Three Essays on Risk in Financial Markets

D I S S E R T A T I O N
of the University of St. Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Management

submitted by

Tobias Nigbur

from
Germany

Approved on the application of

Prof. Dr. Manuel Ammann

and

Prof. Dr. Markus Schmid

Dissertation no. 4295

Digitaldruckhaus GmbH, Konstanz 2014
Three Essays on Risk in Financial Markets

D I S S E R T A T I O N
of the University of St. Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Management

submitted by

Tobias Nigbur

from
Germany

Approved on the application of

Prof. Dr. Manuel Ammann

and

Prof. Dr. Markus Schmid

Dissertation no. 4295

Digitaldruckhaus GmbH, Konstanz 2014
The University of St. Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, May 19, 2014

The President:

Prof. Dr. Thomas Bieger
Acknowledgements

This dissertation presents the research that evolved through the four years I have been a doctoral student at the Swiss Institute of Banking and Finance (SBF). The accomplishment of such a challenging and demanding work would not have been possible without the help of a number of people.

First and foremost, my gratitude goes to my thesis supervisor, Prof. Dr. Manuel Ammann, for his guidance, his inspiration and his encouragements throughout the years. I greatly benefited from our many discussions and his valuable comments and insights helped me improve my work considerably. I would like to thank my co-supervisor Prof. Dr. Markus Schmid as well as my co-authors Christian and Norman, and my other colleagues at the SBF for many interesting and stimulating conversations, for helpful insights and for fruitful comments on my research, which allowed me to learn something new every day. Further, I thank Professor Adlai Fisher for inviting me to University of British Columbia, Canada, encouraging me in my work and involving me in the faculty. This research stay would not have been possible had it not been financially supported by the Swiss National Science Foundation.

Above all, I am truly indebted to my parents for their manifold and unquestioning support and their loving care which has been the foundation of all my achievements. The emotional support of my family, Angelika, Siegfried, Simon and Leonie, and my friends provided me the strength to finalize this dissertation.

St. Gallen, June 2014, Tobias Nigbur
# Contents

## General Introduction

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I Risk Management</td>
<td>13</td>
</tr>
<tr>
<td>I.1 Introduction</td>
<td>14</td>
</tr>
<tr>
<td>I.2 Methodological Framework</td>
<td>16</td>
</tr>
<tr>
<td>I.3 Concept of Displaced Relative Changes</td>
<td>17</td>
</tr>
<tr>
<td>I.3.1 The Displaced Diffusion Model for Stochastic Processes</td>
<td>18</td>
</tr>
<tr>
<td>I.3.1.1 Displace Diffusion Model, Historical Simulation, and Relation to Relative Changes</td>
<td>19</td>
</tr>
<tr>
<td>I.3.2 Displaced Relative Changes with Constant Displacements</td>
<td>20</td>
</tr>
<tr>
<td>I.3.2.1 Displaced Relative Shifting with Time Dependent Displacements</td>
<td>21</td>
</tr>
<tr>
<td>I.4 Displaced Relative Changes with GARCH Filtering</td>
<td>22</td>
</tr>
<tr>
<td>I.5 Estimation and Testing</td>
<td>24</td>
</tr>
<tr>
<td>I.5.1 Maximum Likelihood Estimation for the Displaced Lognormal Model</td>
<td>24</td>
</tr>
<tr>
<td>I.5.2 Backtesting the Historical Simulation</td>
<td>25</td>
</tr>
<tr>
<td>I.5.2.1 Average Backtest Hit Statistic</td>
<td>26</td>
</tr>
<tr>
<td>I.5.2.2 Back-Test Stability for Different Calibration Window Sizes</td>
<td>26</td>
</tr>
<tr>
<td>I.6 Numerical Results</td>
<td>26</td>
</tr>
<tr>
<td>I.6.1 Data</td>
<td>26</td>
</tr>
<tr>
<td>I.6.2 Average Back-Test Hits Statistic Results</td>
<td>27</td>
</tr>
<tr>
<td>I.6.3 Back-Test Stability for Different Calibration Window Sizes</td>
<td>29</td>
</tr>
<tr>
<td>I.7 Conclusion</td>
<td>33</td>
</tr>
<tr>
<td>I.8 References</td>
<td>34</td>
</tr>
<tr>
<td>I.9 Maximum Likelihood Estimation for Itô Processes</td>
<td>36</td>
</tr>
</tbody>
</table>

## II Empirical Corporate Finance

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>II.1 Introduction</td>
<td>39</td>
</tr>
<tr>
<td>II.2 Hypothesis Development</td>
<td>41</td>
</tr>
<tr>
<td>II.3 Data</td>
<td>47</td>
</tr>
<tr>
<td>II.3.1 Sources and Screening Criteria</td>
<td>47</td>
</tr>
<tr>
<td>II.3.2 Sampling and Sample Size</td>
<td>47</td>
</tr>
<tr>
<td>II.3.3 Descriptive Statistics</td>
<td>49</td>
</tr>
<tr>
<td>II.4 Event Study</td>
<td>52</td>
</tr>
<tr>
<td>II.4.1 Trading Volume</td>
<td>52</td>
</tr>
<tr>
<td>II.4.2 Stock Returns</td>
<td>54</td>
</tr>
<tr>
<td>II.4.3 Explanatory Regressions</td>
<td>58</td>
</tr>
</tbody>
</table>
General Introduction

This thesis consists of three individual research topics. It comprises a new approach for risk management, empirical research in corporate finance, and theoretical research on firm structure.

In Chapter I presents empirical and theoretical work in risk management: *Displaced Historical Simulation Improves Modeling of Possibly Negative-Valued Financial Risk Values: Application to VaR in Times of Negative Government Bond Yields*. In this paper we introduce the displaced relative change (DRC) model and the displaced filtered historical simulation (DFHS) model. These models make possible historical simulations based on potentially negative risk factors, such as interest rates or spreads. This is an issue of recent and major interest to the financial sector, both from a regulatory and financial institutions perspective, especially in light of observed negative values for government bond yields. Our empirical results show that compared to other models, models with our proposed displacement feature handle situations of close-to-zero or negative interest rates particularly well.

Chapter II deals with empirical corporate finance: *Calls of Convertible Debt Securities: No Bad News at All*. In this paper, I examine the impact of in-the-money convertible bond calls on stock prices, employing a sample of U.S. convertible bond calls over the period 1994-2011. In contrast to previous literature, I find that conversion forcing convertible bond calls do not significantly influence stock prices. I posit that the discrepancy between my results and those in the literature is caused by amplified screening criteria, especially strong news cleaning. Companies tend to announce calls as side notes to other major corporate news, resulting in an event study bias. Further, convertible bond design, moneyness of the conversion option at the announcement date, and convertible-arbitrage strategies cast doubt on previous literature’s negative abnormal returns.

The last paper *Failure Risk, Lender Control, and Contracting in Project Finance* in Chapter III theoretically answers the question why project firms contract away asset risks. I argue that lenders have a strong influence on the shape of the firm because the financing contribution of the sponsoring firm is low and debt ownership is concentrated among few parties. I propose a simple, but powerful extension of standard real option models to failure risk that can explain the low residual risks in project finance. Failure risk can lead to a total loss of a company’s assets-in-place, a situation that most of contingent claim models ignore. The concept of failure risk is regularly observable in practice, but is especially pronounced in project finance. Abandoned projects, catastrophes, and political instability are just a few examples of how the claims of lenders and equity holders can be rendered worthless. Calibrated to reasonable parameters, this model shows that it can be beneficial for equity holders to contract away asset volatility.
Einleitung

Diese Dissertation besteht aus drei individuellen Forschungsthemen. Sie beinhaltet einen neuen Ansatz für das Risikomanagement, empirische Forschung in Unternehmensfinanzierung und theoretische Forschung im Bereich Firmenstruktur.


