volatility control in an allowance-based separate account fund is largely ensured by the simplicity of having a single risky asset to model. The volatility control mechanism is straightforward: the fund consists of dynamically rebalanced positions in the allowance market and risk free assets (e.g. treasuries and bonds), with investment weights reflecting projected allowance market volatilities under a selected model and a target volatility the fund aims to maintain. Let $F(t)$ denote the unitized time-t fund value. Assume that the allowance price follows:

$$dS(t) = rS(t)dt + \sigma(t)S(t)dW(t), \tag{4.25}$$

where $W(t)$ is a standard Brownian Motion under $\mathcal{Q}$. Note that the LN and MLN allowance price specifications are both instances of this general form. Without the fund management fee, the fund value process is given by:

$$\begin{aligned}\frac{dF(t)}{F(t)} &= w_r(t)rdt + w_s(t)(rdt + \sigma(t)dW(t)) \\ &= rdt + \sigma^* dW(t),\end{aligned} \tag{4.26}$$

where σ^* is the target volatility level. $w_r(t)$ and $w_s(t)$ denote the investment weights in treasury and the allowance markets, which change in continuous time following:

$$w_s(t) = \frac{\sigma^*}{\sigma(t)}, \tag{4.27}$$

$$w_r(t) = 1 - w_s(t) = 1 - \frac{\sigma^*}{\sigma(t)}, \tag{4.28}$$

where a negative value of $w_r(t)$ means a leveraged position in allowance. An advantage of the VCF fund is that it leads to a potentially more robust hedging portfolio, since the vega risk not captured in a delta-hedging strategy is resolved at the fund level. The downside is the heavy reliance on the model used in the volatility control mechanism as well as effectiveness in the fund management process.

In reality, the investment allocation adjustment occurs in discrete-time, for which we assume the same frequency as the time steps of t in the previous discussions. Some modifications to the volatility control mechanism are also made for implementation. Instead of a single target volatility, a target volatility range $[\sigma_{min}, \sigma_{max}]$ is established. Restrictions are also set on the investment weights in the allowance market, denoted by $[w_{min}, w_{max}]$, to prevent excessive leveraging and deleveraging in the fund composition. The volatility control mechanism follows the step below, assuming that the fund starts with a full allocation to the allowances at $t = 0$:

1. Select a volatility model and volatility control parameters described above.
2. At each point $t \geq 0$, project the volatility $\widehat{\sigma(t+1)}$ that applies through the subsequent period $[t, t+1]$. Model recalibration may be performed as needed.
3. Adjust the investment allocations for the subsequent period using the volatility con-

trol strategy, so that:

$$w_s(t+1) = \begin{cases} w_s(t), & w_s(t)\widehat{\sigma(t+1)} \in [\sigma_{min}, \sigma_{max}] \\ \min\left\{\max\left\{\dfrac{\sigma_{min}+\sigma_{max}}{2\widehat{\sigma(t+1)}}, w_{max}\right\}, w_{min}\right\}, & w_s(t)\widehat{\sigma(t+1)} \notin (\sigma_{min}, \sigma_{max}) \end{cases}. \tag{4.29}$$

The process is very intuitive: whenever the projected fund volatility for the subsequent period falls out of the target range, the investment weights are adjusted to bring it to the midpoint of the range.

Simulation of fund values for a VCF under $\mathcal{Q}$ is simple to perform. For a projection horizon of n periods, using the assumed volatility model, generate paths of volatilities and allowance prices simultaneously by discretizing (4.25). The investment weights for each period along each path are calculated based on the simulated volatilities and the volatility control mechanism above. The fund value scenarios are subsequently obtained using the paths of investments weights and allowance prices. Finally, the account value scenarios are derived from the fund value scenarios using the relationship:

$$A(t) = A_0 \frac{F(t)}{F_0}(1-m)^t, \quad \forall t \in \{1, 2, ...n\}. \tag{4.30}$$

The results can be used in the valuation, pricing, and hedging of the various guarantee riders. To simulate fund values of the VCF under $\mathcal{P}$ that are used for outer loop runs in budgeting and cash flow testing purposes, replace the drift parameter in (4.25) by the real world estimate μ. This shall affect the accumulation of the fund value over periods, but does not affect any other components of the procedure.

4.2 Allowance-based fixed index annuities

A fixed index annuity (FIA) contract can be viewed as a hybrid of a fixed and a variable annuity. Being a fixed annuity in nature, the policyholder's account value is backed by the insurer's general account asset. However, each period, the account is credited a rate that is linked to a market index (less a pricing spread), and hence gaining exposures to market returns as for a variable annuity. The credited rate is declared at the beginning

of each period subject to a contractual minimum. At the end of the accumulation phase, the policyholder may annuitize the account under the available annuity options or take the account value as a lumpsum. Lapses during the accumulation phase are subject to applicable surrender charges and market value adjustments, exactly the same as for a fixed annuity contract.

Existing FIA products commonly link their credited rates to various equity indices to provide the policyholders an exposure to the global stock markets. Contemporary studies on FIA also center around these equity-linked annuities (EIA), such as in Hardy (2003)[50] with focuses on their design and valuation. We are interested in a hypothetical FIA design where the credited rate is linked to the allowance market, with a focus on the credited rate strategy design, which differs to some extent from the existing literature.

The major risk for an insurer issuing the FIA contract is the basis risk between the actual earned rate on premium and the rate credited to the account. The former is equal to the general account portfolio return, while the latter is based on the allowance market return to which the portfolio may lack an exposure. However, the fact that the credited rate is periodically declared allows a flexible design to largely (if not completely) hedge the risk. Notice that from an actuarial perspective, the payoffs in the FIA contract resemble those of call options on the allowance. This is in contrast to the VA guarantees in the previous section, which are in nature put options. As an example, consider the simplest Point-to-Point (PTP) indexing, under which the rate credited to the policyholder account for the period $[t, t+1], t \in [0, T-1]$ is given by:

$$c(t) = \max\left\{\frac{S(t+1)}{S_t} - 1, G\right\}, \tag{4.31}$$

where G is the contractual rate guarantee. The dollar amount credited is:

$$A(t)C(t) = GA(t) + \frac{A(t)}{S_t}\left(S(t+1) - (1+G)S_t\right)^+. \tag{4.32}$$

The variable portion of the expression is equivalent to the payoff of a time-t allowance call option with a strike price of $(1+G)S_t$ maturing at $t+1$. Therefore, the insurer selling the FIA contract has a short position in an allowance call option, whose payoffs are theoretically unlimited. To mitigate such a risk, a productive approach is to add features such as cap, participation rate, and offset to the credited rate strategy. We show several designs below along with the method of determining these mitigating parameters. To facilitate the presentation, we assume that t is measured in years, so that all crediting rates parameters

are reset annually. To avoid confusion, we remind the reader that the parameters declared at t apply to year $t+1$ (i.e. the year spanning $[t, t+1]$), while the actual rate crediting for the year occurs at $t+1$ (i.e. the year-end).

The funding for the call option embedded in an FIA policy, which we refer to as the option budget, is provided by the general account return. In practice, the general account portfolio consists of mostly well-managed investment grade instruments that deliver a relatively predictable return on a periodic basis[4]. Denote by $C_g(t)$ the time-t projected general account return for year $t+1$, then the corresponding option budget, denoted by $O(t)$, is obtained by subtracting from the projected return a pricing spread:

$$O(t) = C_g(t) - m. \tag{4.33}$$

The option budget is the key variable in setting the mitigating parameters.

Annual Point-to-Point with Cap: A cap is the most effective risk mitigating parameter declared at the beginning of each period, which puts an upper bound on the credited rate over the period. When it is applied to the simple Annual PTP indexing, the credited rate over year $t+1$ is

$$c(t) = \min\left\{\max\left\{\frac{S(t+1)}{S_t} - 1, G\right\}\kappa\right\}, \tag{4.34}$$

where $\kappa > G$ represents the cap rate. The corresponding dollar amount credited is given by:

$$GA(t) + \frac{A(t)}{S_t}\left((S(t+1) - (1+G)S_t)^+ - (S(t+1) - (1+\kappa)S_t)^+\right). \tag{4.35}$$

The variable part of this payoff is equivalent to that of a bull spread consisting of a long call with strike $(1+G)S_t$ and a short call with strike $(1+\kappa)S_t$. To set the declared cap at time t, we find the available prices of allowance call options maturing at $t+1$ on the market, and solve for the value of κ such that the market cost of (4.35) equals the dollar amount of the option budget:

$$O(t) - G = \frac{1}{S_t}\left(C^m(t, (1+G)S_t, t+1) - C^m(t, (1+\kappa)S_t, t+1)\right), \tag{4.36}$$

where the superscript m indicates the prices are market-quoted. In practice, a precise solution can rarely be found in the traded strikes. In this case, κ may be solved for under a selected allowance price model or via a linear interpolation. The latter approach

[4]At the beginning of period, the return over the period can be projected with a satisfactory level of accuracy, due to the nature of the general account portfolio.

retains the market-consistency of the method but ignores the convexity of option values with respect to the strike price, which is trade off to ponder over. In addition, the right hand side of (4.36) represents the spread positions needed to hedge the basis risk between the earned and credited rates for the FIA policy, which can also be done using allowance futures options with minor adaptations. Futures options are exchange-traded with high liquidity, where maturities in annual intervals are largely available.

Annual Point-to-Point with Participation: A participation rate limits the level of market indexing by applying a multiplier to the allowance market return in the credited rate. Compared to a cap, it is less effective in risk mitigation in that the credited rate remains unbounded above. When it is applied to the simple Annual PTP indexing, the credited rate over year $t+1$ is

$$c(t) = \max\left\{\alpha\left(\frac{S(t+1)}{S_t} - 1\right), G\right\}, \tag{4.37}$$

where $\alpha \in [0,1]$ denotes the participation percentage. The associated dollar amount credited for the period is:

$$GA(t) + \frac{\alpha A(t)}{S_t}\left(S(t+1) - \left(1 + \frac{G}{\alpha}\right)S_t\right)^+. \tag{4.38}$$

The variable portion of this payoff is equivalent to that of an scaled out-of-the-money allowance call option maturing at $t+1$. To set the declared participation rate at t, the same logic above is followed by solving for the α that equates the market cost of (4.38) to the option budget in dollar amount:

$$O(t) - G = \frac{\alpha}{S_t} C^m\left(t, \left(1 + \frac{G}{\alpha}\right)S_t, t+1\right), \tag{4.39}$$

which requires some numerical search method. Once the participation rate is set, the right-hand side of (4.39) gives the required hedge to be built for the year.

Annual Point-to-Point with Offset: An offset represents a direct reduction to the indexed market return in the calculation of the credited rate. Being the simplest in design, it is less effective than a cap as a risk mitigation since the credited rate with an offset is unbounded above. When it is applied to the PTP indexing, the credited rate over year $t+1$ is

$$c(t) = \max\left\{\alpha\left(\frac{S(t+1)}{S_t} - 1 - \theta\right), G\right\}, \tag{4.40}$$

where θ is the offset rate, with a corresponding dollar amount credited:

$$GA(t) + \frac{A(t)}{S_t}\left(S(t+1) - (1+\theta+G)S_t\right)^+. \tag{4.41}$$

The variable portion of this payoff is equivalent to that of an out-of-the-money allowance call option maturing at $t+1$ with strike $(1+\theta+G)S_t$. Exactly the same logic above is used in setting the declared offset rate at t, by solving for the θ equating the market cost of (4.41) to the option budget in dollar amount:

$$O(t) - G = \frac{1}{S_t} C^m(t, (1+\theta+G)S_t, t+1). \tag{4.42}$$

Similar to the cap and participation rate, the right-hand side of (4.42) gives the simple holdings required to hedge the indexed payoff for the year.

We have demonstrated the design, calculation, and hedging required when the mitigating parameters are applied to a Point-to-Point indexing strategy, the most commonly offered mechanism in the market of equity-linked annuities. However, any combinations can be built by applying the mitigating parameter to other indexing strategies, which include:

- Monthly Average Price Return, with a credited rate for year $t+1$:

$$c(t) = \max\left\{\frac{\sum_{j=1}^{12} S(t+\frac{j}{12})}{12S_t} - 1, G\right\}.$$

- Monthly Maximum Price Return, with a credited rate for year $t+1$:

$$c(t) = \max\left\{\frac{\max\left\{S(t+\frac{j}{12}) \mid j = 1, 2..12\right\}}{S_t} - 1, G\right\}.$$

- Point-to-Point Binary Return, with a credited rate for year $t+1$

$$c(t) = \begin{cases} \gamma, & \dfrac{S(t+1)}{S_t} - 1 > G \\ G, & \text{otherwise.} \end{cases} \tag{4.43}$$

  where $\gamma > G$ is reset annually.

One should observe that these indexing strategies correspond to different types of call option payoffs, namely a variable price Asian call, a floating price look-back call, and a binary call. Adding a participation rate or offset to these increases the strike price, while adding a cap transforms the liability to the corresponding bull spread, all of which reduces the liability value to the insurer. Unfortunately, these exotic allowance options are not actively offered. Therefore, the determination of the parameter values to be declared requires a proper allowance price model. Hedge portfolios change accordingly to follow a model-contingent delta-hedging strategy.

Finally, note that we have not discussed the hedging for the guaranteed minimum credited rate G. This is because for an FIA contract, the guaranteed interest credit is comfortably low compared to the general account return to warrant its funding. Under the prevailing low rate economy, a value of G in the range of 0% to 0.5% is common. In addition, a FIA contract typically has a shorter term from issuance than a VA contract, ranging from 10 to 15 years. Hence, the minimum rate guarantee is of a minor concern in practice.

4.3 Examples and numerical illustrations

In this section, we provide several examples to numerically demonstrate and supplement the concept discussed in this Chapter.

Example 4.1: Pricing of a GMWB rider.
Consider a representative VA policy with GMWB issued to a life of age 50 on May 3, 2019. The separate account remains fully invested in the allowance market fund. Initial deposit is €10000. All cash-flows occur monthly. Policy terms are listed below:

- Payout phase: starts at age 60.
- Benefit base: calculated monthly; ratchets every two years during accumulation.
- GMWB amount: 5% annual notional (0.417% monthly) of the month-start benefit base.
- Maximum withdrawal if above GMWB amount: the remaining account value or 3 times GMWB amount, whichever is less.

- GMWB term: 20 years.
- Fund Management Expense Ratio (MER): 0%.
- Account fee: 3% annual notional (0.25% monthly) of the month-start account value.
- Surrender charges and MVA: None.

Scenario generation follows the specifications below:

- Starting allowance price: €24.63 per EEX market data.
- Allowance price model: two-component Lognormal Mixture.
- Parameters: use $\lambda = 0.7$, $\sigma_1 = 0.3$, and $\sigma_2 = 0.7$ (both annual).
- Number of scenarios: 5000.

Actuarial assumptions follow the specifications below:

- Mortality table: 2012 IAM Period Male ANB. Converted to monthly under Constant Force of Mortality (CFM).
- Base lapse: 8%/year in accumulation phase and 4%/year in payout phase. Converted to monthly under Constant Force of Lapse (CFL).
- Dynamic lapse: None.
- Accumulation phase partial withdrawal: None.
- Risk free rate: 0.5% annual continuous.
- GMWB utilization: 100% utilization at age 60.
- GMWB withdrawal: follows the three-state Markov Process specified by Table 4.2 and 4.3, where the state Above, At, and Below respectively means 300%,100%, and 85% of the guaranteed withdrawal amount for the month.

	Starting Probability	Transition Probability		
	Probability	Above	At	Below
Above	3.0%	0.0%	60.0%	40.0%
At	57.0%	2.0%	93.0%	5.0%
Below	40.0%	4.0%	26.0%	70.0%

Table 4.2: GMWB withdrawal rates when AV $>$ 20% of Benefit Base

	Starting Probability	Transition Probability		
	Probability	Above	At	Below
Above	0.0%	0.0%	60.0%	40.0%
At	60.0%	0.0%	95.0%	5.0%
Below	40.0%	0.0%	30.0%	70.0%

Table 4.3: GMWB withdrawal rates when AV $<$ 20% of Benefit Base

In addition, it is assumed that each withdrawal takes place at the beginning of the month after the collection of the account fee, but before the start of accumulation of the account value.

Using the method in Section 4.1.2, the rider is priced at 0.44% per annum. The simulated APV of the rider payoffs has a mean of €3554.38 but a standard deviation of €17575.45. This high uncertainly can be attributed to the nature of a GMWB, as it does not make a payoff until the depletion of account value, the occurrence and timing of which depend on the market return and policyholder withdrawal behaviors that are both stochastically modelled. The crash-o-phobia effect in the allowance market, which called for the use of the MLN-2 generator, also contributed to this observation by delivering an increased number of tail market scenarios. Table 4.1 shows the histogram of the simulated APVs of rider payoffs, where a positive skew highlights the needs of prudent tail risk management for the product. **End of Example 4.1**.

Example 4.2: Hedging of a GMIB rider
Consider a representative VA policy with GMIB issued to a life of age 50 on May 3, 2019. The separate account remains fully invested in the allowance market fund. Initial deposit is €30000. All cash-flows occur monthly. Policy terms are listed below:

- Payout phase: starts at age 66.
- Benefit base: initial premium with a fixed 2% COLA per annum.

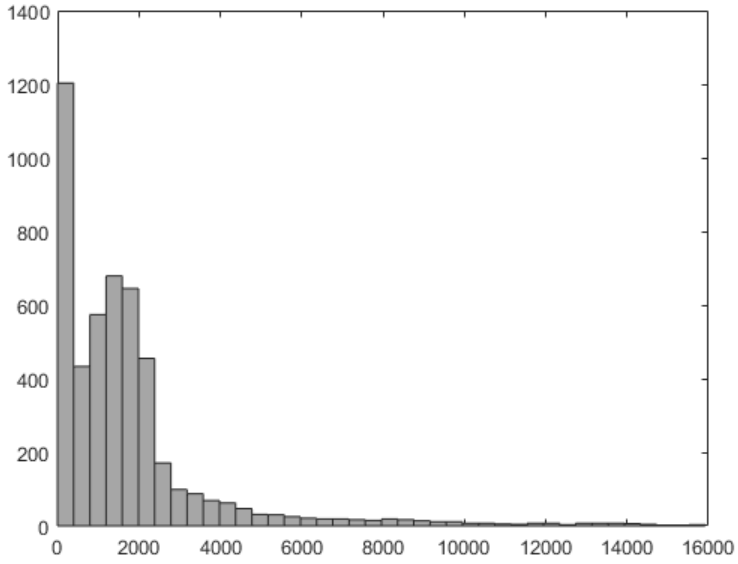


Figure 4.1: Histogram of simulated APV of GMWB payoffs

- GMIB Amount: 6% annual notional (0.5% monthly) of the benefit base.
- GMIB term: 15 years.
- Fund Management Expense Ratio (MER): 0%.
- Account Fee: 3% annual notional (0.25% monthly) of the month-start account value.
- Surrender charges and MVA: None.

Actuarial assumptions are listed below:

- Mortality table: 2012 IAM Period Male ANB. Converted to monthly double decrement under CFM.
- Base lapse: 8%/year in the accumulation phase. Converted to monthly double decrement under CFL. No lapse is assumed for the payout phase.
- Dynamic Lapse: None.

- Accumulation phase partial withdrawal: None.
- Risk free rate: 0.5% annual continuous.
- GMIB utilization: 100% utilization at age 60 if rider is in the money.
- Scenario generation: 5000 scenarios under the LN allowance price model with annual parameters $\mu = 0.08$ and $\sigma = 0.4$.

The rider is priced at 1.44% using (4.13) to (4.14) and under the LN allowance price model, which is made particularly efficient by the closed-form option value expressions. Hedging follows the strategy in Section 4.1.3., assuming the same volatility above as pricing and rebalanced monthly. We are interested in exploring:

- The distribution for the APV of hedging errors, which are expected to be symmetric around 0.
- Hedge effectiveness, measured as the average proportion of the change in rider value that is offset by the change in hedging portfolio value across scenarios.

Figure 4.2 shows the histogram of the APV of hedging errors, expressed as a percentage of the initial premium. The observations are fully consistent with our expectation, where the hedging errors are symmetrically centered around 0, displaying a bell shaped curve with no clearly identifiable skewness or heavy tails. The overall level of accuracy of the hedging is satisfactory.

The hedge effectiveness is estimated to be 87.62%, which is between 85% and 120%, the range it must fall within for the strategy to qualify for a hedge accounting. It is lower than what one may expect for a GMIB on an account based on stock-market funds, due to the high volatility assumed for allowance return. However, it may be improved by increasing the hedge rebalancing frequency. In addition, this example demonstrates the hedge mechanism from a purely economical perspective for a single policy. In practice, a full projection of income statement and balance sheet items must be performed to estimate the impact of the hedge on the reserves, surplus, and distributable earnings, which will require nested model runs. **End of Example 4.2**

Example 4.3: Cap setting for FIA credited rate
Consider a representative FIA policy linked to the allowance market return. Rate crediting occurs annually by following the annual Point-to-Point strategy with Cap reset at the

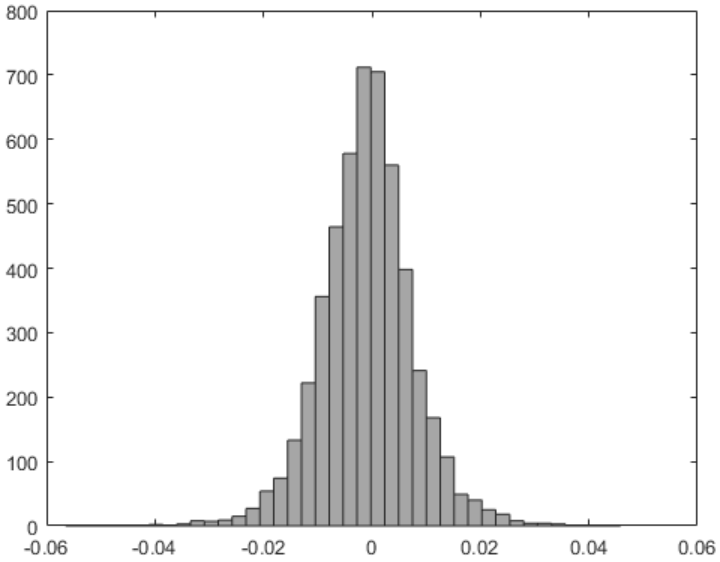


Figure 4.2: Histogram of the APV of GMIB hedging errors

beginning of year. A guaranteed minimum credited rate of 0.5% throughout the accumulation phase. We are interested in determining the cap to declare on December 15, 2017 that applies to the subsequent credited rate of December 13, 2018. We assume that the general account return and the pricing spread for the period are 5% and 1.5% respectively.

The allowance price on December 15, 2017 is €7.21. Due to the lack of OTC market data, we approximate the allowance call option quotes by the corresponding futures option prices, an excerpt of which is given in Table 4.4.

The option budget is 3.5%. Following (4.36) and the specifications in this example, we need to solve for κ so that the following holds:

$$0.2523 = C^m(0, 7.2461, 1) - C^m(0, (1+\kappa)(7.21), 1). \tag{4.44}$$

Linear interpolations are applied to estimate the prices of non-quoted options. Simple calculations leads to $\kappa = 10.38\%$. Therefore, a cap of 10.38% for the subsequent credited rate is declared. To construct the hedge, long $A_0/7.21$ calls with strike 7.2461 and short the same number of calls with strike 7.9584, where A_0 is the account value at December

K(€)	Price (€)	Maturity
6.0	2.06	12/12/2018
6.5	1.79	12/12/2018
7.0	1.55	12/12/2018
7.5	1.36	12/12/2018
8.0	1.19	12/12/2018
8.5	1.05	12/12/2018
9.0	0.93	12/12/2018

Table 4.4: Allowance call option prices on December 15, 2017

15, 2017. This is a bull spread costing exactly the same as the dollar value of the option budget. Finally, in reality, the spot allowance price on December 12, 2017 is €21.47, representing 198% return. Hence, the cap is an extremely effective risk mitigation in this example. **End of Example 4.3**

These examples conclude our discussion on the design and modeling of annuity products linked to the allowance market, established under the premise that allowances under open trading phases can be viewed as an alternative investment instrument. Such innovative products are appealing to policyholders seeking further diversification in market exposures, as well as thematic investors capitalizing on the global trend of emission control. In the subsequent chapter, we propose a framework for building sustainable portfolios under the climate change scheme, which also has a wide range of actuarial application for insurers and pension funds.

Chapter 5

Sustainable portfolio management under climate change

The contents of this chapter are adapted from the paper titled "Sustainable Portfolio Management under Climate Change" by Fang, Tan, and Wirjanto (2018)[43].

Climate change is posing a complex set of emerging risks to the global economy. While the impacts of climate change have attracted the attention of financial institutions, insurance companies, and pension funds, practical effort is mostly focused on liability drivers such as weather-induced mortalities and shifted frequency-severity patterns of extreme losses. On the other hand, relatively little interest has been directed toward the asset side of the balance sheet as demonstrated by Diaz-Rainey, Robertson, and Wilson (2017)[34], who surveyed 20725 articles published in 21 leading finance journals and found only 12 related to climate finance. Among this dearth of research, Kakeu (2017)[61] presents a theoretical model of portfolio decisions that incorporates environmental externalities into an investor's inter-temporal utility maximization goals. Contributing from an institutional perspective, this chapter proposes a framework for managing equity investment portfolios under the climate change scheme. The results presented in this chapter are highly relevant to financial and insurance products backed by managed funds, offering perspectives in the design of sustainable investment funds and the modeling of fund returns under the impacts of climate change.

Climate change affects an investment portfolio through two channels of transmission: directly, it elevates weather-related physical risk to real properties, agricultural production, resource landscapes, and infrastructure assets, which extends to increased market risk in

equity holdings with material business exposures in climate-sensitive regions. Indirectly, it triggers stricter environmental regulations and social norms, leading to a higher cost of emission and lower cost of green technologies for companies, which consequently induce downturns in carbon-intensive industries representing material constituents of a portfolio. This indirect impact has been the focus of most existing studies so far. Unlike many other risks, the impact of climate change risk on portfolios is particularly difficult to quantify due to its long-term nature requiring views and predictions spanning several decades into the future. The result in Oestreich and Tsiakas (2015)[79] is amongst the earliest evidence of market pricing of carbon risk. In this study, 65 German stocks are grouped into three portfolios based on the firms' allowance grants during Phase 1 and 2 of EU ETS. Using factor-based stock return models as filters, the authors discovered a significant carbon risk premium for the portfolio with the largest EUA allocation. Ibikunle and Steffen (2015)[58] examined the problem from a different angle. Under the CAPM and Carhart's Four Factor return model, the authors compared the risk-adjusted performances of green mutual funds to those of black and conventional mutual funds over 1991 to 2014, and found that the risk-adjusted return of green funds improved remarkably to catch up with that of the conventional funds and exceeded that of the black funds over the period of 2012 to 2014. Subsequent studies followed similar paths. For instance, Laplante and Watson (2017)[66], in exploring the underpricing of carbon risk by the market, performed empirical studies on the returns of environmentally conscious indices over the period from 2011 to 2016 and found that four of the five subject indices outperformed the S&P500 benchmark, thus confirming the potential of managing the portfolio profitably under climate change. Henriques and Sadorsky (2017)[53] studied the performance of three portfolios representing combinations of inclusions of the carbon-intensive and clean energy sectors, where the one that replaces the emission-heavy companies with the clean energy firms shows better performance under a series of metrics. Dorfleitner, Utz, and Wimmer (2018)[35] performed cross-sectional studies of equity abnormal returns, and concluded that companies with stronger Corporate Social Responsibility (CSR) rating outperformed the others. From the practitioners' side, Nordea Bank (2017)[83] studied the relative stock performances of firms with top and bottom Environmental, Social, and Governance (ESG) ratings in Europe over 2012 to 2015, whose result confirms the existence of value premiums paid by the market for high ESG ranking.

A common mitigation proposed in the few available studies centers around divesting portfolio holdings in the carbon-intensive sectors. Among those publications, a systematic quantitative approach is given by Anderson, Bolton, and Samama (2016)[3], who presented a multi-factor model for constructing carbon-efficient portfolios with minimal tracking errors to a selected benchmark through divestiture of the most carbon-intensive holdings. A

numerical demonstration was provided using the S&P500 index data to show the effectiveness of the algorithm. The key intuition under their study is the possibility of carbon-free portfolios that delivers a specific benchmark performance. On the other hand, Hunt and Weber (2018)[57] applied different combinations of divestment strategies to the S&P TSX index over 2011 to 2015, where the resultant portfolio show improved excess returns under the Four-Factor model. Despite its effectiveness as a risk mitigation, the simple divestment approach does not comprehensively leverage the complex set of threats and opportunities posed by climate change. For example, the possible downturn of the traditional energy and utility sectors are expected to be accompanied by the boom of the green energy sectors. Such an asymmetric impact of climate change applies both across and within industries, which is discussed in Busch and Hoffmann (2007)[18] through business input and output dimensions, as well as in Battiston et al. (2017)[40] under a financial stress-testing framework. It must be prudently measured and managed by portfolio managers to optimize the fund return in the long run.

This chapter contributes to existing literature in both methodological and applied aspects. In Section 5.2, we depart from existing empirical findings by directly analyzing the risk-adjusted performance of the three carbon-intensive industries (i.e. energy, utility, and material) under climate change. We show that on a stand-alone basis, these industries did not deliver a satisfactory performance compared to other sectors and benchmarks in order to be considered indispensable constituents. This finding enhances the premise of portfolio decarbonization and addresses the concerns of interest groups on the forgone portfolio returns in divesting the emission-heavy holdings for climate risk exposure reductions. In Section 5.3, we introduce carbon intensity as a measure of climate change risk exposure. Comparisons are made between revenue-based and profit-based intensities in their scope of applications and relevance to equity portfolio management. In Section 5.4, we propose a quantification method for stranded asset risk in individual stock selections and analyses, which to the best of our knowledge has not been specifically studied systematically. Finally, in Section 5.5, we present a scenario-based framework for the construction and management of sustainable portfolios, which encompasses views on both the threats and opportunities of climate change. Such a comprehensive approach is not yet leveraged in the existing literature. Extended applications of the framework, such as the generation of equity return scenarios are also discussed. A hypothetical example is provided by using a selected universe of assets for illustration.

5.1 Approach and data

Effective management of climate change risk of a portfolio must consider both the exposures and impacts of the risk. An exposure reduction decreases both the upside and downside potentials of climate change on the investments, making it a direct approach to avoid future uncertainties and hence the preferred method in existing studies. On the other hand, the impact of the risk represents the expected effect on portfolio returns under possible future climate change scenarios. It is a product of likelihood (of a scenario occurring) and severity (i.e. change in portfolio return under the scenario). Our approach to building a sustainable portfolio involves the quantification of both the exposure and the impacts. In particular, the projected impacts are used to derive parameter inputs in optimizing the portfolio, while constraints are set for the portfolio risk exposures to alleviate the extent of the uncertainty. This approach is intended to be comprehensive and leads to a systematic sustainable portfolio management framework.

We emphasize that the scarcity of directly usable data has hindered studies in investment-related climate change risk in general. In fact, even with the abundant market, economic, and environmental data, it is extremely difficult (if not impossible) to filter out the impacts of all non-climate factors explaining equity market returns, so that the price of the risk (in terms of return premiums) can be studied via existing methods. In this research, we make the best possible use of available data from various reliable sources, combined with reasonable subjective inputs. Data validation is done whenever necessary and applicable. A list of major data sources are as follows:

1. Market and sector index prices: the online database of S&P Dow Jones Indices, which provides price data for both the S&P500 (U.S.) and S&P TSX (Canadian) indices
2. U.S. interest rates: the online database of U.S. Treasury Department
3. Canadian interest rates: the Bank of Canada online database for Treasury Bill yields
4. Company financial data: Yahoo Finance and corporate filings
5. Company emission data: 2013 CDP Global 500 Climate Change Report [1]
6. Future climate scenarios: generated by the official WITCH model simulator

[1]This study uses the total annual emissions in the CDP filings, which is the sum of the Scope 1 and Scope 2 emissions reported by the firms. Definitions for emission scopes are given in the official site of the U.S. Environmental Protection Agency at: https://www.epa.gov/greeningepa/greenhouse-gases-epa

More details on the data sets used are given when the results are presented.

5.2 Carbon-intensive sector performance in North America

We begin the analysis with a cross-sectional study of the stock market returns for different industries. Concerns raised by interest groups against divesting carbon-intensive holdings in portfolios mainly involve foregone returns from the divestment, despite the various evidence of the outperformance of carbon-efficient funds. A direct analysis of the risk-adjusted returns of the carbon-intensive industries will address such a concern. The industry sectors are defined by the Global Industry Classification Standards (GIGS), where the carbon-intensive sectors include the energy, utilities, and material industries. Separate analysis are carried out for the U.S. and Canadian stock markets.

Sector performances for the U.S. market are measured by using the S&P 500 sector indices, which comprise of companies from the various market sectors in the S&P 500. The S&P 500 index is used as the benchmark. The daily closing index values are collected for a 10-year study period from July 31, 2007 to July 31, 2017, based on which the performance metrics, namely Sharpe Ratios (SR) and Treynor Ratios (TR), are calculated. This study period provides sufficient data with reasonable quality and reliability for an equity return analysis. The same metrics are calculated using the 5-year historical data for comparison. The 1-year U.S. Treasury Rate serves as a proxy for the risk free rate in the calculations. Results are summarized in Table 5.1.

It is apparent that the carbon-intensive sectors, particularly the energy and material sectors, underperformed the others and the benchmark across the various metrics. Similar conclusions can be drawn from the 5-year rolling Sharpe and Treynor Ratios as shown in Figures 5.1 and 5.2. The rolling metrics for all three carbon-intensive sector indices remain below the benchmark for most of the study period, with the exception of the utility industry, which exhibited slightly higher Treynor Ratios toward the end of the study period.

Sector	10 yr SR	5 yr SR	10-yr TR	5-yr TR
Energy	0.123 [10]	0.018 [10]	0.036 [10]	0.004 [10]
Utilities	0.313 [7]	0.682 [8]	0.107 [6]	0.214 [5]
Material	0.309 [8]	0.806 [7]	0.086 [8]	0.141 [8]
Consumer Discretionary	0.649 [4]	1.573 [1]	0.173 [4]	0.254 [2]
Consumer Staples	0.707 [2]	1.056 [6]	0.208 [1]	0.199 [6]
Financial	0.192 [9]	1.377 [4]	0.055 [9]	0.227 [4]
Health Care	0.712 [1]	1.473 [2]	0.208 [2]	0.260 [1]
IT	0.672 [3]	1.432 [3]	0.182 [3]	0.237 [3]
Real Estate	0.328 [6]	0.582 [9]	0.104 [7]	0.132 [9]
Benchmark	0.430 [5]	1.315 [5]	0.108 [5]	0.194 [7]

Note: numbers in the brackets represent the rankings for the sectors

Table 5.1: Performance metrics of U.S. sector indices

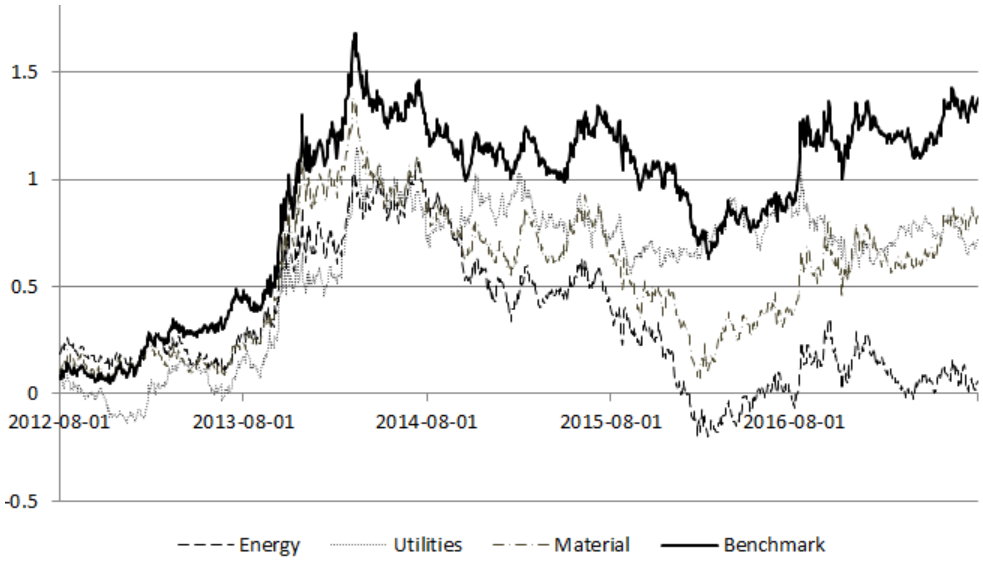


Figure 5.1: US market 5-year Rolling Sharpe Ratio

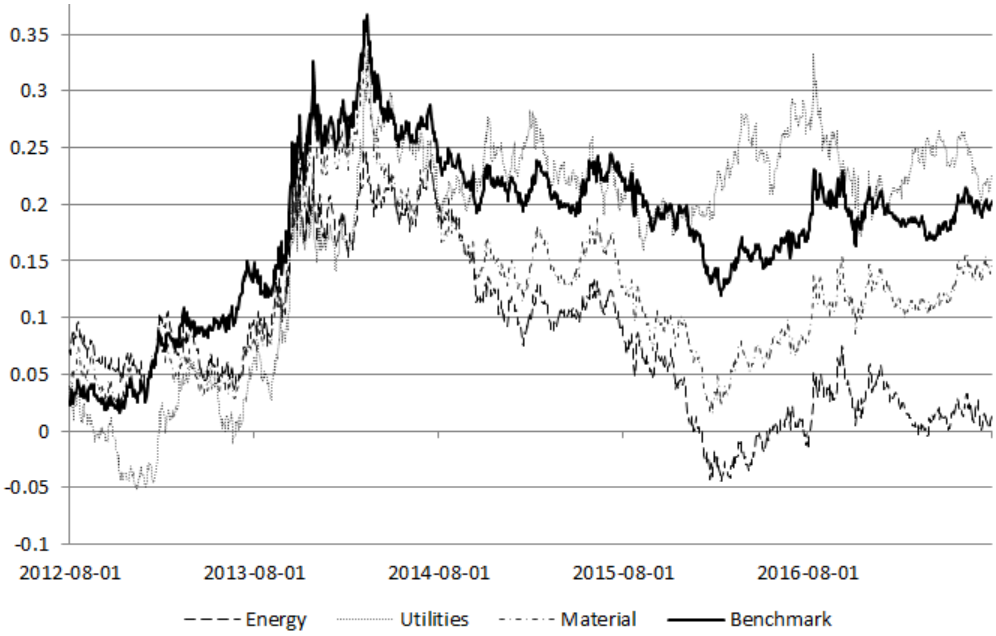


Figure 5.2: US market 5-year Rolling Treynor Ratio

Sector performances for the Canadian market are measured by using the S&P TSX sector indices, which comprise companies from the various market sectors in the S&P TSX Composite. The S&P TSX Composite index is used as the benchmark in this analysis. Daily closing index values are collected for the same study period as the one for the U.S. sectors, from which 10-year and 5-year Sharpe and Treynor Ratios are calculated. The 1-year Canadian Treasury rate serves as a proxy for the risk free rate in the calculations. Results are summarized in Table 5.2.

For the Canadian market, all three carbon-intensive sectors significantly underperformed the others and the benchmark on a risk-adjusted basis. Moreover, the energy and material sectors are associated with negative Sharpe and Treynor Ratios across different time horizons. Similar results are given by the 5-year rolling metrics in Figures 5.3 and 5.4, aside from the outperformance of the utility sector in Treynor Ratio.