Displaced Historical Simulation Improves Modeling of Possibly Negative-Valued Financial Risk Values: Application to VaR in Times of Negative Government Bond Yields

Joint work with Christian P. Fries and Norman J. Seeger.

Abstract
In this paper we introduce the displaced relative change (DRC) model and the displaced filtered historical simulation (DFHS) model. These models make possible historical simulations based on potentially negative risk factors, such as interest rates or spreads. This is an issue of recent and major interest to the financial sector, both from a regulatory and financial institutions perspective, especially in light of observed negative values for government bond yields. Our empirical results show that compared to other models in the literature, models with our proposed displacement feature handle situations of close-to-zero or negative interest rates particularly well.
I.1 Introduction

In this paper we propose an extended historical simulation approach that can handle situations in which a risk variable, such as interest rates, spreads, or security prices, can take on negative values. This is a situation that recently became of major interest when negative sovereign bond yields or negative interest rate spreads have been observed frequently.

In general, the historical simulation approach is used to construct a distribution of possible future risk variable realizations from the time series of past risk variable realizations. In the most standard form of historical simulation, first, a method for calculating changes of past realizations is chosen, for example, absolute or relative changes.\(^1\) Next, independence of the past changes in the risk variable is assumed and the past changes are applied to the current data level to create a distribution of possible future values.

Historical simulations are widely employed in the financial industry to assess market risk. For example, under the Basel Accord, banks use Value-at-Risk (VaR) based on historical simulation to assess their market risk. Similarly, insurance companies and mutual funds frequently measure the market risk of their portfolio positions based on historical simulations. An advantage of historical simulation is that it is relatively easy to implement and has the potential to preserve structures hidden in the time-series data, e.g., correlation structures, that can become lost when fitting the data to a fully parametric model.\(^2\) However, we show in an empirical application of historical simulation to Swiss government bond yield data and U.S. government bond spread data that standard historical simulation and currently available extensions have a difficult time coping with risk variables that can take on close-to-zero or even negative values. This is particularly the case for situations when relative changes, e.g., log returns, are assumed as the starting point of a historical simulation. This is important to notice since relative changes in the form of log returns constitute a standard choice in financial industry which is often applied uncritically when modeling changes to conduct historical simulations. The analysis in this paper is centered around extending historical simulation in such a way that it can handle close-to-zero or negative risk variables.

The contributions of this paper are the following. First, we propose a simple one-parameter “historical simulation model” that we call the displaced relative change model (DRC). The displaced relative change model encompasses the two most prominent change concepts—relative and absolute changes—as special cases and can be viewed as a continuous in-

---

1 For a more detailed discussion on choosing a method to calculate change when performing historical simulations, see Fries et al. (2013).

2 Another approach to generating a distribution of possible future realizations of a risk variable is to use a parametric model dependent approach. That is, one assume some parametric data-generating model, estimates the model parameters based on a time-series of past realizations, and then generates a distribution of possible future values using the fitted model, e.g., analytically or through a Monte Carlo simulation. In such an approach, the resulting distribution is by necessity strongly influenced by the choice of the parametric model imposed on the structures in the data.
terpolation between them. In fact, employing the DRC model makes it unnecessary to explicitly choose between using either absolute or relative changes since the DRC model interpolates between the two concepts depending on the needs of the data. This feature of the model makes it particularly interesting for financial institutions from a regulatory perspective. Changing, e.g., from absolute to relative shifting when conducting a historical simulation poses an operational model change which, in general, has to be approved by the regulator. An advantage of the DRC model is that it makes an automatic and, even more important from the viewpoint of regulatory supervision, methodologically sound decision as to whether to apply absolute or relative changes to the data.

Second, the DRC model is designed to handle negative risk factors, such as negative interest rates or negative spreads, a property other historical simulation models lack. This is of particular interest since we currently face a market regime in which negative interest rates, e.g., for sovereign bonds, and negative spreads are frequently observed. Furthermore, we extend the filtered historical simulation (FHS) model introduced by Barone-Adesi et al. (1999) and Hull and White (1998) to overcome its inability to handle negative state variables. To this end, we combine our displaced relative change approach with FHS and propose the displaced filters historical simulation (DFHS), a model fully capable of handling negative state variables.

Finally, we conduct an empirical test of our proposed models. We apply models equipped with the displaced change feature to a VaR application. We set up a back-testing framework for VaR calculations using Swiss government bond yield (2Y) data for the period 1999–2012 and U.S. government bond spread data (10Y-3M) for the period 1992–2012. We compare the forecasting performance of the DFHS model and a combination of a standard GARCH model combined with our DRC model with that of a standard historical simulation model, a FHS, and a GARCH model. Note that our proposed models are not restricted to VaR applications but can, in fact, be applied to any other situation where distributions of state variables need to be considered.

There is a vast literature on applying historical simulation to financial issues. One strand that is particularly related to our paper attempts to overcome the major assumption of historical simulation that historical observations are identically independently distributed. This assumption is not supported by empirical data since we commonly observe the phenomenon of clustered volatility (see Bollerslev (1986)), which cannot be picked up by historical simulation. To overcome this drawback Boudoukh et al. (1998) propose an exponential weighting of historical time series that puts more weight on more recent observations. Several other papers, e.g., see Hull and White (1998) and Barone-Adesi et al. (1999), suggest a filtered historical simulation approach, in which historical observations are passed through a filter before being used to generate forecasting distributions in order

3For a general discussion of the drawbacks of historical simulation, see, e.g., Christoffersen (2003) or Dowd (2002).
to account for time-series pattern in the variance of observations. The general idea is that GARCH-type models are fitted to historical data and estimated volatility is used to adjust the observed data. For example, Hull and White (1998) use estimated volatility to update the current level of volatility when calculating a forecasting distribution and thereby are able to improve forecasts on stock indices. Barone-Adesi et al. (1999) use the errors resulting from fitting historical data to a GARCH process as shifts when generating a forecasting distribution. Multivariate extensions for filtered historical simulation models are proposed in Christoffersen (2009) and Audrino and Barone-Adesi (2005). Our suggested displaced filtered historical simulation model is in the spirit of the FHS models and is another extension of this model class.

Indeed, a particularly prominent application of historical simulation found in the literature has to do with VaR and management of market risk, which is why we chose VaR as an application in our empirical study also. There are a great many empirical and theoretical studies related to VaR that use historical simulation as a benchmark case. See, e.g., Küster et al. (2006), Bao et al. (2006), Barone-Adesi et al. (2002), and van den Goorbergh and Vlaar (1999).

The remainder of the paper is structured as follows. Section I.2 introduces the notational and methodological framework of historical simulation. Section I.3 presents our proposed displaced relative change model that, in Section I.4, is combined with a filtered historical simulation approach to create a new displaced filtered historical simulation model. Section I.5 lays out the estimation and back-testing procedure for testing the models empirically. Data and empirical results are presented and discussed in Section I.6 and Section I.7 summarizes and concludes.