To examine the statistical significance of these observations, paired-sample t-tests are conducted on the performance metrics, where results are summarized in Table 5.3. For both the U.S. and Canadian markets, the carbon intensive sectors deliver lower average metrics

Sector	10 yr SR	5 yr SR	10-yr TR	5-yr TR
Energy	-0.101 [12]	-0.189 [12]	-0.026 [12]	-0.033 [12]
Utilities	0.117 [10]	0.217 [9]	0.043 [9]	0.050 [10]
Material	0.077 [11]	-0.087 [11]	0.025 [11]	-0.022 [11]
Consumer Discretionary	0.387 [4]	1.728 [3]	0.120 [6]	0.338 [3]
Consumer Staples	0.922 [1]	1.841 [1]	0.390 [1]	0.532 [1]
Financial	0.262 [7]	1.137 [5]	0.074 [7]	0.192 [6]
Health Care	0.326 [6]	0.170 [10]	0.177 [2]	0.060 [9]
IT	0.410 [3]	1.755 [2]	0.160 [3]	0.506 [2]
Real Estate	0.222 [8]	0.508 [8]	0.072 [8]	0.117 [7]
Telecom	0.358 [5]	1.020 [6]	0.146 [4]	0.305 [4]
Industrial	0.460 [2]	1.263 [4]	0.126 [5]	0.229 [5]
Benchmark	0.122 [9]	0.593 [7]	0.028 [10]	0.081 [8]

Note: numbers in the brackets represent the rankings for the sectors

Table 5.2: Performance metrics of Canadian sector indices

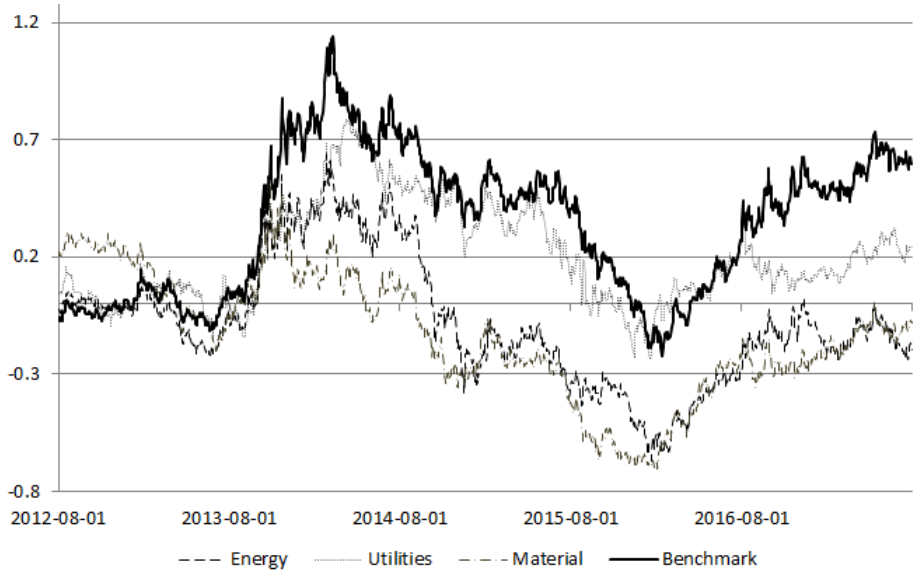


Figure 5.3: Canadian market 5-year Rolling Sharpe Ratio

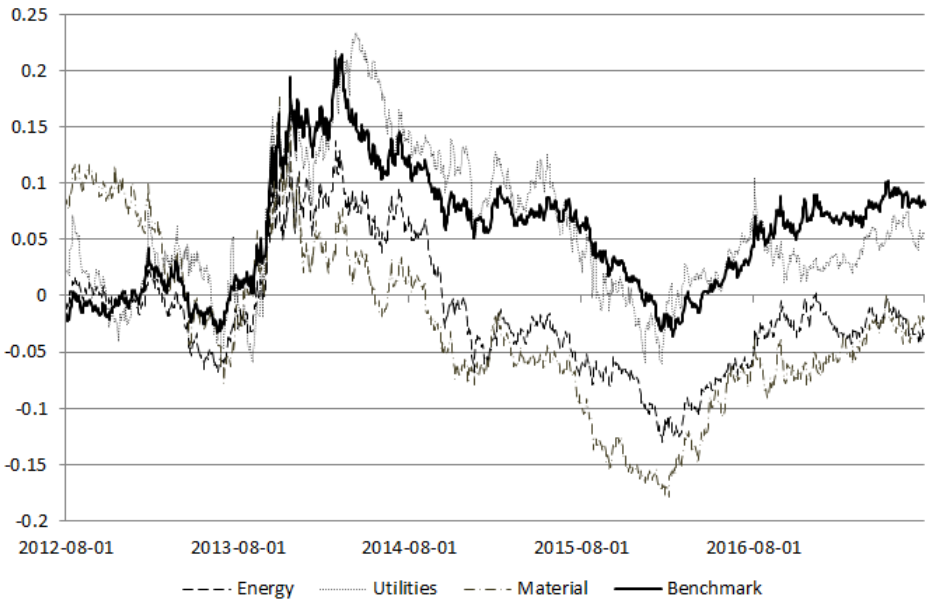


Figure 5.4: Canadian market 5-year Rolling Treynor Ratio

	U.S.		Canada	
	H_0: Mean sector metric $\geq$ Mean benchmark metric			
Sector/Metrics	Sharpe Ratio	Treynor Ratio	Sharpe Ratio	Treynor Ratio
Energy	0.000	0.000	0.000	0.000
Utilities	0.000	0.065	0.000	0.977
Material	0.000	0.000	0.000	0.000

Table 5.3: P-values of t-tests on average performance metrics

than the benchmarks at the 5% level, except for the utility sector Treynor Ratios. Such findings align with the observations above, whose statistical significance is hence confirmed.

In summary, consistent conclusions from the U.S. and Canadian markets can be drawn from results obtained so far. The carbon-intensive sectors underperformed the other industry sectors and benchmarks on a risk-adjusted basis in North America over the past decade. On a stand-alone basis (i.e. without considering the diversification benefit), they are not key performance enhancers in investment portfolios, implying that any forgone portfolio return from their divestment is likely to be minimal. Such findings are in line with those in Trinks et al. (2018) who demonstrated the divestment benefit at portfolio level using a

selected sample of fossil fuel and fossil-free portfolios. Properly reducing exposures to the carbon-intensive industries is hence considered one of the risk mitigation strategies in our framework for building sustainable portfolios under climate change.

5.3 Exposure measurement: carbon intensity

In addition to the overall position or weight of investment, a measurement of a stock's exposure to climate change risk is its carbon intensity, which is given by:

$$\nu_R = \frac{E[Q]}{E[R]}, \tag{5.1}$$

where Q and R represent the issuer's annual emission (in metric tons) and revenue amount respectively. This metric is an estimate of the company's contribution to climate change per unit level of revenue it generates, for which the economy charges a price. Details on deriving this market price of carbon is given in Daniel, Litterman, and Wagner (2017). Intuitively, companies with a higher carbon intensity are expected to suffer more on a per unit revenue basis from climate change risk factors such as stricter environmental regulations and competitions promoting green technologies. It should be pointed out that exposures to factors such as elevated climate-induced physical damages are not captured by this measure. Nevertheless, due to its simplicity and ease of interpretation, carbon-intensity remains a useful metric of climate change risk exposures based on which constraints on particular holdings can be set accordingly.

For the purpose of building sustainable equity portfolios, we propose using a profit-based carbon intensity:

$$\nu_P = \begin{cases} \dfrac{E[Q]}{E[P]} & E[P] > 0 \\ \text{undefined} & E[P] < 0, \end{cases} \tag{5.2}$$

where P in (5.2) is the issuer's annual profit. ν_P has the advantage of being relevant and easy to interpret as a measure of climate risk exposure for equity portfolios. A company's profit directly adds to its shareholder equity (including all earning distributions). As a result, ν_P represent a company's contribution to climate change per unit value it creates for shareholders. The disadvantage is that it is only defined when the expected profit is positive. We differentiate the two versions of carbon intensity in notation by subscripts,

Company	ν_R	ν_P	Sector
Exxon Mobil Corporation	0.347	4.481	Energy
UPS	0.226	2.868	Industrials
AT&T Inc.	0.069	0.480	Telecommunication
eBay Inc.	0.029	0.084	IT
Suncor Energy Inc.	0.526	5.329	Energy
General Motors Company	0.051	1.494	Consumer Discretionary
Intel Corporation	0.059	0.325	IT
Walt Disney Company	0.039	0.288	Consumer Discretionary
Goldman Sachs Group Inc.	0.007	0.036	Financial
PepsiCo, Inc.	0.087	0.858	Consumer Staples
Boeing Company	0.018	0.343	Industrials
Canadian National Railway Company	0.496	2.007	Industrials

Table 5.4: Historical carbon intensities of selected companies

as shown in (5.1) and (5.2).

In practice, the expected values in the formula are estimated using historical averages over the same period. This renders ν_P a wider scope of application since it is unlikely that the stock of a company consistently incurring losses enter the candidate asset universe for portfolio construction in the first place. A sufficiently large series of financial and emission data also permits the construction of asymptotic confidence intervals. Table 5.4 shows the carbon intensities of a few selected companies based on 2013 data[2], where values are expressed as one in a thousand. The energy and industrial firms exceed heavily in carbon intensities compared to other sectors based on the estimates. For example, a ν_P of 4.481 for Exxon Mobil indicates that on average the firm outputs 4.481 tonnes of emission to generate one-thousandth dollar of profit, compared to the 0.325 for Intel Corporation.

The portfolio carbon intensity is calculated as a weighted average of the constituents' carbon intensity. For a portfolio consisting of N stocks indexed by i with corresponding weights $w_i, i \in 1, 2, ...N$, its carbon intensities are given by:

$$\nu_R^p = \sum_{i=1}^{N} w_i \nu_{R,i} \tag{5.3}$$

[2]The revenues and profits in the intensity calculations are from the 2013 corporate financial statements to be consistent with the emission data used, which are from the 2013 CDP filing, the latest publicly accessible emission report.

$$\nu_P^p = \sum_{i=1}^{N} w_i \nu_{P,i}, \tag{5.4}$$

where the superscript p indicates the exposure metrics are at the portfolio level.

Carbon intensity may be used as a relative measure, where exposure reduction by excluding the stocks that rank the top in carbon intensity from the portfolio is reasonable. In addition, the carbon intensities can be used to set sector preferences in the strategic allocation process. An overall limit can be set on the portfolio level carbon intensity to restrict the overall exposure, which nonetheless requires subjective inputs.

5.4 Impacts of stranded asset risk

Stranded asset refers to a broad class of assets whose returns may fall below the existing market expectations due to regulatory, technological, and other socio-economic factors related to climate change risk. This is consistent with the conceptual framework for managing stranded assets given in Buhr (2017)[17], where climate change is a dominant source of risk. A more exhaustive list of factors causing stranded assets is introduced by Caldecott, Tilbury, and Carey (2014)[19], who also proposed and exemplified the type of factor scenarios that are most useful for a stranded asset management. In this section, we constrain the definition of the stranded asset to fossil fuel (coal, oil, and natural gas) reserves possessed by energy producers, where a portion of the reserve cannot be exploited at the current technology due to the emission reduction goals of established climate control schemes. In particular, a consensus is formed that under the generally accepted goal of restricting global warming to 2 degrees Celsius until 2100, about 80% of the world's reserves must remain unexploited. The detailed calculation for this estimate is given in Koehler and Bertocci (2016)[65], which also summarize key findings from Meinshausen et al. (2009)[71] that proposed the original hypothesis of stranded asset in *Nature*, and studies by Carbon Tracker. Though economists have generally reached a consensus on the rising trend of stranded asset risk, the impacts of this risk on individual firms are highly uneven and variable across the next four decades where the emission budget applies, making the quantification of the impact of stranded asset risk difficult.

In general, emission control is implemented through the following mechanisms:

1. Carbon tax: A fixed tax is levied on the energy products sold, which varies on the emission generated per unit of the product.

2. Emission trading: A predetermined amount of emission permits (allowances) are issued or auctioned to companies at the beginning of the year. A company submits sufficient permits at the end of the year to cover its emission or pay penalties on any excess emissions. The permits are traded on a secondary market between companies. The best example is the European Union Emission Trading Scheme (EU ETS).

In both mechanisms, the marginal cost of production is increased in order to reduce the equilibrium production quantity, hence reducing emissions. The energy producers may transfer the tax to consumers, which reduces demand and hence achieves the same effect of emission control.

For the purpose of this study, the impact of stranded asset risk is expressed as a change in expected equity returns caused by the fuel reserve getting stranded, holding all other factors constant. Consistent with the common valuation method for fuel reserves, a Net Asset Value (NAV) approach is adopted in forecasting the change in equity value. See Kibbey et al. (2017)[61] for more guidelines on this. Assuming that the current fair value of the company's assets, liabilities, and the fuel reserve can be estimated with a reasonable degree of accuracy, the impact of stranded asset risk, denoted by $\Delta\mu_S$ is given by

$$\begin{aligned}\Delta\mu_S &= \frac{A - \alpha A_R - D}{A - D} - 1 \\ &= \frac{-\alpha A_R}{A - D},\end{aligned} \tag{5.5}$$

where A, D, and A_R are the current fair value of the company's assets, liabilities, and the fuel reserve respectively. In (5.5), $\alpha \in [0, 1]$ is the stranding ratio representing the expected percentage of values lost in the fuel reserve for getting stranded. Estimation of α can be done deterministically or stochastically, although the latter is preferred due to the variability embedded in the risk drivers. Consider a fuel reserve with a maximum life of T years. Its fair value is given by:

$$A_R = \sum_{t=1}^{T} \frac{CF_t}{(1+k)^t}, \tag{5.6}$$

where k is the discount rate and CF_t represents the optimal year-t net cash flow generated by the reserve, accounting for all benefits and costs without the stranded asset risk.

Stranding occurs as the optimal production amounts and hence the associated cash flows are reduced due to regulations and other risk drivers, under which the stranding ratio is estimated by:

$$\alpha = \min\left\{\max\left\{\frac{1}{A_R}\left(A_R - \sum_{t=1}^{T}\frac{E\left[CF_t'\right]}{(1+k)^t}\right), 0\right\}, 1\right\}, \tag{5.7}$$

where CF_t' represents the optimal year-t net cash flow from the fuel reserve exploration under stranding factors and the remaining reserve amount as of year t. The key factors driving CF_t' include emission regulations, market price of the fuel produced, and technology affecting the emission per unit of production. Let $\Theta_t \in \{1, 2, ...M\}$ be a discrete variable denoting the (finite-state) scenarios of risk factors in year t. The interdependence between the variables must be captured by the specification of a joint distribution of CF_t' and Θ_t $\forall t \in \{1, 2, ...T\}$, expanding (5.7):

$$\alpha = \min\left\{\max\left\{\frac{1}{A_R}\left(A_R - \sum_{t=1}^{T}\frac{E\left[E\left[CF_t' \mid \Theta_t, CF_{t-1}'\right]\right]}{(1+k)^t}\right), 0\right\}, 1\right\}. \tag{5.8}$$

Clearly, (5.8) is best simplified numerically by means of simulations and scenario analysis. Specifications of the distribution functions should be based on both quantitative analysis (e.g. on the forward gas prices) and subjective inputs (e.g. views on trends of emission regulation policies).

An alternative simplified framework is available for estimating the stranding ratio α. This approach leverages the view that the magnitude of the stranding (i.e. 80% of the reserves globally) significantly outweighs the timing effect of the cash flows. Consider a fuel reserve with remaining level L and lifetime $[0, T]$. Denote by $X, X > 0$ the future maximum exploitable amount over the horizon. Note that X is a derived variable based on projected factors such as regulations and fuel prices. Stranding occurs iff $L > X$. The conditional distribution of X given each scenario can be specified based on inputs from subject area experts and calibrated to market data if applicable. The stranding ratio is given by

$$\alpha = \frac{E[L - X]^+}{L} = \frac{1}{L}\sum_{\theta=1}^{M} E\left[L - X \mid X < L, \Theta = \theta\right] P\left[X < L, \Theta = \theta\right]. \tag{5.9}$$

A hypothetical example to illustrate the application of (5.9) is given below.

Example 5.1: An oil producer has a reserve with a remaining level of 2 million liters

and a fair market value of 3 million. The company is governed by an Emission Trading Scheme, where the emission allowance allocation to company A is expected to reduce significantly over the life of the reserve for an accelerated emission reduction goal. The company identified two foreseeable scenarios in view of the regulations, technologies, and oil price projections, based on which distributions of the maximum permissible exploitations are derived where all parameters are in millions:

1. High oil price and low emission price: $X \equiv U(1.8, 2.3)$. Scenario probability is 0.2,
2. Low oil price and high emission price: $X \equiv U(1.6, 2.1)$. Scenario probability is 0.8,

where $U(.,.)$ denote the Uniform distribution. Furthermore, assume that the fair values of the company's assets and debts are 30 and 5 million respectively. From the given information, there is a probability of 0.6 that the reserve asset gets stranded, with a stranding ratio:

$$\alpha = \frac{E[\max\{2 - X\}]}{2} = 0.55. \tag{5.10}$$

The expected impact of stranded asset risk can be easily calculated by using (5.5):

$$\Delta\mu_S = -6.6\% \tag{5.11}$$

In practice, this long run return impact should be graded over the investment horizon to reflect the information discovery and response process of the market. **End of Example 5.1**

To conclude this section, we offer our perspectives on the impact of the climate change risks and market efficiency. Arguments against the climate change risk management are made based on the market being in the semi-strong form of efficiency, so that the risk is already priced. In our view, the pace at which risk premiums drift to their fundamental levels heavily depends on the nature of the risk. Market risks such as interest rate changes tend to be priced immediately while emerging risks such as climate change are priced relatively slowly. Evidence on this can be drawn from two streams of literature. The first stream includes studies confirming the pricing of the risk reflected by an outperformance of carbon-efficient funds over the traditional ones in the recent decade, such as in Ibikunle and Steffen (2015)[58]. In contrast, the second stream examines the market response to major risk events. Griffin et al. (2015)[46] studied the stock market's response to the first formal hypothesis of stranded asset risk published in the prestigious *Nature* magazine, and found limited stock return drops based on a sample of the 63 largest U.S oil and gas firms. The implication of these findings is that climate change risk is priced by the market only

gradually, and there is the potential to profit through the institution of proper management of investment portfolios.

5.5 Scenario generation and sustainable portfolio

In this section, we present a framework for the construction and management of sustainable portfolios under climate change risks. While carbon intensity offers a measure of a portfolio's exposures to climate change risk, the impact of the risk is comprehensive in nature and hence better estimated using scenario-based modular approaches. The flexibility of the framework allows for easy adaptations to different climate change models and mapping functions, for which the choices presented serve illustrative purposes only. The two outputs from the framework are:

1. Projected paths of equity return impacts by climate change scenarios, which has a range of applications in the modeling of equity portfolio returns for separate accounts and managed pension funds.

2. A mean-variance optimal portfolio leveraging the set of emerging risks and opportunities under climate change.

Climate change risk is difficult to quantify as a proper analysis of it would require the interdisciplinary expertise from environmental science, finance, and economics. As discussed in Mercer (2016)[73], solutions to this problem rely on the use of Integrated Assessment Models (IAM), which is a set of scientific models integrating methodologies across multiple disciplines to predict future climate scenarios as well as their potential impacts on the economy. Examples of IAM given in Mercer (2016) include the Climate Framework for Uncertainty, Negotiation and Distribution (FUND), the Dynamic Integrated Climate-Economy Model (DICE), and the World Induced Technical Change Hybrid Model (WITCH). The WITCH model is used in this framework to generate climate scenarios due to its highly interpretable scenario outputs that can be easily transformed to factor values. Details of the WITCH model are given in Bosetti et al. (2007)[12] and Bosetti et. al (2009)[11]. The derivation of the framework is presented below. Note that certain numerical tools require subjective actuarial inputs, which should be validated or improved based on actual market experience.

5.5.1 Equity return scenarios under Climate Change

Climate change risk is driven by three factors identified below:

1. Technology (T): the rate of progress and investment in the development of technology to support the low-carbon economy. It can be interpreted as a measure of the future low-carbon investment flows. The energy sectors are most influenced by this factor as the low-emission energies start to excel in both efficiency and cost.

2. Climate Impact (I): the impacts of climate change on the physical environment that manifest through two channels. The first is shifts in the frequency and severity of large weather-induced (e.g. wildfire, coastal flooding) losses, which is systematic and affects industries including the real estate and insurance sectors. The second is resource scarcity under chronic weather patterns, where economic productions must adapt based on their reliance on resources that are at risk of becoming scarcer or, in rarer cases, more abundant. Sectors such as agriculture and materials are most influenced by this factor.

3. Policy (P): the developments in climate control policies to reduce carbon emissions. In addition to the regulatory facets mentioned in previous sections, it also includes the governments' coordinated effort in items such as building codes to improve energy efficiency and land use regulations to restrict deforestry.

The factors are assumed to follow different paths under future climate change scenarios. Four scenarios are defined under the WITCH model as described in Table 5.5, under which values of selected environmental and economic variables across time are generated. Based on a theoretical representation, proxies for the risk factors are selected from these variables as summarized in Table 5.6.

Given a maximum scenario horizon $[0, T]$ and projection horizon $[0, \tau], \tau \leq T$, it is convenient to express the factors as bounded unitless variables using a continuous function $G : \mathcal{R} \to [0, 100]$. A uniform transformation is adopted in this study so that for any $s \in \Omega$ and $t \in [0, \tau]$:

$$f_{j,s,t} = 100 \left(\frac{y_{j,s,t} - \min \{y_{j,p,k} : \forall p \in \Omega, k \in [0, T]\}}{\max \{y_{j,p,k} : \forall p \in \Omega, k \in [0, T]\} - \min \{y_{j,p,k} : \forall p \in \Omega, k \in [0, T]\}} \right), \quad (5.12)$$

where $f_{j,s,t}$ and $y_{j,s,t}$ represent respectively the values for factor j and its proxy under scenario s at a projection point $t \in [0, \tau]$, where the post-transformation average is taken

Scenario	Description
Fragmentation	A scenario with limited and fragmented action on climate change. Global temperature increases up to 4℃ above pre-industrial era.
Coordination (500 ppm)	A scenario with weak climate control effort until 2020, and moderate collaborative global climate control effort afterwards, aiming at stabilizing the concentration of GHGs at 500 parts per million (ppm) of CO_2 eq by 2100.
Transformation (450 ppm, without permit trading)	A scenario with weak climate control effort until 2020, and moderate collaborative global climate control effort afterwards, aiming at stabilizing the concentration of GHGs at 450 parts per million (ppm) of CO_2 eq by 2100.
Transformation (450 ppm, with permit trading)	A scenario with weak climate control effort until 2020, and moderate collaborative global climate control effort afterwards, aiming at stabilizing the concentration of GHGs at 450 parts per million (ppm) of CO_2 eq by 2100. Emission trading is adopted after 2020.

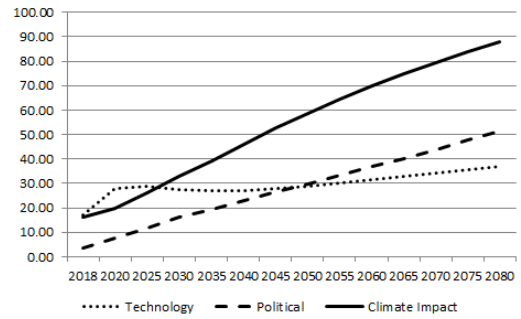
Table 5.5: Climate change scenarios under the WITCH model

for factors with multiple proxies. Here, Ω represents the scenario set (given in Table 5.5). Figure 5.5 shows the generated factor value paths under the transformation function in (5.12).

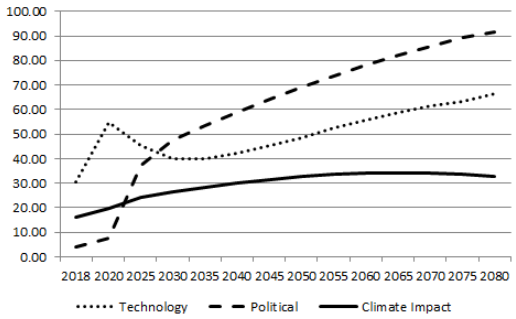
Risk factor	Proxy 1	Proxy 2
Technology (T)	Investment in advanced biofuel (USD)	Investment in energy efficiency (USD)
Political (P)	Green House GAS (GHG) abatement (ton CO_2/year)	None
Climate Impact (I)	Radiative Forcing (W/m^2)	Global mean temperature change (°C)

Table 5.6: WITCH model proxies for climate change risk factors

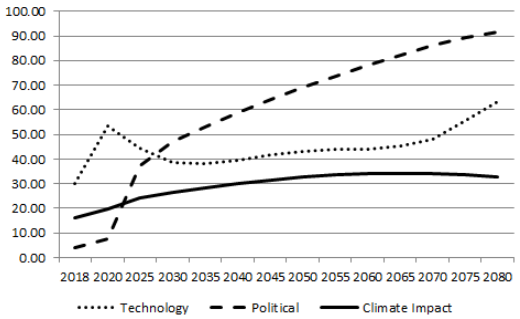
The factors impact the sectors in fundamentally different directions and magnitudes. As examples, development in green technology benefit the green energy sectors but detriment the traditional energy producers. Extreme whether events cause facility damages and



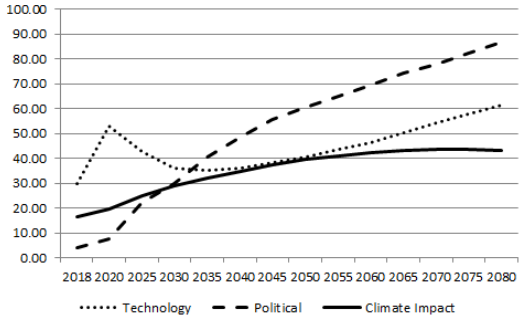
(a) Fragmentation



(b) 450 ppm, no permit trading



(c) 450 ppm, with permit trading



(d) 500 ppm, no permit trading

Figure 5.5: Factor paths under climate change scenarios

production disruptions, directly harming the industrial, real estate, and consumer staples sectors. To account for such differences, for each sector, sensitivities to the four risk factors are assigned, where a positive sensitivity implies that stocks in the sector should benefit from the risk factor, and vice versa. This can be done for the individual stocks if needed. The sector sensitivities in this study are given in Table 5.7 for an illustrative purpose. For simplification, the effect of cross-holdings[3] are not considered here. Methods can be developed to appropriately capture it at the individual stock level if it is deemed to be necessary.

Denote by $\beta_{i,j} \in [-1, 1]$ the sensitivity of sector i to factor j. The aggregate level of factors on (stocks in) sector i at time point $t \in [0, \tau]$ under scenario $s \in \Omega$, denoted by

[3]Such situations include, for example, a bank holding securities issued by institutions that invest in the carbon-intensive sectors, which indirectly impacts the bank's stock return. See Battiston et al. (2017)[40].

Industry Sector	T	I	P
Green Energy	0.75	0	0.5
Consumer Discretionary	0	0	-0.25
Consumer Staples	0	-0.25	0
Energy	-0.5	-0.5	-1
Financials	0	-0.25	0
Health Care	0.25	-0.25	0
Industrial	0	-0.75	-0.5
Information Technology	0.25	0	0
Materials	0	-0.25	-0.75
Real Estate	0	-1	0
Telecommunications	0	-0.25	0
Utilities	-0.25	-0.5	-0.75

Table 5.7: Factor sensitivities to stock market sectors

$\eta_{i,s,t}$ is given by:

$$\eta_{i,s,t} = \sum_{j=1}^{4} \beta_{i,j} f_{j,s,t}. \tag{5.13}$$

Changes in the aggregate factor levels are recognized by the market and reflected as impacts to stock returns, which requires specification of a mapping $F : \mathcal{R} \to (-1, \infty)$ that satisfies the following regularity conditions:

1. $F(x)$ is piecewise continuous.
2. $F(x) \geq 0$, $\forall x \geq 0$ and $F(x) < 0$, $\forall x < 0$.
3. $F(x)$ is non-decreasing in x.
4. $F''(x) \geq 0$, $\forall x \leq 0$ and $F''(x) < 0$, $\forall x > 0$.

Conditions 1 to 3 ensure the proper conversion of the aggregate factor level changes to stock return impacts, so that positive changes are mapped to increase in returns and vice versa. This aligns with how the factor sensitivities are defined. Condition 4 is based on the hypothesis that the market's psychological response is a concave function of the absolute magnitude of the aggregate factor change: it takes more effort and resources to translate large risk factor shifts to proportionate actions. This is consistent with the common

value function specifications under prospect theory in the seminal work of Kahneman and Tversky (1979)[60]. In practice, it is common to formulate the function so that its range is a subinterval of the one specified above to accommodate reasonable expectations. We propose a function in the form of:

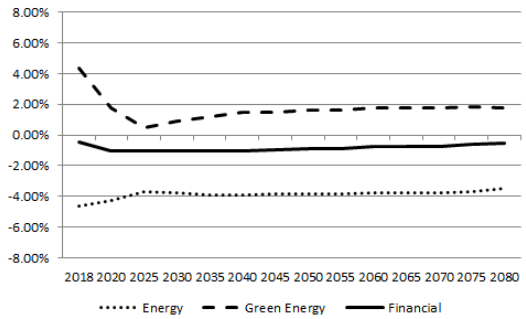
$$F(x) = \gamma \arctan(bx) + J, \tag{5.14}$$

where $\gamma > 0$, $b > 0$ and J are scalars. In this study we set $\gamma = 0.2\pi$, $b = 0.1$ and $J = 0$. Without loss of generality, let $\mathcal{T} = \{0, t_1, t_2, ..t_k, ...t_n\}$ be a partition of $[0, T]$ consistent with that of the scenario horizon. The impacts of climate change risk on the return of (stocks in) sector i for the period $[t_k, t_{k+1}] \in \mathcal{T}$ under scenario $s \in \Omega$, which we denote by $\Delta r_{i,s,t_k}$ follows:

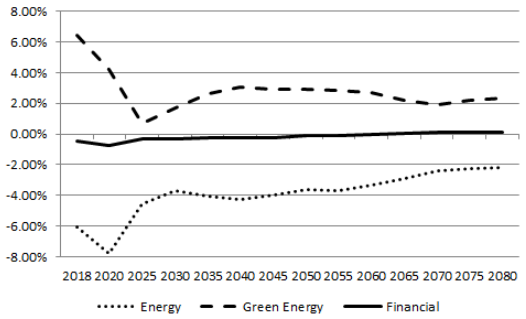
$$\Delta r_{i,s,t_k} = \frac{1}{5\pi} \arctan\left(\frac{\eta_{i,s,t_{k+1}} - \eta_{i,s,t_k}}{10}\right). \tag{5.15}$$

The return impact scenarios for stocks in the energy, green energy, and financial sectors are shown in Figure 5.6. As expected, the energy and green energy stocks are subject to significant return decrease and increase respectively across scenarios, while the financial industry is relatively resistant to climate change. In practice, it may be desirable to smooth the return impacts to reflect the market's stale reaction to climate change risk as discussed in Section 5.4, which can be achieved through a linear grading of the impact over each partition interval. This is a straightforward implementation that is not illustrated here for brevity.

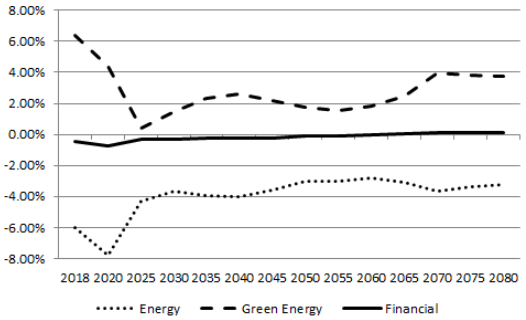
We conclude this subsection with a discussion on the financial applications and limitations of the return impact scenarios obtained. The scenarios can be adapted to any real world scenario-based calculation involving asset return models, where common examples include budgeting, cash flow testing, and asset-liability management. For prescribed scenarios, the stock return impacts are added to the deterministic return paths to create corresponding stressed scenarios capturing the climate change risk. For stochastic scenarios, the return impacts are used to adjust the corresponding drift parameters in the stock return process assumed in the economic scenario generators. These implementations are simple when the fund returns are modeled directly based on the exact portfolio compositions. In practice, many banks and insurers model the fund returns indirectly as a linear function of the returns from a selection of benchmark indices (e.g. S&P500). In this case, under each climate change scenario, it is necessary to calculate for each benchmark index a path of weighted average return impacts based on its compositions in the sectors over the projection period. The resultant paths of impacts are then applied as described above to the benchmarks' return scenarios.



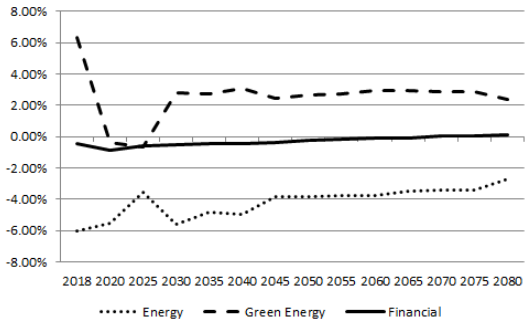
(a) Fragmentation



(b) 450 ppm, no permit trading



(c) 450 ppm, with permit trading



(d) 500 ppm, no permit trading

Figure 5.6: Equity return impact scenarios for selected sectors

Three characteristics of the equity return impact scenarios should be emphasized. First, stocks in the same sector share the same risk factor sensitivities, and hence by construction are subject to the same equity return impacts from climate change. This assumption substantially reduces the effort but can be relaxed if is required to do so. Second, the effect of climate change on the covariances between equity returns are indirectly captured by the systematic risk factors. Future work to directly model this effect is desirable. Lastly, the return impacts are derived from the outputs of the WITCH model under its deterministic climate change scenarios, whose accuracy and adequacy are assumed in this study.

5.5.2 Constructing and managing sustainable portfolios

Carbon intensities, stranded asset risk, as well as equity return impact scenarios discussed in the previous sections are used in constructing sustainable portfolios under climate change. Markowitz mean-variance portfolio optimization framework is assumed to apply: given a universe of n stocks, the goal is to minimize the portfolio variance given a target portfolio return, subject to applicable constraints. The problem is formulated as follows:

$$\begin{aligned} \underset{W}{\text{minimize}} \quad & W^T \Sigma W \\ \text{subject to} \quad & W^T \mu = \mu^* \\ & W^T 1_n = 1 \\ & AW \leq a, \end{aligned}$$

where the last inequality specifies the appropriate constraints and:

- W is the n by 1 portfolio weight matrix. Without short positions (as for insurers and pension funds), entries of W are in $[0, 1]$, with column sums of 1.
- Σ is the n by n covariance matrix of the stock returns, estimated from historical data.
- μ is the n by 1 mean return vector of the stocks, estimated from historical data.
- μ^* is the target portfolio return.

A key assumption under the Markowitz framework is that the stock returns can be modeled by a Multivariate Normal (MVN) distribution with mean parameters μ and Σ. This assumption is also made in this section.

The conventional top-down approach to portfolio management follows four steps. The first is the asset universe selection, where a set of stocks is specified for use in building the portfolio. This is a topside decision that captures applicable regulatory constraints and internal investment policies. In the second step, Σ and μ are estimated from price data. Choice of the estimation method is beyond the scope of this chapter, while the market-implied equilibrium return in Black and Litterman (1991)[7] is used as the estimator of μ in the example we present. The third step is the strategic asset allocation, where decisions are made on the weights of the major stock market sectors. These decisions enter into the optimization problem as constraints. For insurers and pension funds, most portfolios are

backing long-term liabilities for which balanced allocations are appropriate. The final step is optimization using the results from the previous steps to obtain the optimal portfolio weights.

To induce portfolio sustainability under climate change, three additional modules apply before the final optimization step:

1. Risk exposure management. The carbon intensity metrics introduced in Section 5.3 are used to quantify the climate change risk exposures of the stocks in the universe, based on which mitigations are applied, which may include:

   - Tactical asset allocation: adjustments are made to the strategic asset allocations based on sector average exposures.
   - Divestment: stocks with the top exposure are excluded from the asset universe.
   - Portfolio exposure control: a cap is set on the overall portfolio carbon intensity, formulated as an additional constraint equation. With green energy stocks that carry negative carbon intensities, there is a potential to make the portfolio carbon neutral (i.e having 0 carbon intensity).

2. Stranded asset management. Based on available financial data and regulations, identify stocks with a material stranded asset risk (SAR). Quantify the risk as discussed in Section 5.4, which supplements the risk impact management module below.

3. Risk impact management. This final module utilizes the equity return impact scenarios presented in Section 5.5.1. Scenario probabilities are assigned, which we denote by λ_s, $s \in \Omega$. A projection horizon $[t_k, t_{k+1}]$ is fixed according to the portfolio rebalancing schedule. By making the intuitive assumption that the stock returns absent of climate change risks are independent of the impacts of the risks, the posterior return distribution is fully specified (i.e. MVN) with a mean vector:

$$\mu^{post} = \mu + \Delta\mu \tag{5.16}$$

   and a posterior covariance matrix:

$$\Sigma^{post} = \Sigma + \Pi \tag{5.17}$$

   where $\Delta\mu$ denotes the n by 1 view vector of the return impacts whose ith entry is

$$\Delta\mu_i = \sum_{s\in\Omega} \lambda_s \Delta r_{i,s,t_k} \tag{5.18}$$

and Π denotes the n by n covariance matrix of the stock return impacts whose ijth entry is:

$$\Pi_{ij} = \sum_{s\in\Omega} \lambda_s \left(\Delta r_{i,s,t_k} - \sum_{s\in\Omega} \lambda_s \Delta r_{i,s,t_k} \right) \left(\Delta r_{j,s,t_k} - \sum_{s\in\Omega} \lambda_s \Delta r_{j,s,t_k} \right). \tag{5.19}$$

i and j here index the stocks as opposed to the sectors. Posterior estimates are updated periodically along with their priors upon portfolio rebalancing. Further tactical asset allocation can also be made as needed based on the results.

These modules transfer the impacts from climate change to the mean-variance optimization algorithm while controlling the portfolio level risk exposure. Overall, this framework leads to a sustainable portfolio and a best practice for its management. A hypothetical example is provided below to illustrate the framework in application.

Example 5.2: Consider a representative balanced managed fund that aims at achieving an 8% annual growth under climate change. To make the scope of the problem manageable, we take a small subset of 33 stocks covering all sectors from the S&P500 as the market portfolio shown in Table 5.8, where the capitalization weights are proportional to the actuals in the S&P500 index. The portfolio is rebalanced quarterly with no short positions.

The prior mean return vector take values from the implied equilibrium returns derived following Black and Litterman (1991) as follows:

$$\mu = 1_n r_f + \delta \Sigma W_{eq}, \tag{5.20}$$

where δ is the market price of risk, which is set to 3.73 based on the 2017 estimates in Miller and Lu. W_{eq} is the 33 by 1 vector of the market capitalization weights. r_f is the risk free rate set to 2% based on the average 1-year U.S. treasury yield in 2017. Σ is estimated from the daily adjusted closing prices of the stocks from 2012 to 2017. This approach overcomes the issue of the resultant optimal portfolio concentrating in a few sectors.

To minimize the confounding influence of subjective inputs, no strategic asset allocations are assumed in this example. We further assume that no stranded assets are identified. With the target return $\mu^* = 8\%$, the optimal portfolio weights are as shown in the second column of Table 5.9, which is associated with a portfolio volatility of 13%. The three modules of managing climate change risk are then implemented as follows:

1. Risk exposure management: ν_R and ν_P are estimated for the stocks by using the 2013 emission and income statement data. Results are shown in Table 5.8. The

Company	Sector	Cap	μ_R
NextEra	Green Energy	1.30%	5.11%
Walt Disney Company	Consumer Discretionary	2.53%	7.28%
General Electric Company	Consumer Discretionary	1.90%	7.59%
Time Warner Inc.	Consumer Discretionary	1.27%	6.93%
Starbucks Corporation	Consumer Discretionary	1.10%	7.35%
PepsiCo Inc.	Consumer Staples	2.53%	5.56%
Coca Cola	Consumer Staples	3.06%	5.64%
Wal-Mart Stores Inc.	Consumer Staples	4.08%	5.28%
General Mills	Consumer Staples	0.42%	5.48%
Proctor and Gamble	Consumer Staples	3.21%	5.67%
Exxon Mobil Corporation	Energy	5.70%	7.38%
Occidental Petroleum	Energy	1.05%	8.27%
Chevron	Energy	3.90%	8.10%
ConocoPhillips	Energy	1.33%	8.93%
Goldman Sachs Group Inc.	Financial	1.36%	9.54%
Morgan Stanley	Financial	1.36%	11.10%
JPMorgan Chase	Financial	5.77%	9.35%
Citigroup Inc.	Financial	2.76%	10.74%
Colgate Palmolive	Health care	0.92%	6.17%
Johnson&Johnson	Health care	5.39%	5.89%
UPS	Industrials	1.49%	6.42%
Boeing Company	Industrials	3.18%	7.68%
Oracle Corporation	IT	2.91%	8.03%
Intel Corporation	IT	3.77%	8.27%
Google Inc.	IT	12.60%	8.78%
Microsoft	IT	12.46%	9.50%
Dow Dupont	Materials	2.48%	8.87%
Newmont Mining	Materials	0.33%	4.89%
Simon Property Group	Real Estate	1.00%	5.98%
AT&T Inc.	Telecommunications	3.92%	5.85%
Verizon	Telecommunications	3.43%	5.93%
American Electric Power	Utilities	0.57%	5.07%
Duke Energy Corporation	Utilities	0.92%	4.54%

Table 5.8: Market portfolio

metrics for the green energy sector are assigned conservatively negative values. Risk exposures of the original portfolio are $\nu_R^p = 0.1525$ and $\nu_P^p = 1.4942$. As an exposure control, the stock with the top intensity is excluded from the portfolio. In addition, the following caps are imposed:

$$\nu_R^p < 0.07, \quad \nu_P^p < 0.7.$$

These constraints represent over 50% reduction in the climate change risk exposures.

2. Stranded asset management: This is not needed per the assumption made in this example.

3. Risk impact management: The equity return scenarios in Section 5.5.1 are input in this step. A 0.25 probability is assigned to all four scenarios. A linear interpolation is used to extract the return impacts for the first quarter of 2018. Posterior mean vector and covariance matrix are calculated following (5.16) to (5.19).

The posterior optimal portfolio weights under climate change risk and the suggested mitigations are given in the third column of Table 5.9. These allocations sum to one trivially, but vary from those of the initial portfolio. As expected, investment in green energy is substantially increased, while the weights in the energy and utility sectors are heavily reduced. The portfolio has a volatility of 0.1267 and carbon intensities of $\nu_R = 0.0494$ and $\nu_P = 0.4267$, all below the metrics for the initial portfolio.