I.2 Methodological Framework

In this section we introduce our notational framework and offer a formal definition of a historical simulation. Let $t_i$, with $t_i < 0$ and $i = 1, \ldots, n + 1$, denote a series of past discrete observation times and $x_t$ denotes a time series of $n + 1$ past observations of some state variable (risk variable) at the respective times $t_i$. \{${x_t \mid i = 1, \ldots, n + 1}$\} can be thought of as a particular sample path, denoted with $\omega_0$, of some (unknown) stochastic process $X$, i.e.,

$$x_{t_i} = X(t_i, \omega_0) \quad \text{for all } i = 1, \ldots, n + 1.$$ 

A historical simulation consists of generating $n$ possible future scenario samples $x^i_{T_1} := X(T_1, \omega_i)$, with $i = 1, \ldots, n$, as variations of a given state $x^0 = X(T_0, \omega_0)$ by applying data changes derived from $n$ “historical” changes in a time series \{${x_t}$\} to the current state $x^0$. That is, $x^i_{T_1}$ and $x_{t_i}$ are understood to be samples of the same stochastic process $X$, namely,
samples with respect to time (for a fixed path $\omega_0$)

$$x_{T_1}^i = X(T_1, \omega_i) \quad \text{(I.1)}$$

and samples with respect to the state space (at a fixed time $T_1$)

$$x_{t_i} := X(t_i, \omega_0), \quad \text{(I.2)}$$

where $i = 1, 2, \ldots, t_i \leq T_0 < T_1$, and $\omega_i \in F_0 \in \mathcal{F}_{T_0}$ for a fixed $F_0$ (and $X$ being adapted to the filtration $\mathcal{F}$).\(^4\) Given that the process $X$ is known, the two equations allow a time series from the past ($x_{t_i}$ defined by varying the point in time, i.e., varying $t_i$), to be used to generate a distribution in the future ($x_{T_1}^i$ defined by varying the path, i.e., varying $\omega_i$).

Note that we still left open how we derive the past changes from $\{x_{t_i}\}$ and how we apply them to the current state $x^0$ to generate possible future scenarios $x_{T_1}^i$. Two prominent options for generating a possible future distribution $X(T_1)$ are absolute and relative changes.

Absolute data changes are defined as follows: given a discrete time series $\{x_{t_i}\}$ and a state $x^0$, we define a sample of states $x_{T_1}^i$ by adjusting the state $x^0$ according to the “model” of absolute changes as

$$x_{{T_1}}^{\text{abs}, i} := x^0 + (x_{t_i} - x_{t_{i-1}}) \quad \text{(I.3)}$$

for $i = 1, \ldots, n$. That is, we generate a possible future distribution of $X$ based on absolute changes of the process in the past.\(^5\)

A second prominent option is to consider relative changes calculated according to

$$x_{{T_1}}^{\text{rel}, i} := x^0 \left( 1 + \frac{x_{t_i} - x_{t_{i-1}}}{x_{t_{i-1}}} \right), \quad \text{(I.4)}$$

for $i = 1, \ldots, n$.

In practice, the choice of methodology to calculate and apply changes is often arbitrary. Frequently, relative changes are preferred and it corresponds to modeling rate of changes that are seen as more fundamental than absolute values. For example, for equity prices, one usually considers returns, i.e., relative changes of these prices. However, for spreads, i.e., quantities that are by definition differences (e.g., differences of interest rates), it is eminent that a relative change does not make sense. Spreads may become negative, in which case absolute changes appear more appropriate. We will refer to this issue in more detail in Section I.3.2.1.

\(^4\)Note that this cumbersome mathematical notation just means that we know the past, i.e., $X(t_i, \omega) = X(t_i, \omega_0)$ for all $\omega \in F_0$.

\(^5\)Note that the usual situation is such that $\{x_{t_i}\}$ is a sequence of the predecessors of $x^0$. However, $t_i$ could also denote any other time period in the past (e.g., a period with stressed market data). We discuss this special case in more detail in Section I.3.2.1.
I.3 Concept of Displaced Relative Changes

In the following section we introduce the concept of displaced relative changes. We show that this concept arises naturally when applying historical simulation to the \textit{“displaced diffusion model”} (DD).

I.3.1 The Displaced Diffusion Model for Stochastic Processes

In modeling asset price processes, displacing a log-normal diffusion process is a common way of introducing a mixed log-normal/normal model; see Brigo and Mercurio (2001). Rubinstein (1983) introduces the displaced diffusion model into the area of option pricing and recently the model has received extensive use in the LIBOR market literature; see, e.g., Beveridge and Joshi (2009), Errais and Mercurio (2005), Brigo and Mercurio (2003). The model captures a skewed distribution for the underlying stochastic process, which leads to the desirable feature of skewed implied volatility patterns. The model is also appreciated for its analytical tractability, resulting in closed-form pricing solutions. The model may allow for negative values. This was considered an undesired effect in interest rate modeling.\footnote{With respect to interest rate modeling, Rebonato (2002) shows that for reasonable values of the displacement parameter, the probability of obtaining negative values is small. Joshi and Rebonato (2003) show that the displaced diffusion model has characteristics that are very similar to those of the constant elasticity of variance (CEV) model, which is known for its appealing economic properties.} However, in the post-credit crises, modeling with negative rates became a desirable feature. For us, too, the ability to model negative risk factors is an advantage and is why we consider the model. Assume $X$ follows the mixed log-normal/normal process

$$dX(t) = ((1 - |\alpha|)X(t) + \alpha X_0)\sigma^*dW(t), \quad (I.5)$$

with $-1 \leq \alpha \leq 1$. Then $X$ is a log-normal process for $\alpha = 0$ having log-volatility $\sigma^*$, and $X$ is a normal process for $\alpha = 1$ and $\alpha = -1$ having normal volatility $X_0\sigma^*$. For $0 < |\alpha| < 1$, the process is a mixture of a log-normal and a normal process. For $\alpha < 0$, the two processes are anti-correlated. Note that the standard procedure in the literature is to consider the case $0 \leq \alpha \leq 1$, but it is straight-forward to generalize the model. Here $X_0 > 0$ is just a scaling factor to allow different volatilities in the two limiting cases. Without loss of generality, we may assume $X_0 > 1$.

For $|\alpha| \neq 1$, we can rewrite Equation (I.5) in order to obtain the standard definition of the displaced diffusion model:

$$dX(t) = \left(X(t) + \frac{\alpha}{(1 - |\alpha|)}X_0\right)(1 - |\alpha|)\sigma^*dW(t)$$

$$= (X(t) + a)\sigma^*dW(t), \quad (I.6)$$

with $-1 \leq \alpha \leq 1$. Then $X$ is a log-normal process for $\alpha = 0$ having log-volatility $\sigma^*$, and $X$ is a normal process for $\alpha = 1$ and $\alpha = -1$ having normal volatility $X_0\sigma^*$. For $0 < |\alpha| < 1$, the process is a mixture of a log-normal and a normal process. For $\alpha < 0$, the two processes are anti-correlated. Note that the standard procedure in the literature is to consider the case $0 \leq \alpha \leq 1$, but it is straight-forward to generalize the model. Here $X_0 > 0$ is just a scaling factor to allow different volatilities in the two limiting cases. Without loss of generality, we may assume $X_0 > 1$.