Figure 7 shows the efficient frontiers before and after the climate change risk management, respectively labeled as "Initial" and "Posterior". A small gap is visible between the two. The posterior frontier lies above the initial for most of the feasible set until the two curves cross at the high volatility region, which is a desirable result for balanced funds that target long term growth with moderate volatility. **End of Example 5.2**

Company	W	W^{post}	ν_R	ν_P
NextEra	1.06%	11.66%	-0.0010	-0.0090
Walt Disney Company	2.76%	0.83%	0.0392	0.2879
General Electric Company	2.05%	1.94%	0.0430	0.3733
Time Warner Inc.	0.91%	0.59%	0.0082	0.0590
Starbucks Corporation	0.86%	0.00%	0.0628	0.4993
PepsiCo Inc.	1.28%	1.53%	0.0871	0.8581
Coca Cola	2.25%	3.37%	0.0788	0.4299
Wal-Mart Stores Inc.	3.30%	4.15%	0.0445	1.3223
General Mills	0.09%	0.72%	0.0556	0.5462
Proctor and Gamble	2.14%	4.06%	0.0783	0.5005
Exxon Mobil Corporation	5.77%	0.00%	0.3469	4.4813
Occidental Petroleum	1.00%	0.00%	0.8130	3.3621
Chevron	5.03%	0.00%	0.2833	2.9132
ConocoPhillips	1.48%	1.95%	0.4697	2.8188
Goldman Sachs Group Inc.	0.76%	1.11%	0.0071	0.0362
Morgan Stanley	2.54%	3.49%	0.0093	0.1175
JPMorgan Chase	5.84%	5.35%	0.0142	0.0599
Citigroup Inc.	3.85%	3.92%	0.0143	0.0751
Colgate Palmolive	2.37%	2.48%	0.0383	0.2979
Johnson&Johnson	3.90%	8.01%	0.0166	0.0856
UPS	0.03%	0.00%	0.2262	2.8680
Boeing Company	3.52%	0.00%	0.0182	0.3433
Oracle Corporation	3.08%	5.66%	0.0123	0.0418
Intel Corporation	4.24%	6.74%	0.0593	0.3249
Google Inc.	13.85%	14.18%	0.0214	0.0932
Microsoft	14.40%	14.61%	0.0161	0.0574
Dow Dupont	3.04%	2.39%	0.6278	7.4853
Newmont Mining	0.00%	0.00%	0.7102	1.0279
Simon Property Group	0.87%	0.00%	0.1140	0.3924
ATT Inc.	3.42%	0.00%	0.0687	0.4801
Verizon	3.60%	1.26%	0.0485	0.5085
American Electric Power	0.71%	0.00%	8.2196	82.3556
Duke Energy Corporation	0.00%	0.00%	5.4241	46.3152

Table 5.9: Exposure metrics and optimal portfolio weights

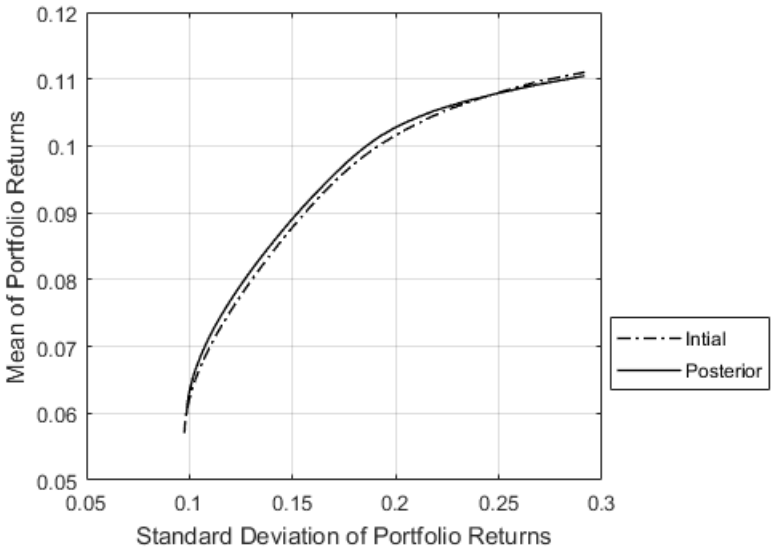


Figure 5.7: Efficient frontiers: initial vs. posterior

Chapter 6

Conclusion

In this thesis, we have presented valuation frameworks and models of emission allowances as well as related derivatives under the prevailing emission trading schemes, which are broadly categorized into being operating in a closed or open trading phase. A closed trading phase, as best exemplified by Phase 1 of EU ETS, mandates the lapse of unused allowances at the end of the phase. Three structural models representing different assumptions on the market aggregate emission process were proposed, which included the ABM emission rate, Vasicek emission, and the Vasicek emission rate models. Closed-form expressions were derived for allowance, allowance futures, and allowance option prices under these models, which have the advantage of being analytically-tractable and easy to implement compared to most of the prior studies in the literature. Despite the implications of market information incompleteness, the models remain useful in deriving the emission values from market prices via model calibration, the results of which concluded the superior performance of the ABM emission rate model among the three in both reliability and accuracy. The framework was extended to accommodate allowance valuation for individual firms, whose implementation was illustrated using real data of a representative U.S. power plant.

In contrast, an open trading phase, as best exemplified by Phase 2 and 3 of EU ETS, permits inter-phase banking of any unused allowances. We performed empirical analyses of historical allowance prices under the open phase, which reveal similar stylized facts as the stock market that includes excess return kurtosis, volatility clustering, and volatility smiles. Three reduced-form models were proposed, namely the Lognormal allowance price model (the LN model), the Skewness-kurtosis-modified Lognormal allowance price model (the SKM Model), and the two-component Mixture Lognormal allowance price model (the MLN-2 model). Closed-form option price expressions were derived under these models,

which share the advantage of being analytically-tractable. Numerical implementation was performed through calibrating the models to actual allowance futures call option prices. While all three models tend to misprice the out-of-the-money options, both the SKM and MLN-2 models outperformed the LN model. In addition, the MLN-2 model tends to overprice the out-of-the-money options, where the opposite applies to the SKM model. These findings are important factors to consider for practitioners in model selections for different applications.

In addition, we presented the design and modeling considerations of variable annuities (VA) backed by allowance-based funds, as well as fixed indexed annuities (FIA) whose credited rate are based on the allowance return under open trading phases. Both the non-lapsing feature and low return correlations with the stock market make emission allowance a suitable alternative asset class for portfolio diversification. Hence incorporating it into the annuity product designs represents an innovative step forward for insurers and pension plans. For the allowance-based VA products, we presented the pricing, valuation, and hedging of common guarantee riders. Under simplification assumptions, we obtained closed-form results for GMMB, GMDB, and GMIB riders utilizing the allowance put option valuation expressions under the open-phase models introduced. A volatility control mechanism for active fund management is also introduced, which shall improve hedging effectiveness by restricting the variations in the fund volatility. For the allowance-based FIA products, we presented strategies on product design to mitigate the market risk exposures of insurers as well as the associated hedging methodologies. We illustrated these modeling considerations and logic through examples with numerical implementations. These results can also be easily incorporated into the existing actuarial platforms adopted by the insurers for VA and FIA models, since the only adaptation required is transferring to the allowance market, whose scenarios can be efficiently generated following the algorithm we proposed for both outer and inner loop runs.

Finally, we introduced a complete framework for the management of sustainable equity investment portfolios under the climate change scheme, where risk exposures and stranded asset risk are quantified through the proposed metrics. We empirically demonstrated the inferior risk-adjusted performance of carbon-intensive sectors in the past decade to justify the mitigation of portfolio carbon risk via divestment from these industries. Furthermore, we presented in detail a method to generate equity return impact scenarios that can be incorporated into scenario-based calculations and models in various actuarial practices involving managed funds. The scenarios are also core inputs to formulate the posterior equity return distributions in the mean-variance optimization framework, from which optimal sustainable portfolios are built.

Several paths of future work are desirable for the subject areas covered in this thesis. For valuation of allowances and related derivatives, the rapid growth of emission trading schemes in both developed and emerging markets should hopefully give rise to reliable data to further validate the framework and models presented in the thesis. In particular, technology advances have made emission reporting possible on higher-frequency basis, which (upon actual implementation) shall inspire a review of the structural models we presented for closed trading phases via a direct calibration to actual emission processes. On the other hand, numerical studies of the stability of parameter estimates are highly desired for the option valuation models under open trading phases. For the allowance-based VA and FIA products we presented, it is of interest to analyze the profitability of the designs and validate the modeling approaches proposed as policy experiences become available. Risk metrics should also be examined through numerical studies and historical market scenarios to provide insights to regulators for these innovative products. Lastly, many models and functions used in our framework of sustainable equity portfolio management are based on certain theories and subjective inputs. Albeit the full flexibility of the framework in accommodating alternative specifications for these components, their calibration and validation constitute an interesting area of future research as actual experience data becomes available.

References

[1] European Environmental Agency. EU emissions trading system (ETS) data viewer, June 2016.

[2] E. Alberola, J. Chevallier, and B. Cheze. Price drivers and structural breaks in european carbon prices 2005-2007. *Energy Policy*, 36(2), 2008.

[3] M. Andersson, P. Bolton, and F. Samama. Hedging climate change risk. *Financial Analysts Journal*, 27(3), 2016.

[4] European Central Bank. Key ECB interest rates, 2016.

[5] N. Bellamy and M. Jeanblanc. Incompleteness of markets driven by a mixed diffusion. *Finance and Stochastics*, 4(2):209–222, 2000.

[6] T. Bjork. *Arbitrage Theory in Continuous Time*. Oxford University Press, 2 edition, 2004.

[7] F. Black and R. B. Litterman. Asset allocation, combining investor views with market equilibrium. *Journal of Fixed Income*, 1(2), 1991.

[8] F. Black and R. B. Litterman. Global portfolio optimization. *Financial Analyst Journal*, 48(5), 1992.

[9] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[10] K. Borovkov, G. Decrouez, and J.Hinz. Jump-diffusion modeling in emission markets. *Stochastic Models*, 27, 2010.

[11] T. Bosetti, M. Tavoni, and E. D. Cian. The 2008 WITCH model: New model features and baseline. *FEEM Working Paper*, 2009.

[12] V. Bosetti, E. Massetti, and M. Tavoni. The WITCH model: Structure, baseline, solutions. *FEEM Working Paper*, 2007.

[13] P. Boyle. Options: A monte carlo approach. *Journal of Financial Economics*, 4:323–338, 1997.

[14] F. Branger, J. Ponssard, O. Sarter, and M. Sato. EU ETS, free allocations and activity level thresholds, the devil lies in the tails. *GRI working paper series*, (16), 2014.

[15] D. Brigo and F. Mercurio. Lognormal-mixture dynamics and calibration to market volatility smiles. *International Journal of Theoretical and Applied Finance*, 5(4):427–446, 2002.

[16] C. Brown and B. David. Skewness and kurtosis implied by option prices: A correction. *Journal of Financial Research*, 25(2):279–282, 1995.

[17] B. Burr. Assessing the sources of stranded asset risk: A proposed framework. *Journal of Sustainable Finance and Investment*, 7(1):37–53, 2017.

[18] T. Busch and V. Hoffmann. Emerging carbon constraints for corporate risk management. *Ecological Economics*, 62(3):518–528, 2007.

[19] B. Caldecott, J. Tilbury, and C. Carey. Stranded assets and scenarios. *Stranded Assets Programme*, 2014.

[20] R. Carmona, M. Fehr, and J. Hinz. Optimal stochastic control and carbon price formation. *SIAM Journal of Control and Optimziation*, 48(4):2168–2190, 2009.

[21] R. Carmona, M. Fehr, J. Hinz, and A. Porchet. Market design for emission trading schemes. *SIAM Review*, 52(3):403–452, 2010.

[22] R. Carmona and J. Hinz. Risk-neutral models for emission allowance prices and option valuation. *Journal of Management Science*, 57(8):1453–1468, 2011.

[23] P. Carr, H. Geman, and D. Madan. Pricing and hedging in incomplete markets. *Journal of Financial Economics*, 62(1):131–167, 2001.

[24] D. Cassidy, M. Hamp, and R. Rachid. Pricing european options with a log student's t-distribution: A gosset formula. *Physica A Statistical Mechanics and its Applications*.

[25] U. Cetin and M. Verschuere. Pricing and hedging in carbon emission markets. *International Journal of Theoretical and Applied Finance*, 12(7):949–967, 2009.

[26] M. Chesney and L. Taschini. The endogenous price dynamics of emission allowances and an application to option pricing. *Applied Mathematical Finance*, 19(5), 2012.

[27] European Commission. The EU emissions trading system (EU ETS), October 2013.

[28] C. Corrado and T. Su. Skewness and kurtosis in SP500 index returns implied by option prices. *The Journal of Financial Research*, 19(2), 1996.

[29] C. Corrado and T. Su. Implied volatility skews and stock return skewness and kurtosis implied by stock option prices. *The European Journal of Finance*, 1997.

[30] K. Daniel, R. B. Litterman, and G. Wagner. Applying asset pricing theory to calibrate the price of climate risk. *National Bureau of Economic Research working paper series*, 2016.

[31] G. Daskalakis and R. N. Markellos. Are the european carbon markets efficient? *Review of Future Markets*, 17(2):103–128, 2008.

[32] G. Daskalakis, D. Psychoyios, and R. N. Markellos. Modeling CO_2 emission allowance prices and derivatives: evidence from the European trading scheme. *Journal of Banking and Finance*, 33(7), 2009.

[33] M.H.A Davis. Option pricing in incomplete markets. In *Mathematics of Derivative Securities*, pages 216–226. Cambridge University Press, 1997.

[34] I. Diaz-Rainey, B. Robertson, and C. Wilson. Stranded research? Leading finance journals are silent on climate change. *Climatic Change*, 143(1):243–260.

[35] G. Dorfleitner, S. Utz, and M. Wimmer. Patience pays off - corporate social responsibility and long-term stock returns. *Journal of Sustainable Finance and Investment*, 8(2):132–157.

[36] J. Duan. The GARCH option pricing model. *Mathematical Finance*, 5(1):13–32, 1995.

[37] B. Dupire. Pricing with a smile. *Risk*, 7:28–20, 1994.

[38] SinoCarbon Education and Training. China carbon market monitor. *Partnership for Market Readiness*, 3, 2016.

[39] R. Engle and C. Mustafa. Implied ARCH models from options prices. *Journal of Econometrics*, 52(1-2):289–311, 1992.

[40] Battiston et al. A climate stress-test of the financial system. *Nature Climate Change*, 7(4), 2017.

[41] M. Fang. Lognormal mixture model for option pricing with applications to exotic options. Master's thesis, University of Waterloo, 2012.

[42] M. Fang, K. S. Tan, and T. Wirjanto. Managing climate and carbon risk in investment portfolios. *Society of Actuaries Research Reports*, 2018.

[43] M. Fang, K.S. Tan, and T. Wirjanto. Sustainable portfolio management under climate change. *Journal of Sustainable Finance and Investment*, 9(1):45–67, 2019.

[44] S. Figlewski. What does an option pricing model tell us about option prices? *Financial Analysts Journal*, 44(5), 1989.

[45] M. Gorgen, A. Jacob, M. Nerlinger, R. Riordan, M. Rohleder, and M. Wilkens. Carbon risk. *SSRN*.

[46] P. Griffin, A. Jaffe, D. Lont, and R. Dominguez-Faus. Science and the stock market: Investors' recognition of unburnable carbon. *Energy Economics*, (52), 2015.

[47] G. Grull and R. Kiesel. Pricing CO_2 permits using approximation approaches. *SSRN*, 2009.

[48] G. Grull and L. Taschini. Cap-and-trade properties under different hybrid scheme designs. *Journal of Environmental Economics and Management*, 61:107–118, 2011.

[49] M. Hardy. A regime-switching model of long-term stock returns. *North American Actuarial Journal*, 5(1), 2001.

[50] M. Hardy. *Investment Guarantees*. John Wiley and Sons Inc., 2003.

[51] L. Harris. Transactions data tests of the mixture of distributions hypothesis. *Journal of Financial and Quantitative Analysis*, 22(2), 1987.

[52] X. He and S. Zhu. Pricing european options with stochastic volatility under the minimal entropy martingale measure. *European Journal of Applied Mathematics*, 27(2):233–247, 2015.

[53] I. Henriques and P. Sadorsky. The relationship between environmental commitment and managerial perceptions of stakeholder importance. *Academy of Management Journal*, 42(1), 2017.

[54] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:237–243, 1993.

[55] Y. Huang. The price dynamics in the emissions market and the valuation of allowance derivatives. *SSRN*, 2010.

[56] J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42:281–300, 1987.

[57] C. Hunt and O. Weber. Fossil fuel divestment strategies: Financial and carbon-related consequences. *Organization and Environment*, 2018.

[58] G. Ibikunle and T. Steffen. European green mutual fund performance: A comparative analysis with their conventional and black peers. *Journal of Business Ethics*, 145(2), 2015.

[59] J.Seifert, M. Uhrig-Homberg, and M. Wagner. Dynamic behaviors of CO_2 spot prices. *Journal of Environmental Economics and Management*, 56(2):180–194, 2008.

[60] D. Kahneman and A. Tversky. Prospect Theory: An analysis of decision under risk. *Econometrica*, 47(2), 1979.

[61] J. Kakeu. Environmentally conscious investors and portfolio choice decisions. *Journal of Sustainable Finance and Investment*, 7(4):1–19, 2017.

[62] C. Kettner, A. Koppl, S. Schleicher, and G. Thenius. Stringency and distribution in the EU emissions trading scheme - the 2005 evidence. *SSRN*, 2007.

[63] A. Kibbey, R. Rasor, and B. Bostwick. Valuing oil and gas assets: the complexities and key considerations. *Finance Monthly*, 2017.

[64] J. Kilander. Calibrating an option pricing model under regime-switching volatility. Master's thesis, Stockholm School of Economics, 2007.

[65] D. Koehler and B. Bertocci. Stranded asset: What lies beneath. *UBS Asset Management Sustainable Investment Team Research Report*, 2016.

[66] A. LaPlante and C. Watson. Managing carbon risk: A look at environmentally conscious indices. *Global Risk Institute Report*, 2017.

[67] D. Leisen. Mixed lognormal distributions for derivatives pricing and risk-management. *Computing in Economics and Finance 2004*, 48, 2004.

[68] C. Li, S. Chen, and S. Lin. Pricing derivatives with modeling CO_2 emission allowance using a regime-switching jump diffusion model: with regime-switching risk premium. *The European Journal of Finance*, 22(10):887–908, 2016.

[69] F. Longstaff. Option pricing and the martingale restriction. *Review of Financial Studies*, 8(4):1091–1124, 1995.

[70] S. Markose and A. Alentorn. Option pricing and the implied tail index with the generalized extreme value (GEV) distribution. Computing in Economics and Finance 397, Society for Computational Economics, 2005.

[71] M. Meinshausen, N. Meinshausen, W. Hare, S. Raper, K. Frieler, R. Knutti, D. Frame, and Myles. Allen. Greenhouse gas emission targets for limiting global warming to 2 degree celsius. *Nature*, 458, 2009.

[72] W. Melick and C. Thomas. Recovering an asset's implied PDF from option prices: An application to crude oil during the gulf crisis. *Journal of Financial and Quantitative Analysis*, 32(1):91–115, 1997.

[73] Mercer. Investing in a time of climate change. *Mercer Investment Research Report*, 2015.

[74] L. Miller and W. Lu. S&P500's risk adjusted return was close to world best in 2017. *Bloomberg Market News*, 2018.

[75] W. Mnif. Incomplete market models of carbon emission markets. Master's thesis, The University of Western Ontario, 2012.

[76] W. Mnif, M. Davison, and D. Wu (Editor). *Quantitative Financial Risk Management*. Springer-Verlag, 2011.

[77] D. Nel. Carbon trading. Master's thesis, Oxford University, 2009.

[78] M. Neumann. Option pricing under the mixture of distributions hypothesis. *Preliminary Discussion Paper*, 2002.

[79] M. Oestreich and I. Tsiakas. Carbon emissions and stock returns: Evidence from the EU Emissions Trading Scheme. *Journal of Banking and Finance*, 58(9), 2015.

[80] United Nations Framework Conventions on Climate Change. *Kyoto Protocol Reference Manual on Accounting of Emissions and Assigned Amount*, 2008.

[81] M.S. Paolella and L. Taschini. An econometric analysis of emission allowance prices. *Journal of Banking and Finance*, 32(10):2022–2032, 2008.

[82] P. Pellizzari and A. Gamba. Utility based pricing of contingent claims in incomplete markets. *Applied Mathematical Finance*, 9(4):241–260, February 2002.

[83] Nordea Equity Research. Cracking the ESG code. *Nordic Ideas*, 2017.

[84] United Nations Climate Change Secretariat. Paris Agreement.

[85] G. Sevendson and M. Vesterdal. How to design greenhouse trading in the EU? *Energy Policy*, 31(14), 2003.

[86] Y. Shi. Modeling the time series dynamics of carbon emission markets. Master's thesis, Durham University, 2014.

[87] D. Shimko. Bounds of probability. *Risk*, 6(4), 1993.

[88] G. Sun. Modeling the spot price dynamics of CO_2 emission allowances: Evidence from the european union emission trading system. Master's thesis, Tilburg University, 2010.