For $|\alpha| \neq 1$, we can rewrite Equation (I.5) in order to obtain the standard definition of the displaced diffusion model:

$$dX(t) = \left(X(t) + \frac{\alpha}{(1 - |\alpha|)}X_0\right)(1 - |\alpha|)\sigma^*dW(t)$$

$$= (X(t) + a)\sigma^*dW(t), \quad (I.6)$$

with $-1 \leq \alpha \leq 1$. Then $X$ is a log-normal process for $\alpha = 0$ having log-volatility $\sigma^*$, and $X$ is a normal process for $\alpha = 1$ and $\alpha = -1$ having normal volatility $X_0\sigma^*$. For $0 < |\alpha| < 1$, the process is a mixture of a log-normal and a normal process. For $\alpha < 0$, the two processes are anti-correlated. Note that the standard procedure in the literature is to consider the case $0 \leq \alpha \leq 1$, but it is straight-forward to generalize the model. Here $X_0 > 0$ is just a scaling factor to allow different volatilities in the two limiting cases. Without loss of generality, we may assume $X_0 > 1$.
where the coordinate transformation is given by
\[
a = \frac{\alpha}{(1 - |\alpha|)}X_0 \quad \sigma^{(a)} = (1 - |\alpha|)\sigma^*
\]
\[
\alpha = \frac{\frac{\alpha}{X_0}}{(1 + \frac{\alpha}{X_0})} \quad \sigma^* = \left(1 + \frac{\alpha}{X_0}\right)\sigma^{(a)}.
\]
with \(-\infty < a < \infty\) corresponding to \(-1 < \alpha < 1\) and (I.6) being equivalent to
\[
d(X(t) + a) = (X(t) + a)\sigma^{(a)} dW(t).
\] (I.7)

We consider the model in the form
\[
d(X(t) + a) = \frac{(X(t) + a)}{(1 + \frac{\alpha}{X_0})}\sigma^* dW(t).
\] (I.8)

The latter form has the advantage of allowing us to consider a log-normal model for \(X(t) + a\) with a fixed (bounded) volatility parameter \(\sigma^*\) (note that \(\sigma^{(a)}\) is unbounded for fixed \(\sigma^*\)). This is advantageous when applying numerical algorithms, e.g., when estimating \(\sigma^*\) through a GARCH model.

I.3.1.1 Displace Diffusion Model, Historical Simulation, and Relation to Relative Changes

Let us assume that we observe past values of \(X(t_i, \omega_0)\) with \(i = 1, \ldots, n\) and we wish to generate “realistic” future samples of \(X(T_1)\) from it by means of historical simulation. Further, let us assume that \(X\) follows the stochastic process in Equation (I.8) for some displacement parameter \(a\). Using the concept of historical simulation, what we ultimately wish to do is to extract the stochastic driver that generates the historical data to forecast the possible future distribution of \(X\). In the case of the displaced diffusion model, we rearrange Equation (I.8) to back out the stochastic driver \(\sigma^{(a)} dW\) of the displaced diffusion model. That is, we calculate the stochastic driver as
\[
\sigma^{(a)} dW(t) = \frac{d(X(t) + a)}{X(t) + a}.
\]

In other words, to extract the stochastic driver of the displaced diffusion model from historical data, the displaced diffusion model implies that we have to calculate relative changes of the displaced historical values of the process \(X\). On the other hand, the SDE (I.8) leads to
\[
X(t + dt) + a = (X(t) + a) + (X(t) + a) \left[\sigma^{(a)} dW(t)\right]
\]
\[
= (X(t) + a) \left\{1 + \left[\frac{d(X(t) + a)}{X(t) + a}\right]\right\}.
\]
That is, using a historical driver to calculate a possible value of the displaced diffusion process at time $t + dt$ is consistent with generating a forecasting distribution by means of relative changes of the displaced value of $X$.

To summarize: the notion of “change”, consistent with the model approach (I.8) (equivalently (I.5)), is that of a “relative change of the displaced value with displacement parameter $a$”. For our time-discrete process defined in Equations (I.1) and (I.2), using the relative displaced changes approach translates to

\[
\begin{align*}
\Delta t_i & := \frac{(x_{t_i} + a) - (x_{t_{i-1}} + a)}{(x_{t_{i-1}} + a)} \\
\end{align*}
\]

and

\[
\begin{align*}
x_t^i + a & := (x^0 + a) + (x^0 + a)(\Delta t_i) \\
& = (x^0 + a)(1 + \Delta t_i).
\end{align*}
\]

**I.3.2 Displaced Relative Changes with Constant Displacements**

Let us apply changes in the variable $V^i := x^i_{T_1} + a$ using relative changes of past scenarios $W^i := x_i + a$, i.e., we define

\[
\begin{align*}
V^i_{T_1} & := V_0\left(1 + \frac{W_i - W_{i-1}}{W_{i-1}}\right).
\end{align*}
\]

In other words, we apply relative changes on a displaced variable, $X + a$ with displacement parameter $a$. By definition, this is equivalent to

\[
\begin{align*}
x^i_{T_1} + a & := (x^0 + a)\left(1 + \frac{x_{t_i} - x_{t_{i-1}}}{x_{t_{i-1}} + a}\right).
\end{align*}
\]
From this we have

\[
X^T_{t_1} := (x^0 + a) \left( 1 + \frac{x_t - x_{t-1}}{x_{t-1} + a} \right) - a
\]

\[
= x^0 + (x^0 + a) \frac{x_t - x_{t-1}}{x_{t-1} + a}
\]

\[
= x^0 \left( 1 + \frac{x_t - x_{t-1}}{x_{t-1} + a} \right) + a \frac{x_t - x_{t-1}}{x_{t-1} + a}
\]

\[
= x^0 \left( 1 + \frac{x_t - x_{t-1}}{x_{t-1} + a} \right) + a \frac{x_t - x_{t-1}}{x_{t-1} + a} \frac{1}{x_{t-1} + a} (x_t - x_{t-1})
\]

\[
= x^0 \left( 1 + \frac{x_t - x_{t-1}}{x_{t-1} + a} \right) x_{t-1} + a + x^0 \frac{a}{x_{t-1} + a} + a \frac{a}{x_{t-1} + a} (x_t - x_{t-1})
\]

\[
= x^0 \left( 1 + \frac{x_t - x_{t-1}}{x_{t-1} + a} \right) x_{t-1} + a + \left( x^0 + (x_t - x_{t-1}) \right) \frac{a}{x_{t-1} + a}
\]

\[
= \alpha \left[ x^0 \left( 1 + \frac{x_t - x_{t-1}}{x_{t-1} + a} \right) \right] + (1 - \alpha) \left[ x^0 + (x_t - x_{t-1}) \right]
\]

with \( \alpha = x_{t-1} / (x_{t-1} + a) \).

We can see that the displaced relative changes model is just a linear interpolation of relative and absolute changes applied to the underlying stochastic process without displacement. The two limit cases can be recovered. For \( a = 0 \), we obtain relative changes; for \( a = \infty \), we obtain absolute changes.

Two features of our model are worth special note at this point. First, we can use the DRC model to discover whether it is more appropriate to apply relative or absolute changes to a data series by the following approach. Apply the DRC model to the considered data series and estimate the parameter \( a \) (with respect to some optimality criteria, like a maximum likelihood). Then, for \( a = 0 \), relative changes should be used, and for \( a = \infty \), absolute changes should be used. Such a method provides an automatic and, even more important, methodologically sound (in the sense of a regulatory supervision) basis for deciding whether to apply absolute or relative changes to the data. Second, our DRC model has the ability to handle negative risk factors, an attribute that has become of major importance recently. Note that in general it is not sensible to apply the model of relative changes to negative risk factors. The DRC model handles this situation in the following intuitive way. In an empirical application, the displacement parameter is estimated in such a way that it shifts the risk factor into the positive realm of \( \mathbb{R} \); see Section I.5. Thereby, it is again possible to apply the relative changes to the risk factor and a historical simulation can be conducted.
I.3.2.1 Displaced Relative Shifting with Time Dependent Displacements

We can generalize the result from Section I.3.2. Let us consider that the current regime of \( x^0 \) follows a model with displacement \( a \), but the past regime of \( x_{t_i} \), from which we calculate the data changes, follows a model with displacement \( b \).

This situation can arise when there is a larger time lag between the observation time of \( x^0 \) and \( x_{t_i} \). For example, this is the case if we want to calculate a Stress-VaR, taking the samples from a past stress period.

Let us consider the general case of a local, time-dependent displacement parameter where we calculate the \( a(i) \)-displaced relative change of \( x_{t_i} \) and apply it as an \( a(0) \)-displaced relative shift to \( x^0 \) to generate \( X_{t_i} \). In other words, we apply relative changes to the variable \( V_i = x^0_t + a(0) \) using relative changes of scenarios \( W_{t_i} := x_{t_i} + a(i) \), i.e., we define

\[
V_{t_i} := V_0 \left( 1 + \frac{W_{t_i} - W_{t_i-1}}{W_{t_i-1}} \right).
\]

From

\[
X_{t_i}^i + a(0) := (x^0 + a(0)) \left( 1 + \frac{x_{t_i} - x_{t_i-1}}{x_{t_i-1} + a(i)} \right)
\]

we then, analogously to the derivation in Section I.3.2, obtain

\[
X_{t_i}^i := x^0 \left( 1 + \frac{x_{t_i} - x_{t_i-1}}{x_{t_i-1}} \right) \frac{x_{t_i-1}}{x_{t_i-1} + a(i)} + x^0 \frac{a(i)}{x_{t_i-1} + a(i)}
\]

\[
+ \frac{a(0)}{x_{t_i-1} + a(i)} \left( x_{t_i} - x_{t_i-1} \right)
\]

\[
= x^0 \left( 1 + \frac{x_{t_i} - x_{t_i-1}}{x_{t_i-1}} \right) \frac{x_{t_i-1}}{x_{t_i-1} + a(i)} + \left( x^0 + \left( x_{t_i} - x_{t_i-1} \right) \right) \frac{a(i)}{x_{t_i-1} + a(i)}
\]

\[
+ \left( x_{t_i} - x_{t_i-1} \right) \frac{a(0) - a(i)}{x_{t_i-1} + a(i)}
\]

\[
= \alpha \left[ x^0 \left( 1 + \frac{x_{t_i} - x_{t_i-1}}{x_{t_i-1}} \right) \right]
\]

\[
+ (1 - \alpha) \left[ x^0 + \left( x_{t_i} - x_{t_i-1} \right) \right] + \frac{x_{t_i} - x_{t_i-1}}{x_{t_i-1} + a(i)} (a(0) - a(i))
\]

with \( \alpha = x_{t_i-1} / (x_{t_i-1} + a(i)) \).

Thus, by taking different displacements into account we obtain an additional term. This term can be reinterpreted in terms of the continuous model. The additional regime change term is the product of the relative changes of our displaced variable and the change of the displacement from the observed historical regime \( (a(i)) \) to the current regime \( (a(0)) \).
I.4 Displaced Relative Changes with GARCH Filtering

The displaced relative changes model discussed in Section I.3 defines a method for calculating “changes” in a historical simulation by calculating relative changes on a displaced variable. Instead of considering “standard returns” of the form \((x_t - x_{t-1})/(x_{t-1})\), we considered \(((x_t + a) - (x_{t-1} + a))/(x_{t-1} + a)\).

We now propose an extension of the model class of filtered historical simulation models by combining the displaced relative changes model and the filtered historical simulation approach proposed in Barone-Adesi et al. (2002). We call this the displaced filtered historical simulation (DFHS) model. That is, we apply a GARCH-type filtering to displaced relative changes. To show how this model extension is constructed, we (1) reproduce the original model by Barone-Adesi et al. (2002), (2) rewrite the model slightly so as to (3) apply the displacement extension to the model.

In Barone-Adesi et al. (1999), an ARMA-GARCH(1,1) model is considered:

\[
\begin{align*}
\eta_{t+1} &= \mu \eta_t + \theta \sigma_t \varepsilon_t + \sigma_{t+1} \varepsilon_{t+1} \\
\sigma_{t+1}^2 &= \omega + \alpha (\sigma_t^2 + \gamma^2) + \beta \sigma_t^2,
\end{align*}
\]

where \(\mu\) denotes the auto-regressive, \(\theta\) denotes the moving-average factor, and \(\sigma_0\) denotes the standard GARCH volatility; \(\sigma_{t+1} \varepsilon_{t+1}\) is the error term with \(\varepsilon_{t+1} \sim N(0, 1)\) and \(\omega, \alpha, \gamma\) are the usual constants controlling the behavior of \(\sigma_{t+1}\).

To remove the path dependency introduced by the time-lagged \(\varepsilon_t\) in Equation (I.9) and (I.10), we apply a state space extension and rewrite equations (I.9) and (I.10) using \(m_t := \sigma_t \varepsilon_t\), as explained below.

Using Equation (I.9) shifted back in time by one period, that is,

\[
y_t = \mu y_{t-1} + \theta m_{t-1} + m_t
\]

we obtain

\[
m_t = -\theta m_{t-1} + (y_t - \mu y_{t-1})
\]

and thus (replacing all \(\sigma_t \varepsilon_t\) by \(m_t\))

\[
\begin{align*}
y_{t+1} &= \mu y_t + \theta m_t + \sigma_{t+1} \varepsilon_{t+1} \\
\sigma_{t+1}^2 &= \omega + \alpha (m_t + \gamma^2) + \beta \sigma_t^2 \\
m_t &= -\theta m_{t-1} + (y_t - \mu y_{t-1}).
\end{align*}
\]

After rewriting, the state space transition is driven by a single factor \(\varepsilon_{t+1}\) only, which is the component that is to be sampled by means of historical simulation, where \(\sigma\) and \(m\) are interpreted as model parameters known at time \(t\).

The historical simulation producing equations based on relative changes can then be written
as:

\[
x^{i}_{T_1} := x_i \left( 1 + \mu \frac{x_t - x_{t-1}}{x_{t-1}} + \theta m_{t_i} + \sigma_{t_i} \sqrt{\Delta_{t+1} z_{t+1}} \right)
\]

\[
\sigma^2_{t_{i+1}} = \omega + \alpha \left( m_i + \gamma \right)^2 + \beta \sigma^2_i
\]

\[
m_{t} = -\theta m_{t-1} + \left( \frac{x_t - x_{t-1}}{x_{t-1}} - \mu \frac{x_{t-1} - x_{t-2}}{x_{t-2}} \right).
\]

In the last step, we combine filtered historical simulation with our proposed displacement extensions, resulting in the following system of equations.

\[
x^{i}_{T_1} := (x_i + a) \left( 1 + \mu \frac{x_t - x_{t-1}}{x_{t-1}} + \theta m_{t_i} + \sigma_{t_i} \sqrt{\Delta_{t+1} z_{t+1}} \right)
\]

\[
\sigma^2_{t_{i+1}} = \omega + \alpha \left( m_i + \gamma \right)^2 + \beta \sigma^2_i
\]

\[
m_{t} = -\theta m_{t-1} + \left( \frac{x_t - x_{t-1}}{x_{t-1}} + a - \mu \frac{x_{t-1} - x_{t-2}}{x_{t-2} + a} \right).
\]

This resulting model encompasses all the approaches that we consider in our empirical analysis. For \( \alpha = \beta = \theta = 0 \) and \( \mu = 1 \), we recover the DRC model, which itself comprises the modeling of relative changes and absolute changes through the parameter \( a \). For \( a = \theta = 0 \) and \( \mu = 1 \), we recover a FHS model based on a GARCH-type filtering. For \( a = 0 \), we recover a FHS model based on an ARMA-GARCH-type filtering.