[89] A. Trinks, B.Scholtens, M. Mulder, and L. Dam. Fossil fuel divestment and portfolio performance. *Ecological Economics*, 146(1):740–748, 2018.

[90] S. Truck and E. Benz. Modeling the price dynamics of CO_2 emission allowances. *Energy Economics*, 31(1):4–15, 2009.

[91] Q. Wu, M. Neelis, and C. Casanova. Chinese emission trading - initial assessment on allocation. *Ecofys*, 2014.

[92] Q. Zhang and J. Han. Option pricing in incomplete markets. *Applied Mathematics Letters*, 26(10):975–978, October 2013.

Appendices

This section contains proofs of the key results presented in the thesis.

Appendix A

Proof of (2.8)

Here we provide the proof of the distributional relationship presented in Equation (2.8). To evaluate the stochastic integral, we use the Riemann-Stieltjes sum approximation.

Partition the time interval (s,t) into equal-length intervals given by: $t_0 = s, t_1 = s+\Delta t, t_2 = s+2\Delta t, ..., t_i = s+i\Delta t, ...t_N = s+N\Delta t$, where $\Delta t = \frac{t-s}{N}$. This gives:

$$\begin{aligned}
L(\Delta t) &= \sum_{i=0}^{N-1} W(t_i)\Delta t \\
&= \sum_{i=0}^{N-1} \left[W(t_0) + (W(t_1) - W(t_0)) + (W(t_2) - W(t_1)) ... + (W(t_i) - W(t_i - 1))\right] \Delta t \\
&= \sum_{i=0}^{N-1} \left[W(t_0) + \sum_{j=1}^{i} \Delta W(t_j)\right] \Delta t \\
&= W(s)N\Delta t + \Delta W(t_1)(N-1)\Delta t + \Delta W(t_2)(N-2)\Delta t + ... + \Delta W(t_{N-1})\Delta t \\
&= (t-s)\left[W(s) + \Delta W(t_1)\frac{N-1}{N} + \Delta W(t_2)\frac{N-2}{N} + ... + \Delta W(t_{N-1})\frac{1}{N}\right]. \qquad \text{(A.1)}
\end{aligned}$$

Due to the property of Brownian Motions, we know that $\Delta W(t_i)$ are independent and identically distributed:

$$\Delta W(t_i) \equiv N\left(0, \frac{t-s}{N}\right), \quad \forall i \in \{1, 2, ...N-1\}. \qquad \text{(A.2)}$$

Hence, we arrive at the result that

$$\begin{aligned} L(\Delta t) &\equiv N\left(W(s)(t-s), \sum_{i=1}^{N-1}(N-i)^2\frac{(t-s)^3}{N^3}\right) \\ &\equiv N\left(W(s)(t-s), \frac{(N-1)(2N-1)}{6N^2}(t-s)^3\right). \end{aligned} \tag{A.3}$$

Taking the limit will give us the result:

$$\int_s^t (W(u))du = \lim_{N\to\infty} L(\Delta t) \equiv N\left(W(s)(t-s), \frac{(t-s)^3}{3}\right). \tag{A.4}$$

This is identical to what we have claimed in Equation (2.8) except for the trivial parameterization difference (the second parameter being variance instead of standard deviation). In addition, the way we approximate the integral within each partition interval does not affect the result. For example, we may also define the approximation sum as:

$$L(\Delta t) = \sum_{i=1}^{N}\frac{1}{2}\left(W(t_i) + W(t_{i-1})\right)\Delta t,$$

which is perhaps a more natural way for an integral approximation. This will give:

$$L(\Delta t) \equiv N\left(W(s)(t-s), \frac{N(2N-1)(2N+1)}{12N^3}(t-s)^3\right).$$

So the result remains the same upon taking the limit.

Appendix B

Proof of Proposition 2.2

From (2.3), given $\mid \mathcal{F}_t$:

$$S(t) = e^{-r(T-t)} E\left[G 1_{Q(t,T)>\alpha - Q(0,t)} \mid \mathcal{F}_t\right], \tag{B.1}$$

Let $\alpha(t) = \alpha - Q(0,t)$, then:

$$S(t) = e^{-r(T-t)} GPr\left(Q\left(t,T\right) > \alpha\left(t\right)\right). \tag{B.2}$$

From (2.8), we know $Q(t,T)$ follows a Normal distribution:

$$Q(t,T) \mid \mathcal{F}_t \sim N\left(\beta(t,T), \sigma\sqrt{\frac{(T-t)^3}{3}}\right), \tag{B.3}$$

where

$$\beta(t,T) = y(t)(T-t) + 0.5\mu\left(T^2 - t^2\right) - \mu t(T-t). \tag{B.4}$$

This gives:

$$\begin{aligned} S(t) &= e^{-r(T-t)} G\left(1 - Pr\left(Q\left(t,T\right) < \alpha\left(t\right)\right)\right) \\ &= e^{-r(T-t)} G\Phi\left(\frac{\beta(t,T) - (\alpha(t))}{\sigma\sqrt{\frac{(T-t)^3}{3}}}\right), \end{aligned} \tag{B.5}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function for a standard Normal variable.

Appendix C

Proof of (2.10)

First, we rearrange the result of Proposition 2.2 to give us:

$$\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)\sigma\sqrt{(T-t)^3} = \sqrt{3}\left(\beta(t,T)-\alpha(t)\right), \tag{C.1}$$

Then, using the expression of $Q(s,t)$ in (2.9), we get:

$$\begin{aligned} d\alpha(t) &= d\left(\alpha - \int_0^t y(u)du\right) \\ &= -y(t)dt. \end{aligned} \tag{C.2}$$

Now we need the process followed by $\beta(t,T)$. Notice that here the double differential $\frac{\partial^2\beta(t,T)}{\partial y(t)^2} = 0$. Hence, we have:

$$\begin{aligned} d\beta(t,T) &= d\left(y(t)(T-t) + 0.5\mu\left(T^2-t^2\right) - \mu t(T-t)\right) \\ &= (T-t)dy(t) - y(t)dt - \mu(T-t)dt. \end{aligned} \tag{C.3}$$

In addition, we define the following notation:

$$\eta(\alpha,\beta,t) \equiv \Phi\left(\frac{\beta(t,T)-\alpha(t)}{\sigma\sqrt{\frac{(T-t)^3}{3}}}\right). \tag{C.4}$$

Rewriting the result of Proposition 2.2 using this notation and taking the differential on both sides, we have:

$$\begin{aligned} dS(t) &= d(Ge^{-r(T-t)})\eta(\alpha,\beta,t) + Ge^{-r(T-t)}d(\eta(\alpha,\beta,t)) + d(Ge^{-r(T-t)})d(\eta(\alpha,\beta,t)) \\ &= rGe^{-r(T-t)}\eta(\alpha,\beta,t)dt + Ge^{-r(T-t)}d(\eta(\alpha,\beta,t)) \\ &= rS(t)dt + Ge^{-r(T-t)}d(\eta(\alpha,\beta,t)). \end{aligned} \tag{C.5}$$

Now we work out the differential of $\eta(\alpha,\beta,t)$. First observe that from (C.2) and (C.3) we have:

$$(d\alpha(t))^2 = 0$$

and

$$d\alpha(t)d\beta(t,T) = 0.$$

This simplifies the derivation. Taking differential on both sides of (C.4), we get:

$$d(\eta(\alpha,\beta,t)) = \frac{\partial\eta(\alpha,\beta,t)}{\partial t}dt + \frac{\partial\eta(\alpha,\beta,t)}{\partial\alpha}d\alpha(t) + \frac{\partial\eta(\alpha,\beta,t)}{\partial\beta}d\beta(t,T) + \frac{\partial^2\eta(\alpha,\beta,t)}{2\partial\beta^2}\left(d\beta(t,T)\right)^2. \tag{C.6}$$

We derive the expression for each term separately. Using the result in (C.2), we get:

$$\frac{\partial\eta(\alpha,\beta,t)}{\partial\alpha}d\alpha(t) = \phi\left(\frac{\beta(t,T) - (\alpha(t))}{\sigma\sqrt{\frac{(T-t)^3}{3}}}\right)\frac{y(t)dt}{\sigma\sqrt{\frac{(T-t)^3}{3}}}, \tag{C.7}$$

where $\phi(x)$ denotes the standard Normal density function evaluated at x, or, equivalently:

$$\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$

Substituting in the result of (C.1), (C.7) becomes:

$$\frac{\partial\eta(\alpha,\beta,t)}{\partial\alpha}d\alpha(t) = \frac{1}{\sqrt{2\pi}}e^{-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2}\frac{y(t)dt}{\sigma\sqrt{\frac{(T-t)^3}{3}}}. \tag{C.8}$$

Similarly, from the result of (C.3), we get:

$$\frac{\partial \eta(\alpha,\beta,t)}{\partial \beta} d\beta(t) = \frac{1}{\sqrt{2\pi}} e^{-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2} \frac{(T-t)dy(t) - y(t)dt - \mu(T-t)dt}{\sigma\sqrt{\frac{(T-t)^3}{3}}}. \quad \text{(C.9)}$$

For the second-order differential term, we have:

$$\begin{aligned}\frac{\partial^2 \eta(\alpha,\beta,t)}{2\partial \beta^2} &= \frac{1}{2\sqrt{2\pi}} e^{-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2} \left(\frac{-1}{\frac{\sigma^2(T-t)^3}{3}}\right) \Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right) \\ &= \frac{-3}{2\sigma^2(T-t)^3\sqrt{2\pi}} e^{-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2} \Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right). \quad \text{(C.10)}\end{aligned}$$

Based on (C.3) and (2.5), we also have:

$$(d\beta(t,T))^2 = (T-t)^2\sigma^2 dt. \quad \text{(C.11)}$$

Combining this result with (C.10), we get:

$$\frac{\partial^2 \eta(\alpha,\beta,t)}{2\partial \beta^2} (d\beta(t,T))^2 = \frac{-3}{2(T-t)\sqrt{2\pi}} e^{-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2} \Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right) dt. \quad \text{(C.12)}$$

Lastly, we have:

$$\begin{aligned}\frac{\partial \eta(\alpha,\beta,t)}{\partial t} dt &= \frac{1}{\sqrt{2\pi}} e^{-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2} \frac{3}{2}(T-t)^{-\frac{5}{2}} \frac{\beta(t,T) - \alpha(t)}{\sigma\sqrt{\frac{1}{3}}} dt \\ &= \frac{1}{\sqrt{2\pi}} e^{-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2} \left(\frac{3}{2(T-t)} \Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)\right) dt. \quad \text{(C.13)}\end{aligned}$$

Now we can substitute the results of (C.6), (C.8), (C.9), (C.11), (C.12) into (C.5), and obtain:

$$
\begin{aligned}
dS(t) = rS(t)dt &+ \frac{G}{\sqrt{2\pi}} e^{-r(T-t)-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2} \left(\frac{(T-t)dy(t) - \mu(T-t)dt}{\sigma\sqrt{\frac{(T-t)^3}{3}}} \right) \\
&+ \frac{G}{\sqrt{2\pi}} e^{-r(T-t)-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2} \frac{-3}{2(T-t)} \Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right) dt \\
&+ \frac{G}{\sqrt{2\pi}} e^{-r(T-t)-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2} \left(\frac{3}{2(T-t)} \Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right) \right) dt \\
&= rS(t)dt + \frac{\sqrt{3}G}{\sqrt{2\pi(T-t)}} e^{-\left(r(T-t)+0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2\right)} dW(t). \qquad \text{(C.14)}
\end{aligned}
$$

This proves (2.10).

Appendix D

Proof of (2.11)

We first prove the unconditional version of the result. From Proposition 2.2:

$$S(t) = e^{-r(T-t)} G\Phi\left(\frac{\beta(t,T) - \alpha(t)}{\sigma\sqrt{\frac{(T-t)^3}{3}}}\right), \tag{D.1}$$

$$S(0) = S_0 = e^{-rT} G\Phi\left(\frac{\beta(0,T) - \alpha(0)}{\sigma\sqrt{\frac{T^3}{3}}}\right). \tag{D.2}$$

Expand the numerator of the expression within the Normal Cumulative Distribution Function for $S(t)$. Notice that:

$$\beta(t,T) - \alpha(t) = y(t)(T-t) + 0.5\mu\left(T^2 - t^2\right) - \mu t(T-t) - \alpha + Q(0,t). \tag{D.3}$$

The results in (2.5) to (2.7) give:

$$y(t) = y_0 + \mu t + \sigma W(t),$$

$$\begin{aligned} Q(0,t) &= \beta(0,t) + \sigma\int_0^t W(u)du \\ &= y_0 t + 0.5\mu t^2 + \sigma\int_0^t W(u)du. \end{aligned} \tag{D.4}$$

So (D.3) can be further simplified as:

$$\begin{aligned}\beta(t,T)-\alpha(t) &= \left(y_0+\mu t+\sigma W(t)\right)(T-t)+0.5\mu\left(T^2-t^2\right)-\mu t(T-t)-\alpha \\ &\quad + y_0 t+0.5\mu t^2+\sigma\int_0^t W(u)du \\ &= y_0 T+0.5\mu T^2-\alpha+\sigma(T-t)W(t)+\sigma\int_0^t W(u)du. \end{aligned} \tag{D.5}$$

Then, from (D.2), we have

$$\beta(0,T)-\alpha(0)=\sigma\sqrt{\frac{T^3}{3}}\Phi^{-1}\left(\frac{S_0e^{rT}}{G}\right). \tag{D.6}$$

We also have the result:

$$\beta(0,T)-\alpha(0)=y_0T+0.5\mu T^2-\alpha. \tag{D.7}$$

Substituting the results of (D.6) and (D.7) to (D.5) gives:

$$\beta(t,T)-\alpha(t)=\sigma\sqrt{\frac{T^3}{3}}\Phi^{-1}\left(\frac{S_0e^{rT}}{G}\right)+\sigma(T-t)W(t)+\sigma\int_0^t W(u)du. \tag{D.8}$$

Substituting the result of (D.8) to (D.1) and simplifying terms gives:

$$S(t)=e^{-r(T-t)}G\Phi\left(\frac{\sqrt{\frac{T^3}{3}}\Phi^{-1}\left(\frac{S_0e^{rT}}{G}\right)+(T-t)W(t)+\int_0^t W(u)du}{\sqrt{\frac{(T-t)^3}{3}}}\right). \tag{D.9}$$

We then generalize this solution by conditioning on the appropriate σ-fields. First obtain the time-s unconditional expression:

$$S(s)=e^{-r(T-s)}G\Phi\left(\frac{\sqrt{\frac{T^3}{3}}\Phi^{-1}\left(\frac{S_0e^{rT}}{G}\right)+(T-s)W(s)+\int_0^s W(u)du}{\sqrt{\frac{(T-s)^3}{3}}}\right), \tag{D.10}$$

which can be rearranged to get:

$$\sqrt{\frac{T^3}{3}}\Phi^{-1}\left(\frac{S_0e^{rT}}{G}\right)=\Phi^{-1}\left(\frac{S(s)e^{r(T-s)}}{G}\right)\sqrt{\frac{(T-s)^3}{3}}-(T-s)W(s)-\int_0^s W(u)du. \tag{D.11}$$

Substituting this result into Equation (D.9) and simplifying terms gives:

$$S(t) = e^{-r(T-t)} G\Phi \left(\frac{\sqrt{\frac{(T-s)^3}{3}}\Phi^{-1}\left(\frac{S_s e^{r(T-s)}}{G}\right) + \lambda(s,t)}{\sqrt{\frac{(T-t)^3}{3}}} \right), \tag{D.12}$$

where

$$\begin{aligned} \lambda(s,t) &= (T-t)W(t) - (T-s)W(s) + \int_s^t W(u)du \\ &= T\left(W(t) - W(s)\right) + sW(s) - tW(t) + \int_s^t W(u)du. \end{aligned} \tag{D.13}$$

This result is the same as that claimed in (2.11).

Appendix E

Proof of (2.19)

It suffices to find the distribution of the stochastic component in (2.18) and the rest of the results directly follow. Partition the time interval (s,t) into equal-length intervals given by: $t_0 = s, t_1 = s + \Delta t, t_2 = s + 2\Delta t, ..., t_i = s + i\Delta t, ...t_N = s + N\Delta t$, where $\Delta t = \frac{t-s}{N}$. Applying the Riemann-Stieltjes definition gives the distribution of the Ito Integral under the risk neutral measure $\mathcal{Q}$:

$$\int_s^t e^{\kappa u} dW(u) = \lim_{N\to\infty} \sum_{i=0}^{N-1} e^{\kappa t} \left(W(t_i + \Delta t) - W(t_i)\right).$$

By the property of standard Brownian motions, the summands are independently Normally distributed with a common mean of 0. The Ito Integral hence converges to a zero-mean normal variable. To find the variance, apply Ito's Isometry:

$$\begin{aligned} E\left[\left(\int_s^t e^{\kappa u} dW(u)\right)^2\right] &= E\left[\int_s^t e^{2\kappa u} du\right] \\ &= \frac{1}{2\kappa}\left(e^{2\kappa t} - e^{2\kappa s}\right). \end{aligned} \tag{E.1}$$

It thus follows that the stochastic component is also normally distributed:

$$\sigma e^{-\kappa t} \int_s^t e^{\kappa u} dW(u) \equiv N\left(0, \sigma\sqrt{\frac{\left(1 - e^{-2\kappa(t-s)}\right)}{2\kappa}}\right). \tag{E.2}$$

Hence, by adding the deterministic component in (2.18) to the mean parameter, the following result is reached:

$$Q(0,t) \mid \mathcal{F}_s \sim N\left(\beta(s,t), \sigma\sqrt{\frac{\left(1-e^{-2\kappa(t-s)}\right)}{2\kappa}}\right), \tag{E.3}$$

where

$$\beta(s,t) = Q(0,s)e^{-\kappa(t-s)} + \theta\left(1-e^{-\kappa(t-s)}\right). \tag{E.4}$$

To obtain the result in (2.19), filter out the known emission amount in $\mathcal{F}_s$.

$$\begin{aligned} Q(s,t) &= Q(0,t) - Q(0,s) \\ &= \beta(s,t) - Q(0,s) + \sigma e^{-\kappa t}\int_s^t e^{\kappa u} dW(u) \\ &= (\theta - Q(0,s))\left(1-e^{-\kappa(t-s)}\right) + \sigma e^{-\kappa t}\int_s^t e^{\kappa u} dW(u) \\ &= \beta^*(s,t) + \sigma e^{-\kappa t}\int_s^t e^{\kappa u} dW(u). \end{aligned} \tag{E.5}$$

Therefore,

$$Q(s,t) \sim N\left(\beta^*(s,t), \sigma\sqrt{\frac{\left(1-e^{-2\kappa(t-s)}\right)}{2\kappa}}\right), \tag{E.6}$$

where

$$\beta^*(s,t) = (\theta - Q(0,s))\left(1-e^{-\kappa(t-s)}\right). \tag{E.7}$$

Appendix F

Proof of Proposition 2.3

Similar to the proof for Proposition 2.2., this is a relatively trivial derivation. From (2.4), we have:

$$S(t) \mid \mathcal{F}_t = e^{-r(T-t)} E\left[G 1_{Q(t,T)>\alpha - Q(0,t)} \mid \mathcal{F}_t\right] \tag{F.1}$$

Let $\alpha(t) = \alpha - Q(0,t)$, then:

$$S(t) \mid \mathcal{F}_t = e^{-r(T-t)} G Pr\left(Q\left(t,T\right) > \alpha\left(t\right)\right) \tag{F.2}$$

From (2.19), we know $Q(t,T)$ follows a normal distribution:

$$Q(t,T) \mid \mathcal{F}_t \sim N\left(\beta^*(t,T), \sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}}\right), \tag{F.3}$$

where

$$\beta^*(t,T) = \left(\theta - Q(0,t)\right)\left(1-e^{-\kappa(T-t)}\right). \tag{F.4}$$

This gives:

$$\begin{aligned} S(t) &= e^{-r(T-t)} G\left(1 - Pr\left(Q\left(t,T\right) < \alpha\left(t\right)\right)\right) \\ &= e^{-r(T-t)} G \Phi\left(\frac{\left(\beta^*(t,T) - \alpha(t)\right)}{\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}}}\right), \end{aligned} \tag{F.5}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function for a standard Normal variable.

Appendix G

Proof of (2.23)

First, rearrange the result of (2.21) to obtain

$$\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)\sigma\sqrt{1-e^{-2k(T-t)}}=\sqrt{2\kappa}\left(\beta(t,T)-\alpha\right), \tag{G.1}$$

where $\beta(t,T)$ is defined in (2.22).