Moreover, every parameter set can be associated, to some extent, with an interpretation. The parameter \( a \) decides which quantity is assumed to be i.i.d.-returns or absolute changes. The parameter \( \mu \) and \( \theta \) decide whether, on that quantity, a drift (or trend) is considered. The parameters \( \omega \), \( \beta \), and \( \theta \) decide whether filtering based on a GARCH volatility estimation will be performed.

### I.5 Estimation and Testing

In a practical application of a “historical simulation,” one would choose a data change model (say, relative or absolute changes) and then test its predictive power, e.g., via model back testing. Endowed with our one-parameter family of models, we are able to replace the choice of the data change model with an estimation of the displacement parameter \( a \).

In this section we describe the techniques related to estimating the displacement parameter and explain how to test its properties. For parameter estimation, we use maximum likelihood estimation. A more formal derivation of the maximum likelihood estimator of our displaced change model is given in Section I.9 of the Appendix. For testing, we consider back-testing Value-at-Risk calculation based on possible future scenarios generated by the different historical simulation models.

We use a likelihood function to calibrate parameters of the change function \( F \) and the
function $G$ of our historical simulation model. This likelihood function comes from a distributional assumption, namely, on the argument $\Delta Z$ of the function $F$. We choose this procedure because it is a practically feasible way of obtaining some estimates for model parameters. The likelihood function is just one way to define a target function for the optimization problem. Another way is to derive such a target function from a back-test procedure, a method that works without any distributional assumption on $\Delta Z$. However, due to space constraints, improvements to the estimation process are left for future research.

I.5.1 Maximum Likelihood Estimation for the Displaced Lognormal Model

In this section we derive the likelihood function for the displaced lognormal model given in Equation (I.6). We do not consider the GARCH filtering here, but this likelihood function can be used to estimate the GARCH parameters as well, and for the DFHS model we use this likelihood function to jointly estimate all parameters. Following Section I.3.1, we consider a time series $x_i$ and assume that it follows the (time-discrete) model

$$y_i := \log(x_{i+1} + a) - \log(x_i + a) = \mu \Delta t + \sigma \sqrt{\Delta t} Z_i,$$

where $Z_i$ are samples from a standard normal distribution. We wish to estimate $a$, $\mu$, and $\sigma$ in such a way as to maximize the likelihood of observing $y_i$. We find for fixed $a$ that $\mu$ and $\sigma$ are determined by

$$\mu(a) \Delta t := \frac{1}{n} \sum_{i=1}^{n} y_i,$$

$$\sigma(a)^2 \Delta t := \frac{1}{n} \sum_{i=1}^{n} (y_i - \mu(a))^2.$$

These formulas provide the maximum likelihood estimators for samples of a normally distributed random variable with mean $\mu(a) \Delta t$ and variance $\sigma^2(a) \Delta t$.

Given $\mu(a)$ and $\sigma(a)$, we may determine $a$ as

$$a := \arg \max_{a} L(a) = \arg \max_{a} \left( \sum_{i=1}^{n} \log \left( f(x_{i+1}; x_i, a) \right) \right), \quad (I.11)$$

25
with $L(a)$ denoting the log-likelihood function and where

$$
f(x_{i+1}; x_i, a) := \phi(x, \mu(a), \sigma(a)) \frac{dx}{dx_{i+1}} \big|_{x = \log(x_{i+1} + a) - \log(x_i + a)}
= \frac{1}{\sqrt{2\pi\sigma(a)}} \exp \left( -\frac{1}{2} \left( \frac{x - \mu(a)}{2\sigma(a)} \right)^2 \right) \frac{dx}{dx_{i+1}} \big|_{x = \log(x_{i+1} + a) - \log(x_i + a)}
= \frac{1}{\sqrt{2\pi\sigma(a)(x_{i+1} + a)}} \exp \left( -\frac{1}{2} \left( \frac{\log(x_{i+1} + a) - \log(x_i + a) - \mu(a)}{2\sigma(a)} \right)^2 \right).
$$

Note that $f$ is just the probability density of the log-normal distribution. There is a noteworthy subtlety to the likelihood function for the displaced model: when using log returns, the usual procedure is to consider the likelihood function of the normal distribution applied to the log-returns. The difference between the two approaches is the factor $1/(x_{i+1})$, which is a constant with respect to a parameter optimization and hence not relevant in the maximum likelihood optimization. Here, the factor $1/(x_{i+1} + a)$ is not a constant with respect to the parameter optimization.

### I.5.2 Backtesting the Historical Simulation

In the following subsection, we describe the back-testing used to test the models.

#### I.5.2.1 Average Backtest Hit Statistic

We take a fixed date $t_{k-1}$ based on which scenarios for the future date $t_k$ are generated.\(^7\)

To generate future scenarios we use the historical simulation approach on the previous $N$ dates $t_i$ (calibration period), with $k - N \leq i \leq k - 1$, and where $N$ determines the size of the calibration window. A back-test hit is defined as the value at date $t_k$ falling below the $p$-quantile predicted from the scenarios derived from the simulation. That is, the quantile hit indicator process $q$ is defined such that $q(t_k) = 1$ if the value at $t_k$ falls below the $p$-quantile and $q(t_k) = 0$ otherwise.

In theory, the probability of $q(t_k)$ being 1 is $p$. In our empirical analysis we generate $q(t_k)$ predictions for $M$ consecutive dates up to $t_n$, that is, $n - M < k \leq n$. For the date $t_n$, we then calculate the average of the back-test hit indicator values by summing up the series of $M$ back-test hit indicator values and dividing them by $M$. We call this number the average back-test hit statistic (ABHS) (a moving average over a sequence of 0s and 1s), which, ideally, should remain close to the analytically value of the quantile $p$.\(^8\) Then the model is tested by comparing the ABHS to the theoretical value $p$. That is, we check whether

$$ABHS = \frac{1}{M} \sum_{j=n-M+1}^{n} q(t_j) \approx p.$$  

---

\(^7\)For illustrative purpose, one can think of $t_{k-1}$ being the current data point.

\(^8\)Note that if the data set to be analyzed is of length $L$, the first $N + M$ observations are needed to calculate a first ABHS value and in total we can calculate $L - (N + M)$ ABHS values.
Note that this test does not require any distributional assumption about the time series.

I.5.2.2 Back-Test Stability for Different Calibration Window Sizes
In the first step of our analysis we choose $N = M = 250$ working days, that is, the size of calibration window equals the size of the averaging window. Our choice of the averaging period $M = 250$ is motivated by practice: back-test deviations are usually monitored and reported on an annual basis. However, the choice of the calibration window size $N = 250$ is arbitrary and thus we also analyze the overall stability of the model as a function of the calibration window size. We vary $N$ from 100 days to 1,000 days and compare model performance to the theoretical value, as well as comparing the performance of the different models.

I.6 Numerical Results
I.6.1 Data
To empirically analyze our model’s performance for different security classes, we use Swiss government bond yield (2-year maturity) data for the period 1999–2012 obtained from the Swiss National Bank and U.S. government bond spread data (10Y-3M) for the period 1992–2012. We choose this data series since one particular strength of our model is its ability to handle non-positive risk factors and these data have had non-positive values frequently in the past, as illustrated by Figure 1.

![Swiss Government Bond Yield Data](image1)

Figure 1: Swiss Government 2-Year Bond Yield Data for 1999–2012 (left) and U.S. Government 10-Year 3-Month Bond Yield Spread Data for 1992–2012 (right).

I.6.2 Average Back-Test Hits Statistic Results
For our empirical analysis we implement following models: the FHS model as proposed by Barone-Adesi et al. (1999), the DFHS model, a standard GARCH(1,1) model (G), a
displaced version of the GARCH model (DG), the displaced relative change model (DRC), and a standard historical simulation model (HS) with relative changes. In Figure 2, the 1% quantile back-test results for the Swiss government 2-year bond yield data are presented. In the five graphs of Figure 2, from left to right and from top to bottom, we test the models against each other in different combinations: the DFHS model against the FHS model, the displaced GARCH model against the GARCH model, the DFHS model against the displaced GARCH model, the DFHS model against the HS model, the displaced GARCH model against the HS model, and, finally, DRC model against HS model.

In each graph, the horizontal and vertical axis show the time period of the data set analyzed (in this case, 1999–2012) and the value of the average back-test hit statistics (ABHS), respectively. In the upper-left graph we analyze the ABHS for the DFHS (solid black line) and the FHS (solid gray line) model. The horizontal black dashed line shows the theoretical quantile based on which we test the models. If the ABHS value is below the dashed line, it means the model predicts quantile values that overestimate the actually occurring realizations of the state variable. That is, if we compare the predicted value of the state variable, then on average the occurring realization will be less negative than the predicted quantile value. In this sense, the model can be interpreted as giving a conservative prediction for the respective quantile realizations of the risk factor. Vice versa, if the ABHS value is above the dashed line, the model underestimates actually occurring quantile realizations of the risk factor: the occurring realizations violate the predicted quantile values more frequently than implied by the theoretical quantile.

Although these graphs present only a superficial visual comparison, they are indicative of some interesting results. When looking at the upper-left graph again, we observe that for the time period mid 2011 to mid 2012 the FHS model severely fails to predict the future quantile realizations. Even more noticeable, however, is that the FHS model consistently underestimates the quantile values, which produces a bias in the prediction that will expose anyone using that model for any kind of risk management to much larger risks than expected. Since during this period, negative Swiss government bond yields are observed, it is a particularly important period for testing our suggested model extension. And, in fact, the DFHS model does appear to do a much better job in dealing with these negative yields since the ABHS moves closely around the theoretic quantile value. The same result can be seen in the upper-right graph where the displaced GARCH model is compared with the standard GARCH model. The result is even more obvious in the two middle graphs and the lower-left graph where the DFHS, the displaced GARCH, and the DRC model are compared to the standard historical simulation. It is clear that the standard historical simulation model totally fails at handling close-to-zero and negative bond yields. Looking at the upper two graphs for the rest of the overall sample time period reveals, at most, that the displaced model alternative moves, on average, more closely around the theoretical quantile than the model without displacement and that it is more conservative
Figure 2: 1% Quantile Back-Test Results for Swiss Government 2-Year Bond Yield Data for 1999–2012

The figure presents the 1% quantile back-test results for the Swiss government 2-year bond yield data. The six graphs from left to right and from top to bottom show pair-wise model comparison for the DFHS model against the FHS model, the displaced GARCH model against the standard GARCH model, the DFHS against the HS model, the displaced GARCH against the HS model, the DRC against the HS model, and the DFHS model against the displaced GARCH model. In each graph, the calendar day of the forecast (Reference Date) on the x-axis is plotted against the value of the average back-test hit statistics (ABHS) on the y-axis. The ABHS is calculated based on a calibration window size of 250 days and an averaging time period of the same length.

In its predictions. Of course, drawing any conclusions from this rough visual test must be done with caution; however, the result is confirmed in a later analysis (see Section I.6.3). If we directly compare the two displaced model alternatives in the lower-right graph we again are tempted to say that the displaced GARCH model moves on average more closely around the theoretical quantile than does the DFHS model. This result, too, hinted at by the
Figure 3: 5% Quantile Back-Test Results for Swiss Government 2-Year Bond Yield Data for 1999–2012

The figure presents the 5% quantile back-test results for the Swiss government 2-year bond yield data. The six graphs from left to right and from top to bottom show pair-wise model comparison for the DFHS model against the FHS model, the displaced GARCH model against the standard GARCH model, the DFHS against the HS model, the displaced GARCH against the HS model, the DRC against the HS model, and the DFHS model against the displaced GARCH model. In each graph, the calendar day of the forecast (Reference Date) on the x-axis is plotted against the value of the average back-test hit statistics (ABHS) on the y-axis. The ABHS is calculated based on a calibration window size of 250 days and an averaging time period of the same length.

The smaller amplitude of the ABHS for the GARCH-type models compared to the FHS-type models, is confirmed by results presented in Section I.6.3.

All results discussed for the 1% quantile backtest hold in similar fashion for the 5% quantile back-test results, which are presented in Figure 3. If we look at the two graphs the superior performance of the displaced models over the models without the displacement features
becomes even more pronounced. Now, we can identify several time periods, e.g., years 2000–2004 in the left-upper graph, in which the bond yield is not negative and yet the DFHS model still outperforms the FHS model. The same result can be seen in the two middle graphs and the lower-left one comparing the performance of the DFHS, displaced GARCH, and the DRC model with that of the standard HS model. However, directly comparing DFHS against DG models (lower-right graph) does not result in the same obvious interpretation, as was also the case for the 1% quantile.

Looking at graphs (see 4 and 5) for the U.S. government spread data reveals results similar to, and even more consistent than, those derived for the Swiss government bond yield data. In general, the models with the displacement feature consistently outperform the models without this feature, both for the 1% and 5% quantile case. The displacement models’ stronger performance is especially pronounced for periods when the spread is close to zero or negative, for example, 1996, 1999, 2001, and 2007. Also the inability of the standard HS model to cope with situations where the spreads are close to zero or negative is clearly visible in the middle graphs and the lower-left graph of figures 4 and 5. Again, however, there is no major difference in performance between the DFHS and the DG models (see 4 and 5).

To summarize, our results indicate that, on average, a model with the displacement feature performs better than a model without this feature, most especially when risk factors are close to zero or even become negative. Among the two displaced model alternatives, the displaced GARCH model occasionally seems to do slightly better at predicting the VaR quintiles, but there is no indication yet that this is true for longer time periods.

I.6.3 Back-Test Stability for Different Calibration Window Sizes

The back-test results presented in Section I.6.2 are based on a averaging window size of \( M = 250 \), which, as mentioned in Section I.5.2.2, is motivated by the use of historical simulation methods in practice. However, the choice of the calibration window \( N = 250 \) is to some extent arbitrary. Therefore, in this section we test the robustness of our results with respect to calibration window size, \( N \), by varying it from 100 to 1,000. To make results comparable across data sets with varying time periods, we proceed as follows. First, we calculate all daily ABHS values for the whole data set based on a specific calibration window size. Then we sum up all daily ABHS values and divide them by the number of daily ABHS values, resulting in what we call the “summarized ABHS.” A summarized ABHS value of, e.g., 0.5 means that, on average, the daily ABHS value deviates from the theoretical quantile by 0.5%.