Now we need the diffusion process followed by $\beta(t,T)$. Notice that here the double differential $\frac{\partial^2\beta(t,T)}{\partial Q(0,t)^2}=0$. Hence, we have:

$$\begin{aligned} d\beta(t,T) &= \frac{\partial\beta}{\partial Q}dQ(0,t)+\frac{\partial\beta}{\partial t}dt \\ &= e^{-\kappa(T-t)}dQ(0,t)+\kappa Q(0,t)e^{-\kappa(T-t)}-\kappa\theta e^{-\kappa(T-t)} \\ &= \sigma e^{-\kappa(T-t)}dW(t), \end{aligned} \tag{G.2}$$

where to get the last line we bring in the $Q(0,t)$ dynamics in (2.17) and simplify. In addition, we define the following notation:

$$\eta(\beta,t)\equiv\Phi\left(\frac{\beta(t,T)-\alpha}{\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}}}\right). \tag{G.3}$$

Rewrite (2.21) using this notation and take the differential on both sides, we have:

$$\begin{aligned} dS(t) &= d(Ge^{-r(T-t)})\eta(\beta,t) + Ge^{-r(T-t)}d(\eta(\beta,t)) + d(Ge^{-r(T-t)})d(\eta(\beta,t)) \\ &= rGe^{-r(T-t)}\eta(\beta,t)dt + Ge^{-r(T-t)}d(\eta(\beta,t)) \\ &= rS(t)dt + Ge^{-r(T-t)}d(\eta(\beta,t)). \end{aligned} \tag{G.4}$$

Now we work out the differential of $\eta(\beta,t)$.

$$d(\eta(\beta,t)) = \frac{\partial\eta(\beta,t)}{\partial t}dt + \frac{\partial\eta(\beta,t)}{\partial\beta}d\beta(t,T) + \frac{\partial^2\eta(\beta,t)}{2\partial\beta^2}\left(d\beta(t,T)\right)^2. \tag{G.5}$$

The expression for each term is then derived separately. Using the result in (G.2), we get:

$$\frac{\partial\eta(\beta,t)}{\partial\beta}d\beta(t,T) = \phi\left(\frac{\beta(t,T)-\alpha}{\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}}}\right)\frac{e^{-\kappa(T-t)}dW(t)}{\sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}}}, \tag{G.6}$$

where $\phi(x)$ denotes the standard Normal density function evaluated at x, or, equivalently:

$$\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$

Substituting in the result of (G.1), (G.6) becomes:

$$\frac{\partial\eta(\beta,t)}{\partial\beta}d\beta(t,T) = \frac{1}{\sqrt{2\pi}}e^{-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2}\frac{e^{-\kappa(T-t)}dW(t)}{\sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}}}. \tag{G.7}$$

For the second-order differential term, we have:

$$\begin{aligned} \frac{\partial^2\eta(\beta,t)}{2\partial\beta^2} &= \frac{1}{2\sqrt{2\pi}}e^{-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2}\left(\frac{-2\kappa}{\sigma^2\left(1-e^{-2\kappa(T-t)}\right)}\right)\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right) \\ &= \frac{-\kappa e^{-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2}}{\sqrt{2\pi}\sigma^2\left(1-e^{-2\kappa(T-t)}\right)}\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right). \end{aligned} \tag{G.8}$$

In addition, we have:

$$\left(d\beta(t,T)\right)^2 = \sigma^2e^{-2\kappa(T-t)}dt. \tag{G.9}$$

This gives:

$$\frac{\partial^2\eta(\beta,t)}{2\partial\beta^2}\left(d\beta(t,T)\right)^2 = \frac{-\kappa e^{-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2-2\kappa(T-t)}}{\sqrt{2\pi}\left(1-e^{-2\kappa(T-t)}\right)}\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)dt. \tag{G.10}$$

Lastly, we have:

$$\begin{aligned}\frac{\partial\eta(\beta,t)}{\partial t}dt &= \frac{dt}{\sqrt{2\pi}}e^{-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2}\left(\frac{2\kappa e^{-2\kappa(T-t)}}{2(1-e^{-2\kappa(T-t)})}\right)\left(\frac{\beta(t,T)-\alpha}{\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}}}\right)\\ &= \frac{\kappa e^{-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2-2\kappa(T-t)}}{\sqrt{2\pi}\left(1-e^{-2\kappa(T-t)}\right)}\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)dt. \end{aligned}\tag{G.11}$$

We now substitute the results of (G.5), (G.6), (G.10), and (G.11) into (G.4) and simplify to get:

$$\begin{aligned}dS(t) &= rS(t)dt + Ge^{-r(T-t)}\left(\frac{\partial\eta(\beta,t)}{\partial t}dt + \frac{\partial\eta(\beta,t)}{\partial\beta}d\beta(t,T) + \frac{\partial^2\eta(\beta,t)}{2\partial\beta^2}\left(d\beta(t,T)\right)^2\right)\\ &= rS(t)dt + Ge^{-r(T-t)}\frac{1}{\sqrt{2\pi}}e^{-0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2}\frac{e^{-\kappa(T-t)}dW(t)}{\sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}}}\\ &= rS(t)dt + \frac{G\sqrt{\kappa}e^{-\kappa(T-t)}}{\sqrt{\pi\left(1-e^{-2\kappa(T-t)}\right)}}e^{-\left(r(T-t)+0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2\right)}dW(t)\\ &= rS(t)dt + \frac{G\sqrt{\kappa}}{\sqrt{\pi\left(e^{2\kappa(T-t)}-1\right)}}e^{-\left(r(T-t)+0.5\Phi^{-1}\left(\frac{S(t)e^{r(T-t)}}{G}\right)^2\right)}dW(t)\end{aligned}\tag{G.12}$$

where the initial condition is $S(0) = S_0$ and the terminal condition is given by (2.3). This proves (2.23).

Appendix H

Proof of (2.24)

We first prove the unconditional version of the result. From (2.21), we get:

$$S(t) = e^{-r(T-t)} G \Phi \left(\frac{(\beta(t,T) - \alpha)}{\sigma \sqrt{\frac{\left(1 - e^{-2\kappa(T-t)}\right)}{2\kappa}}} \right), \tag{H.1}$$

$$S(0) = e^{-rT} G \Phi \left(\frac{(\beta(0,T) - \alpha)}{\sigma \sqrt{\frac{\left(1 - e^{-2\kappa T}\right)}{2\kappa}}} \right), \tag{H.2}$$

and

$$\beta(t,T) - \alpha = Q(0,t) e^{-\kappa(T-t)} + \theta \left(1 - e^{-\kappa(T-t)}\right) - \alpha. \tag{H.3}$$

Rearranging (H.2) leads to:

$$\sigma \sqrt{\frac{\left(1 - e^{-2\kappa T}\right)}{2\kappa}} \Phi^{-1} \left(\frac{S_0 e^{rT}}{G} \right) = \beta(0,T) - \alpha = \theta \left(1 - e^{-\kappa T}\right) - \alpha. \tag{H.4}$$

Combining the results of (H.4) and (H.3) gives:

$$\beta(t,T) - \alpha = e^{\kappa t} \left(\Phi^{-1} \left(\frac{S_0 e^{rT}}{G} \right) \sigma \sqrt{\frac{\left(1 - e^{-2\kappa T}\right)}{2\kappa}} \right) + \left(e^{\kappa t} - 1\right)(\alpha - \theta) + Q(0,t) e^{-\kappa(T-t)}. \tag{H.5}$$

The solution to $Q(0,t)$ based on the initial condition is given by (2.18):

$$Q(0,t) = \theta(1 - e^{-\kappa t}) + \sigma e^{-\kappa t} \int_0^t e^{au} dW(u). \tag{H.6}$$

Substituting this result into Equation (H.5) and simplifying terms gives:

$$\begin{aligned}
\beta(t,T) - \alpha &= e^{\kappa t} \left(\Phi^{-1} \left(\frac{S_0 e^{rT}}{G} \right) \sigma \sqrt{\frac{\left(1 - e^{-2\kappa T}\right)}{2\kappa}} \right) + \sigma^{-\kappa T} \int_0^t e^{\kappa u} dW(u) \\
&\quad + \left(e^{\kappa t} - 1\right)(\alpha - \theta) + e^{-\kappa T}\left(e^{\kappa t} - 1\right)\theta \\
&= e^{\kappa t} \left(\Phi^{-1} \left(\frac{S_0 e^{rT}}{G} \right) \sigma \sqrt{\frac{\left(1 - e^{-2\kappa T}\right)}{2\kappa}} \right) + \delta(t,T),
\end{aligned}$$

where

$$\delta(t,T) = (e^{\kappa t} - 1)\left(\alpha - (1 - e^{-\kappa T})\theta\right) + \sigma e^{-\kappa T} \int_0^t e^{\kappa u} dW(u). \tag{H.7}$$

Substituting this result into (H.1) gives:

$$S(t) = e^{-r(T-t)} G \Phi \left(\frac{\sigma e^{\kappa t} \sqrt{\frac{\left(1 - e^{-2\kappa T}\right)}{2\kappa}} \Phi^{-1} \left(\frac{S_0 e^{rT}}{G} \right) + \delta(t,T)}{\sigma \sqrt{\frac{\left(1 - e^{-2\kappa (T-t)}\right)}{2\kappa}}} \right), \tag{H.8}$$

We then generalize this solution by conditioning on the appropriate σ-fields. First obtain the time-s unconditional expressions:

$$S(s) = e^{-r(T-s)} G \Phi \left(\frac{\sigma e^{\kappa s} \sqrt{\frac{\left(1 - e^{-2\kappa T}\right)}{2\kappa}} \Phi^{-1} \left(\frac{S_0 e^{rT}}{G} \right) + \delta(s,T)}{\sigma \sqrt{\frac{\left(1 - e^{-2\kappa (T-s)}\right)}{2\kappa}}} \right), \tag{H.9}$$

where

$$\delta(s,T) = (e^{\kappa s} - 1)\left(\alpha - (1 - e^{-\kappa T})\theta\right) + \sigma e^{-\kappa T} \int_0^s e^{\kappa u} dW(u), \tag{H.10}$$

which can be rearranged to obtain:

$$\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-s)}\right)}{2\kappa}}\Phi^{-1}\left(\frac{S_s e^{r(T-s)}}{G}\right)=\sigma e^{\kappa s}\sqrt{\frac{\left(1-e^{-2\kappa T}\right)}{2\kappa}}\Phi^{-1}\left(\frac{S_0 e^{rT}}{G}\right)+\delta(s,T). \quad \text{(H.11)}$$

Rearranging Equation (H.8) and substituting in this result leads to:

$$\begin{aligned}&\sigma e^{\kappa t}\sqrt{\frac{\left(1-e^{-2\kappa T}\right)}{2\kappa}}\Phi^{-1}\left(\frac{S_0 e^{rT}}{G}\right)+\delta(t,T)\\&=e^{\kappa(t-s)}\left(\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-s)}\right)}{2\kappa}}\Phi^{-1}\left(\frac{S_s e^{r(T-s)}}{G}\right)-\delta(s,T)\right)+\delta(t,T). \quad \text{(H.12)}\end{aligned}$$

Using (H.7) and (H.10), we define the term:

$$\begin{aligned}&\nu(s,t)=\delta(t,T)-e^{\kappa(t-s)}\delta(s,T)=\left(e^{\kappa(t-s)}-1\right)\left(\alpha-(1-e^{-\kappa T})\theta\right)+\sigma e^{-\kappa T}\int_0^t e^{\kappa u}dW(u)\\&-\sigma e^{-\kappa T}e^{\kappa(t-s)}\int_0^s e^{\kappa u}dW(u)\\&=\left(e^{\kappa(t-s)}-1\right)\left(\alpha-(1-e^{-\kappa T})\theta\right)+\sigma e^{-\kappa T}\left(\int_s^t e^{\kappa u}dW(u)+\left(1-e^{\kappa(t-s)}\right)\int_0^s e^{\kappa u}dW(u)\right). \quad \text{(H.13)}\end{aligned}$$

Substituting the result of (H.13) into (H.12) gives:

$$\begin{aligned}&\sigma e^{\kappa t}\sqrt{\frac{\left(1-e^{-2\kappa T}\right)}{2\kappa}}\Phi^{-1}\left(\frac{S_0 e^{rT}}{G}\right)+\delta(t,T)\\&=e^{\kappa(t-s)}\left(\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-s)}\right)}{2\kappa}}\Phi^{-1}\left(\frac{S_s e^{r(T-s)}}{G}\right)\right)+\nu(s,t). \quad \text{(H.14)}\end{aligned}$$

Substituting this result into (H.1) gives

$$S(t)=e^{-r(T-t)}G\Phi\left(\frac{\sigma e^{\kappa(t-s)}\sqrt{\frac{\left(1-e^{-2\kappa(T-s)}\right)}{2\kappa}}\Phi^{-1}\left(\frac{S_s e^{r(T-s)}}{G}\right)+\nu(s,t)}{\sigma\sqrt{\frac{\left(1-e^{-2\kappa(T-t)}\right)}{2\kappa}}}\right), \quad \text{(H.15)}$$

where $\nu(s,t)$ is defined by (H.13). This proves the result of (2.24).

Appendix I

Proof of (2.32) and (2.33)

Using the result of (2.31), given $\mathcal{F}_s$, $0 \leq s < t < T$:

$$\begin{aligned} Q(s,t) &= \int_s^t y(u)du \\ &= \int_s^t \left(y(s)e^{-\kappa(u-s)} + \theta(1 - e^{-(u-s)}) + \sigma e^{-\kappa u} \int_s^u e^{kv} dW(v) \right) du. \\ &= \frac{y(s)}{\kappa} \left(1 - e^{-\kappa(t-s)}\right) + \theta \left(t - s - \frac{1}{\kappa} \left(1 - e^{-\kappa(t-s)}\right) \right) + \sigma\psi \end{aligned} \tag{I.1}$$

where

$$\psi = \int_s^t \int_s^u e^{-\kappa(u-v)} dW(v)du. \tag{I.2}$$

The double integral in (I.2) can be evaluated using integration by parts. Define the following:

$$M = \int_s^u e^{\kappa v} dW(v). \tag{I.3}$$

$$N = -\frac{e^{-\kappa u}}{\kappa}. \tag{I.4}$$

Then, observe that:

$$\begin{aligned}
\psi &= \int_s^t M dN = \int_s^t d(MN) - \int_s^t d(M)N - \int_s^t d(M)d(N) \\
&= \int_s^t d\left(\left(\int_s^u e^{\kappa v} dW(v)\right)\left(\frac{-e^{-\kappa u}}{\kappa}\right)\right) - \int_s^t \left(\frac{-e^{-\kappa u}}{\kappa}\right) e^{\kappa u} dW(u) \\
&= -\frac{-e^{-\kappa t}}{\kappa}\int_s^t e^{\kappa v} dW(v) + \int_s^t \frac{dW(u)}{\kappa} \\
&= \frac{1}{\kappa}\int_s^t \left(1 - e^{-\kappa(t-u)}\right) dW(u). \qquad \text{(I.5)}
\end{aligned}$$

Substituting (I.5) into (I.1) gives the result of (2.32).

To prove (2.33), notice that stochasticity in the expression for $Q(s,t)$ is driven by the integral component of (I.5), which is an Ito integral with a driftless integrand and hence a conditionally normal variable with zero mean. The variance of the integral is calculated by using Ito's Isometry as follows:

$$\begin{aligned}
E\left(\left(\int_s^t \left(1 - e^{-\kappa(t-u)}\right) dW(u)\right)^2\right) &= \int_s^t \left(1 - e^{-\kappa(t-u)}\right)^2 du \qquad \text{(I.6)} \\
. &= t - s - \frac{2}{\kappa}\left(1 - e^{-\kappa(t-s)}\right) + \frac{1}{2\kappa}\left(1 - e^{-2\kappa(t-s)}\right).
\end{aligned}$$

Combining these results with (2.32), (2.33) directly follows.

Appendix J

Proof of Proposition 2.4

This is again a relatively trivial derivation. From (2.4), we have:

$$S(t) \mid \mathcal{F}_t = e^{-r(T-t)} E\left[G 1_{Q(t,T)>\alpha - Q(0,t)} \mid \mathcal{F}_t\right]. \tag{J.1}$$

Let $\alpha(t) = \alpha - Q(0,t)$, then:

$$S(t) \mid \mathcal{F}_t = e^{-r(T-t)} G Pr\left(Q\left(t,T\right) > \alpha\left(t\right)\right). \tag{J.2}$$

From (2.33), we know $Q(t,T)$ follows a Normal distribution:

$$Q(t,T) \mid \mathcal{F}_t \sim N\left(\beta(t,T), \frac{\sigma}{\kappa}\sqrt{T-t-\frac{2}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{1}{2\kappa}\left(1-e^{-\kappa(T-t)}\right)}\right), \tag{J.3}$$

where

$$\beta(t,T) = \frac{y(t)}{\kappa}\left(1-e^{-\kappa(T-t)}\right) + \theta\left(T-t-\frac{1}{\kappa}\left(1-e^{-\kappa(T-t)}\right)\right). \tag{J.4}$$

This gives:

$$\begin{aligned} S(t) &= e^{-r(T-t)} G\left(1 - Pr\left(Q\left(t,T\right) < \alpha\left(t\right)\right)\right) \\ &= e^{-r(T-t)} G \Phi\left(\frac{\beta(t,T)-\alpha(t)}{\frac{\sigma}{\kappa}\sqrt{T-t-\frac{2}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-t)}\right)}}\right), \end{aligned} \tag{J.5}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function for a standard Normal variable.

Appendix K

Proof of (2.34)

To simplify the notation, first define:

$$\gamma(t,T) = \frac{\sigma}{\kappa}\sqrt{T-t-\frac{2}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-t)}\right)} \tag{K.1}$$

$$\eta(\alpha,\beta,\gamma) = \Phi\left(\frac{\beta(t,T)-\alpha(t)}{\gamma(t,T)}\right), \tag{K.2}$$

where $\beta(t,T)$ and $\alpha(t)$ are given in Proposition 2.4.

Take the differential of $S(t)$, whose expression is given in Proposition 2.4, we obtain:

$$dS(t) = rS(t)dt + Ge^{-r(T-t)}d\eta(\alpha,\beta,\gamma), \tag{K.3}$$

where, upon eliminating the cross differentials that evaluate to zero, we have:

$$d\eta(\alpha,\beta,\gamma) = \frac{\partial\eta}{\partial\alpha}d\alpha(t) + \frac{\partial\eta}{\partial\beta}d\beta(t,T) + \frac{\partial\eta}{\partial\gamma}d\gamma(t,T) + \frac{\partial^2\eta}{2\partial\beta^2}d\beta(t,T)^2. \tag{K.4}$$

The terms are evaluated individually. For the first term:

$$\frac{\partial\eta}{\partial\alpha}d\alpha(t) = \phi\left(\frac{\beta(t,T)-\alpha(t)}{\gamma(t,T)}\right)\frac{y(t)}{\gamma(t,T)}dt. \tag{K.5}$$

For the second term:

$$\frac{\partial\eta}{\partial\beta} = \phi\left(\frac{\beta(t,T)-\alpha(t)}{\gamma(t,T)}\right)\frac{1}{\gamma(t,T)}, \tag{K.6}$$

and

$$d\beta(t,T) = \frac{\partial \beta}{\partial y} dy(t) + \frac{\partial \beta}{\partial t} dt \tag{K.7}$$
$$= \frac{1 - e^{-\kappa(T-t)}}{\kappa} dy(t) + (\theta e^{-\kappa(T-t)} - \theta - y(t) e^{-\kappa(T-t)}) dt$$
$$= -y(t)dt + \frac{1 - e^{-\kappa(T-t)}}{\kappa} dW(t).$$

For the third term:

$$\frac{\partial \eta}{\partial \gamma} = -\phi \left(\frac{\beta(t,T) - \alpha(t)}{\gamma(t,T)} \right) \frac{\beta(t,T) - \alpha(t)}{\gamma(t,T)^2}, \tag{K.8}$$

and

$$d\gamma(t,T) = \frac{\sigma}{2\kappa} \left(\frac{\gamma(t,T)\kappa}{\sigma} \right)^{-1} d\left(T - t - \frac{2}{\kappa} \left(1 - e^{-\kappa(T-t)} \right) + \frac{1}{2\kappa} \left(1 - e^{-2\kappa(T-t)} \right) \right) \tag{K.9}$$
$$= -\frac{\sigma^2 \left(1 - e^{-\kappa(T-t)} \right)^2}{2\kappa^2 \gamma(t,T)} dt.$$

Finally, for the second-order differential:

$$\frac{\partial^2 \eta}{\partial \beta^2} = -\phi \left(\frac{\beta(t,T) - \alpha(t)}{\gamma(t,T)} \right) \frac{\beta(t,T) - \alpha(t)}{\gamma(t,T)^3}, \tag{K.10}$$

$$d\beta(t,T)^2 = \left(\frac{1 - e^{-\kappa(T-t)}}{\kappa} \right)^2 \sigma^2 dt. \tag{K.11}$$

Substituting (K.5) - (K.11) into (K.4) and simplifying the result gives

$$d\eta(\alpha, \beta, \gamma) = \phi \left(\frac{\beta(t,T) - \alpha(t)}{\gamma(t,T)} \right) \frac{\sigma \left(1 - e^{-\kappa(T-t)} \right)}{\kappa \gamma(t,T)} dW(t). \tag{K.12}$$

In addition, reorganizing the expression of $S(t)$ in Proposition 2.4 gives:

$$\left(\frac{\beta(t,T) - \alpha(t)}{\gamma(t,T)} \right) = \Phi^{-1} \left(\frac{S(t) e^{r(T-t)}}{G} \right). \tag{K.13}$$

Substituting (K.1), (K.12), and (K.13) into (K.3) and simplifying the result give the allowance price process in (2.34).

Appendix L

Proof of (2.35)

We first prove the unconditional version of the result. From Proposition 2.4:

$$S(t) = e^{-r(T-t)} G\Phi\left(\frac{\beta(t,T) - \alpha(t)}{\frac{\sigma}{\kappa}\sqrt{T-t-\frac{2}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-t)}\right)}}\right), \tag{L.1}$$

$$S(0) = S_0 = e^{-rT} G\Phi\left(\frac{\beta(0,T) - \alpha}{\frac{\sigma}{\kappa}\sqrt{T-\frac{2}{\kappa}\left(1-e^{-\kappa T}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa T}\right)}}\right), \tag{L.2}$$

where

$$\beta(t,T) = \frac{y(t)}{\kappa}\left(1-e^{-\kappa(T-t)}\right) + \theta\left(T-t-\frac{1}{\kappa}\left(1-e^{-\kappa(T-t)}\right)\right), \tag{L.3}$$

$$\alpha(t) = \alpha - Q(0,t), \tag{L.4}$$

$$\beta(0,T) = \frac{y_0}{\kappa}\left(1-e^{-\kappa T}\right) + \theta\left(T-\frac{1}{\kappa}\left(1-e^{-\kappa T}\right)\right). \tag{L.5}$$

Given only the initial conditions (i.e. $\mathcal{F}_0$), from (2.31)

$$y(t) = y_0 e^{-\kappa t} + \theta(1-e^{-\kappa t}) + \sigma e^{-\kappa t}\int_0^t e^{ku} dW(u). \tag{L.6}$$

From (2.32)

$$Q(0,t) = \beta(0,t) + \frac{\sigma}{\kappa}\int_0^t \left(1 - e^{-\kappa(t-u)}\right) dW(u), \tag{L.7}$$

where

$$\beta(0,t) = \frac{y_0}{\kappa}\left(1 - e^{-\kappa t}\right) + \theta\left(t - \frac{1}{\kappa}\left(1 - e^{-\kappa t}\right)\right).$$

Substitute (L.6) into (L.3) and (L.7) into (L.4). The following result is obtained after a tedious algebraic simplification:

$$\beta(t,T) - \alpha(t) = \beta(0,T) - \alpha + \frac{\sigma}{\kappa}\xi(0,t), \tag{L.8}$$

where

$$\xi(0,t) = \int_0^t \left(1 - e^{-\kappa(T-u)}\right) dW(u). \tag{L.9}$$

Finally, from (L.2), we have:

$$\beta(0,T) - \alpha = \Phi^{-1}\left(\frac{S_0 e^{rT}}{G}\right)\frac{\sigma}{\kappa}\sqrt{T - \frac{2}{\kappa}\left(1 - e^{-\kappa T}\right) + \frac{1}{2\kappa}\left(1 - e^{-2\kappa T}\right)}. \tag{L.10}$$

Substituting (L.8) and (L.10) into (L.1) leads to:

$$S(t) = e^{-r(T-t)}G\Phi\left(\frac{\sqrt{T - \frac{2}{\kappa}\left(1 - e^{-\kappa T}\right) + \frac{1}{2\kappa}\left(1 - e^{-2\kappa T}\right)}\Phi^{-1}\left(\frac{S_0 e^{rT}}{G}\right) + \xi(0,t)}{\sqrt{T - t - \frac{2}{\kappa}\left(1 - e^{-\kappa(T-t)}\right) + \frac{1}{2\kappa}\left(1 - e^{-2\kappa(T-t)}\right)}}\right). \tag{L.11}$$

We then generalize this solution by conditioning on $\mathcal{F}_s$. First obtain the unconditional expression:

$$S(s) = e^{-r(T-s)}G\Phi\left(\frac{\sqrt{T - \frac{2}{\kappa}\left(1 - e^{-\kappa T}\right) + \frac{1}{2\kappa}\left(1 - e^{-2\kappa T}\right)}\Phi^{-1}\left(\frac{S_0 e^{rT}}{G}\right) + \xi(0,s)}{\sqrt{T - s - \frac{2}{\kappa}\left(1 - e^{-\kappa(T-s)}\right) + \frac{1}{2\kappa}\left(1 - e^{-2\kappa(T-s)}\right)}}\right). \tag{L.12}$$

For notational convenience, define:

$$\gamma(t,T) = \sqrt{T-t-\frac{2}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-t)}\right)}. \tag{L.13}$$

Hence, given $\mathcal{F}_s$, (L.12) leads to:

$$\xi(0,s) = \Phi^{-1}\left(\frac{S_s e^{r(T-s)}}{G}\right)\gamma(s,T) - \Phi^{-1}\left(\frac{S_0 e^{rT}}{G}\right)\gamma(0,T). \tag{L.14}$$

(N.9) can be decomposed as:

$$\xi(0,t) = \delta(0,s) + \int_s^t \left(1-e^{-\kappa(T-u)}\right) dW(u). \tag{L.15}$$

(L.11), (L.13), (L.14), and (L.15) thus collectively give:

$$S(t) = \frac{G}{e^{r(T-t)}}\Phi\left(\frac{\sqrt{T-s-\frac{2}{\kappa}\left(1-e^{-\kappa(T-s)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-s)}\right)}\Phi^{-1}\left(\frac{S_s e^{r(T-s)}}{G}\right)+\xi(s,t)}{\sqrt{T-t-\frac{2}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{1}{2\kappa}\left(1-e^{-2\kappa(T-t)}\right)}}\right), \tag{L.16}$$

where

$$\xi(s,t) = \int_s^t \left(1-e^{-\kappa(T-u)}\right) dW(u). \tag{L.17}$$

This proves (2.35).

Appendix M

Proof of Proposition 2.5

Under risk neutral valuation, we have:

$$C(t,K,\tau) = E\left[e^{-r(\tau-t)}(S(\tau)-K)^+ \mid \mathcal{F}_t\right]. \tag{M.1}$$

Following the condition in the proposition, given $\mathcal{F}_t$, we have

$$S(\tau) = Ge^{-r(T-\tau)}\Phi\left(cX(\tau)+\zeta\right), \tag{M.2}$$

and

$$X(\tau) \mid \mathcal{F}_t \sim N(\mu_x, \sigma_x), \tag{M.3}$$

where $c > 0, \zeta, \mu_x, \sigma_x$ are $\mathcal{F}_t$-measurable. Substituting this result into (M.1) and simplify, we get

$$C(t,K,\tau) = Ge^{-r(T-t)}E\left[\left(\Phi\left(cX(\tau)+\zeta\right)-K^*\right)^+\right]. \tag{M.4}$$

Since $c > 0$, the term $cX(t,\tau)+\zeta$ in the expectation is a monotonically increasing function of $X(t,\tau)$. We may thus write the expectation term as:

$$\begin{aligned} &E\left[\left(\Phi\left(cX(t,\tau)+\zeta\right)-K^*\right)^+\right] \\ &= \int_{x^*}^{\infty} \Phi\left(cx+\zeta\right) f_X(x)dx - K^* Pr\left(X(t,\tau) > x^*\right), \end{aligned} \tag{M.5}$$

where $f_X(x)$ is the conditional density of $X(t,\tau)$ given $\mathcal{F}_t$, or more specifically

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_x}\exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\}, \tag{M.6}$$

and

$$x^* = \frac{\left(\Phi^{-1}\left(K^*\right) - \zeta\right)}{c}, \tag{M.7}$$

which is the unique value of $X(t,\tau)$ satisfying:

$$\Phi\left(cX(t,\tau) + \zeta\right) = K^*.$$

The integral expression in (M.5) is then evaluated. Note that the normality of $X(t,\tau)$ ensures the existence of a closed-form option value expression.

First, (M.3) leads to the immediate result:

$$K^* Pr\left(X(t,\tau) > x^*\right) = K^* \Phi\left(\frac{\mu_x - x^*}{\sigma_x}\right). \tag{M.8}$$

Hence, the integral expression becomes:

$$\begin{aligned} L &= \int_{x^*}^{\infty} \Phi\left(cx + \zeta\right) f_X(x) dx \\ &= \int_{x^*}^{\infty} \left(\int_{-\infty}^{cx+\zeta} \phi(z) dz\right) f_X(x) dx, \end{aligned} \tag{M.9}$$

where $\phi(z)$ denotes the density of a standard normal variable at z, i.e.:

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}. \tag{M.10}$$

We then apply change of variable transformation and decomposition. First, define

$$u = \frac{\mu_x - x}{\sigma_x}.$$

This leads to:

$$L = \int_{-\infty}^{\frac{\mu_x - x^*}{\sigma_x}} \phi(u) \int_{-\infty}^{-cu\sigma_x + c\mu_x + \zeta} \phi(z) dz du. \tag{M.11}$$

Next, we apply the reverse of Cholesky decomposition to define:

$$v = \rho u + \sqrt{1 - \rho^2} z, \tag{M.12}$$

where

$$\rho = \frac{c\sigma_x}{\sqrt{1 + c^2\sigma_x^2}}. \tag{M.13}$$

Equivalently, this means:

$$z = \frac{v - \rho u}{\sqrt{1-\rho^2}}. \tag{M.14}$$

This leads to:

$$\begin{aligned} L &= \int_{-\infty}^{\frac{\mu_x - x^*}{\sigma_x}} \phi(u) \int_{-\infty}^{(-cu\sigma_x + c\mu_x + \zeta)\sqrt{1-\rho^2} + u\rho} \frac{1}{\sqrt{1-\rho^2}} \phi\left(\frac{v - \rho u}{\sqrt{1-\rho^2}}\right) dvdu \\ &= \int_{-\infty}^{\frac{\mu_x - x^*}{\sigma_x}} \int_{-\infty}^{\frac{c\mu_x + \zeta}{\sqrt{1+c^2\sigma_x^2}}} \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} exp\left(\frac{-(v^2 - 2uv\rho + u^2)}{2(1-\rho^2)}\right) dvdu \\ &= \Phi_2\left(\frac{\mu_x - x^*}{\sigma_x}, \frac{c\mu_x + \zeta}{\sqrt{1+c^2\sigma_x^2}}, \rho\right), \end{aligned} \tag{M.15}$$

where by definition $\Phi_2(x_1, x_2, \rho)$ denotes the bivariate standard Normal distribution function with correlation coefficient ρ evaluated at (x_1, x_2).

Substitute the results of (M.5), (M.8), and (M.15) into (M.4). This gives:

$$C(t, K, \tau) = Ge^{-r(T-t)} \left(\Phi_2\left(\frac{\mu_x - x^*}{\sigma_x}, \frac{c\mu_x + \zeta}{\sqrt{1+c^2\sigma_x^2}}, \rho\right) - K^*\Phi\left(\frac{\mu_x - x^*}{\sigma_x}\right)\right), \tag{M.16}$$

which is exactly as claimed in Proposition 2.5.

Appendix N

Proof of Corollary 2.2

The proof is very trivial. By a risk neutral valuation:

$$\begin{aligned} C_F(t,K,\tau,\tau') &= E\left[e^{-r(\tau-t)}(F(\tau,\tau')-K)^+ \mid \mathcal{F}_t\right] \\ &= e^{r(\tau'-\tau)}E\left[e^{-r(\tau-t)}(S(\tau)-K')^+ \mid \mathcal{F}_t\right], \end{aligned} \tag{N.1}$$

where

$$K' = Ke^{-r(\tau'-\tau)}.$$

Similarly, the relationship between the put options can be shown to hold.

Appendix O

Proof of (2.45)

We first find the distribution of $X(t,\tau)$. By (2.19)

$$X \equiv X(t,\tau) = (T-\tau)W(\tau) + \int_t^\tau W(u)du.$$

Conditioning on $\mathcal{F}_t$, we have:

$$W(\tau) \equiv N\left(W(t), \sqrt{(\tau-t)}\right). \tag{O.1}$$

In addition, (2.8) gives:

$$\int_t^\tau (W(u))du \equiv N\left(W(t)(\tau-t), \sqrt{\frac{(\tau-t)^3}{3}}\right). \tag{O.2}$$

Therefore, we know that

$$X \equiv N(\mu_x, \sigma_x), \tag{O.3}$$

where

$$\begin{aligned}\mu_x &= (T-\tau)W(t) + W(t)(\tau-t),\\ &= (T-t)W(t)\end{aligned} \tag{O.4}$$

$$\sigma_X^2 = (T-\tau)^2(\tau-t) + \frac{(\tau-t)^3}{3} + 2\,(T-\tau)\,Cov\left(W(\tau), \int_t^\tau W(u)du\right). \tag{O.5}$$

To find the covariance term, we use the Riemann-Stieltjes sum approximation. Partition the time interval (t, τ) into equal-length intervals given by: $t_0 = t, t_1 = t + \Delta t, t_2 = t + 2\Delta t, ..., t_i = t + i\Delta t, ...t_N = t + N\Delta t$, where $\Delta t = \frac{\tau - t}{N}$. This gives, for N large enough:

$$\begin{aligned} W(\tau) &\approx L_1(\Delta t) \\ &= W(t_0) + (W(t_1) - W(t_0)) + (W(t_2) - W(t_1)) ... + (W(t_i) - W(t_i - 1)) \\ &= W(t) + \Delta W(t_1) + \Delta W(t_2) + ... + \Delta W(t_{N-1}). \end{aligned} \tag{O.6}$$

$$\begin{aligned} \int_t^\tau W(u)du \approx L_2(\Delta t) &= \sum_{i=0}^{N-1} W(t_i)\Delta t \\ &= \sum_{i=0}^{N-1} \left[W(t_0) + (W(t_1) - W(t_0)) + (W(t_2) - W(t_1)) ... + (W(t_i) - W(t_i - 1))\right] \Delta t \\ &= \sum_{i=0}^{N-1} \left[W(t_0) + \sum_{j=1}^{t} \Delta W(t_j)\right] \Delta t \\ &= W(t)N\Delta t + \Delta W(t_1)(N-1)\Delta t + \Delta W(t_2)(N-2)\Delta t + ... + \Delta W(t_{N-1})\Delta t \\ &= (\tau - t)\left[W(t) + \Delta W(t_1)\frac{N-1}{N} + \Delta W(t_2)\frac{N-2}{N} + ... + \Delta W(t_{N-1})\frac{1}{N}\right]. \end{aligned} \tag{O.7}$$

We know that $\Delta W(t_i)$ are independent and identically distributed variables, where

$$\Delta W(t_i) \equiv N(0, \Delta t). \tag{O.8}$$

Therefore

$$\begin{aligned} Cov\left(L_1(\Delta t), L_2(\Delta t)\right) &= (\tau - t)\Delta t \sum_{i=1}^{N-1} \frac{N-i}{N} \\ &= \frac{N(N-1)(\tau - t)^2}{2N^2}. \end{aligned} \tag{O.9}$$

This leads to the result:

$$\begin{aligned} Cov\left(W(\tau), \int_t^\tau W(u)du\right) &\approxeq \lim_{N\to\infty} Cov\left(L_1(\Delta t), L_2(\Delta t)\right) \\ &= \lim_{N\to\infty} \frac{N(N-1)(\tau - t)^2}{2N^2} \\ &= \frac{(\tau - t)^2}{2}. \end{aligned} \tag{O.10}$$

Substituting this result into (G.5) gives:

$$\sigma_X^2 = (T-\tau)^2(\tau-t) + (\tau-t)^2(T-\tau) + \frac{(\tau-t)^3}{3}. \tag{O.11}$$

Hence (2.45) to (2.47) are proved.

Appendix P

Proof of Proposition 3.3

The proof for Proposition 3.3 is straightforward. Without loss of generality, it is sufficient to demonstrate the proof for call options. Under risk neutral valuation:

$$C^{MLN}(t,K,\tau) = E\left[e^{-r(\tau-t)}\left(S(\tau)-K\right)^{+} \mid \mathcal{F}_t\right]. \tag{P.1}$$

Based on the model specification in (3.26) and (3.29):

$$S(t) = \sum_{i=1}^{N} 1_{C=i} S_i(t), \quad i \in 1,2...N, \tag{P.2}$$

where

$$dS_i(t) = S_i(t)rdt + \sigma_i S_i(t)dZ(t). \tag{P.3}$$

Therefore, we have

$$\begin{aligned} C^{MLN}(t,K,\tau) &= E\left[E\left[e^{-r(\tau-t)}\left(S(\tau)-K\right)^{+} \mid \mathcal{F}_t, C\right]\right] && \text{(P.4)} \\ &= \sum_{i=1}^{N} \lambda_i E\left[e^{-r(\tau-t)}\left(S_i(\tau)-K\right)^{+} \mid \mathcal{F}_t\right] \\ &= \sum_{i=1}^{N} \lambda_i C^{LN}(t,K,\tau,\sigma_i). && \text{(P.5)} \end{aligned}$$

where the last line follows from that each component price process being a Geometric Brownian Motion in (P.3), coinciding with the specification of the Lognormal allowance price model in (3.1).

Appendix Q

Proof of (3.41)

From the specification in (3.36), we have:

$$E\left[\ln\left(\frac{S(\tau)}{S_t}\right) \mid C=1\right] = r - \frac{1}{2}\sigma_1^2, \tag{Q.1}$$

$$Var\left(\ln\left(\frac{S(\tau)}{S_t}\right) \mid C=1\right) = \sigma_1^2, \tag{Q.2}$$

$$E\left[\ln\left(\frac{S(\tau)}{S_t}\right) \mid C=2\right] = r - \frac{1}{2}\sigma_2^2, \tag{Q.3}$$

$$Var\left(\ln\left(\frac{S(\tau)}{S_t}\right) \mid C=2\right) = \sigma_2^2, \tag{Q.4}$$

Therefore, total return variance is given by:

$$\begin{aligned} Var\left(\ln\left(\frac{S(\tau)}{S_t}\right)\right) &= E\left[Var\left(\ln\left(\frac{S(\tau)}{S_t}\right) \mid C\right)\right] + VaR\left(E\left[\ln\left(\frac{S(\tau)}{S_t}\right) \mid C\right]\right) \\ &= \lambda\sigma_1^2 + (1-\lambda)\sigma_2^2 + \lambda(1-\lambda)\left(\frac{1}{2}\sigma_1^2 - \frac{1}{2}\sigma_2^2\right)^2. \end{aligned} \tag{Q.5}$$

Substituting (Q.1) to (Q.5) into (3.40) gives the result in (3.41).

Appendix R

Option Greeks under LN and MLN allowance price models

Under the Lognormal allowance price model, the option Greek expressions are direct replications of the Black-Scholes results. We use the subscripts c and p to differentiate between call and put options.

- Delta:

$$\Delta_c^{LN} = \frac{\partial C^{LN}(t,K,T)}{\partial S} = N(d_1), \tag{R.1}$$

$$\Delta_p^{LN} = \frac{\partial P^{LN}(t,K,T)}{\partial S} = -N(-d_1). \tag{R.2}$$

- Gamma:

$$\Gamma_c^{LN} = \Gamma_p^{LN} = \frac{\partial^2 C^{LN}(t,K,T)}{\partial S^2} = \frac{N'(d_1)}{S_t\sigma\sqrt{T-t}}. \tag{R.3}$$

- Vega:

$$\mathcal{V}_c^{LN} = \mathcal{V}_p^{LN} = \frac{\partial C^{LN}(t,K,T)}{\partial \sigma} = S_t N'(d_1)\sqrt{T-t}. \tag{R.4}$$

- Rho

$$\rho_c^{LN} = \frac{\partial C^{LN}(t,K,T)}{\partial \sigma} = e^{-r(T-t)}K(T-t)N(d_2), \tag{R.5}$$

$$\rho_p^{LN} = \frac{\partial P^{LN}(t,K,T)}{\partial \sigma} = -e^{-r(T-t)}K(T-t)N(-d_2). \tag{R.6}$$

d_1 and d_2 follow the expressions in Proposition 3.1.

Under the Mixture Lognormal allowance price model, each option Greek is the λ-weighted average of the LN Greeks calculated with the component parameters. For example, in the $MLN-2$ model, the put option delta is given by:

$$\Delta_P^{MLN2} = \lambda \frac{\partial P^{LN}(t, K, T, \sigma_1)}{\partial S} + (1-\lambda) \frac{\partial P^{LN}(t, K, T, \sigma_2)}{\partial S}. \tag{R.7}$$

Special attention should be given to vega, which must be defined in terms of the component volatilities under the Mixture Lognormal model.