Table 1 reports results of the summarized ABHS for the 1% quantile back test of the Swiss 2-year government bond yields. For varying calibration periods, Columns 2–7 give the results for the DFHS, FHS, displaced GARCH, standard GARCH, DRC, and a standard
Figure 4: 1% Quantile Back-Test Results for U.S. Government Spreads (10Y-3M) for 1992–2012
The figure presents the 5% quantile back-test results for the U.S. government spreads (10Y-3M) data. The six graphs from left to right and from top to bottom show pair-wise model comparison for the DFHS model against the FHS model, the displaced GARCH model against the standard GARCH model, the DFHS against the HS model, the displaced GARCH against the HS model, the DRC against the HS model, and the DFHS model against the displaced GARCH model. In each graph, the calendar day of the forecast (Reference Date) on the x-axis is plotted against the value of the average back-test hit statistics (ABHS) on the y-axis. The ABHS is calculated based on a calibration window size of 250 days and an averaging time period of the same length.

historical simulation model. To compare the models to each other, we calculate several combinations of ratios of summarized ABHS values in Columns 8–13. These ratios can be interpreted as percentage relations between the performances of the models. If the ratio is less than 1, the model in the numerator is best; if it is larger than 1, the model in the denominator is best. A ratio of, e.g., 0.8 means that the model in the numerator has on
Figure 5: 5% Quantile Back-Test Results for U.S. Government Spreads (10Y-3M) for 1992–2012
The figure presents the 5% quantile back-test results for the U.S. government spreads (10Y-3M) data. The six graphs from left to right and from top to bottom show pair-wise model comparisons for the DFHS model against the FHS model, the displaced GARCH model against the standard GARCH model, the DFHS against the HS model, the displaced GARCH against the HS model, the DRC against the HS model, and the DFHS model against the displaced GARCH model. In each graph, the calendar day of the forecast (Reference Date) on the x-axis is plotted against the value of the average back-test hit statistics (ABHS) on the y-axis. The ABHS is calculated based on a calibration window size of 250 days and an averaging time period of the same length.

A comparison of the results for the DFHS and FHS models in Columns 2 and 3, and the displaced GARCH and the GARCH model in Columns 4 and 5, reveals an interesting result. For calibration window sizes up to 500 days, the model with the displaced feature always average a 20% lower daily deviation from the theoretical quantile than the model in the denominator.
Table 1: Swiss Government 2-Year Bond Yield Back Test: Rolling Windows, 1% Quantile

<table>
<thead>
<tr>
<th>Wind</th>
<th>DFHS</th>
<th>FHS</th>
<th>DG</th>
<th>G</th>
<th>DRC</th>
<th>HS</th>
<th>DFHS</th>
<th>FHS</th>
<th>DG</th>
<th>G</th>
<th>DRC</th>
<th>HS</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.53</td>
<td>0.62</td>
<td>0.46</td>
<td>0.60</td>
<td>0.46</td>
<td>0.79</td>
<td>0.87</td>
<td>0.77</td>
<td>1.15</td>
<td>0.59</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>150</td>
<td>0.49</td>
<td>0.78</td>
<td>0.46</td>
<td>0.47</td>
<td>0.43</td>
<td>1.00</td>
<td>0.63</td>
<td>0.98</td>
<td>1.07</td>
<td>0.50</td>
<td>0.46</td>
<td>0.43</td>
</tr>
<tr>
<td>200</td>
<td>0.53</td>
<td>0.90</td>
<td>0.42</td>
<td>0.70</td>
<td>0.55</td>
<td>1.18</td>
<td>0.59</td>
<td>0.61</td>
<td>1.25</td>
<td>0.45</td>
<td>0.36</td>
<td>0.46</td>
</tr>
<tr>
<td>250</td>
<td>0.68</td>
<td>0.99</td>
<td>0.50</td>
<td>0.70</td>
<td>0.65</td>
<td>1.30</td>
<td>0.69</td>
<td>0.73</td>
<td>1.34</td>
<td>0.52</td>
<td>0.39</td>
<td>0.50</td>
</tr>
<tr>
<td>300</td>
<td>0.87</td>
<td>1.03</td>
<td>0.49</td>
<td>0.97</td>
<td>0.78</td>
<td>1.61</td>
<td>0.84</td>
<td>0.50</td>
<td>1.79</td>
<td>0.54</td>
<td>0.30</td>
<td>0.49</td>
</tr>
<tr>
<td>350</td>
<td>0.45</td>
<td>1.06</td>
<td>0.49</td>
<td>0.64</td>
<td>0.80</td>
<td>1.58</td>
<td>0.43</td>
<td>0.77</td>
<td>0.92</td>
<td>0.29</td>
<td>0.31</td>
<td>0.50</td>
</tr>
<tr>
<td>400</td>
<td>0.52</td>
<td>0.74</td>
<td>0.54</td>
<td>1.23</td>
<td>0.82</td>
<td>1.82</td>
<td>0.71</td>
<td>0.44</td>
<td>0.97</td>
<td>0.29</td>
<td>0.29</td>
<td>0.45</td>
</tr>
<tr>
<td>450</td>
<td>0.53</td>
<td>1.23</td>
<td>0.58</td>
<td>0.64</td>
<td>0.86</td>
<td>1.86</td>
<td>0.43</td>
<td>0.90</td>
<td>0.91</td>
<td>0.28</td>
<td>0.31</td>
<td>0.46</td>
</tr>
<tr>
<td>500</td>
<td>0.68</td>
<td>1.30</td>
<td>0.49</td>
<td>0.78</td>
<td>0.89</td>
<td>1.95</td>
<td>0.52</td>
<td>0.63</td>
<td>1.39</td>
<td>0.35</td>
<td>0.25</td>
<td>0.46</td>
</tr>
<tr>
<td>550</td>
<td>0.55</td>
<td>0.55</td>
<td>0.49</td>
<td>0.68</td>
<td>0.97</td>
<td>2.02</td>
<td>1.00</td>
<td>0.72</td>
<td>1.12</td>
<td>0.27</td>
<td>0.24</td>
<td>0.48</td>
</tr>
<tr>
<td>600</td>
<td>0.65</td>
<td>0.56</td>
<td>0.63</td>
<td>0.52</td>
<td>0.91</td>
<td>2.12</td>
<td>1.17</td>
<td>1.22</td>
<td>1.03</td>
<td>0.31</td>
<td>0.30</td>
<td>0.43</td>
</tr>
<tr>
<td>650</td>
<td>0.61</td>
<td>0.62</td>
<td>0.53</td>
<td>0.55</td>
<td>0.91</td>
<td>2.27</td>
<td>0.99</td>
<td>0.95</td>
<td>1.17</td>
<td>0.27</td>
<td>0.23</td>
<td>0.40</td>
</tr>
<tr>
<td>700</td>
<td>0.63</td>
<td>0.58</td>
<td>0.56</td>
<td>0.53</td>
<td>0.91</td>
<td>2.26</td>
<td>1.08</td>
<td>1.05</td>
<td>1.13</td>
<td>0.28</td>
<td>0.25</td>
<td>0.40</td>
</tr>
<tr>
<td>750</td>
<td>0.58</td>
<td>0.57</td>
<td>0.57</td>
<td>0.55</td>
<td>0.84</td>
<td>2.29</td>
<td>1.01</td>
<td>1.04</td>
<td>1.01</td>
<td>0.25</td>
<td>0.25</td>
<td>0.36</td>
</tr>
<tr>
<td>800</td>
<td>0.60</td>
<td>0.55</td>
<td>0.57</td>
<td>0.48</td>
<td>0.86</td>
<td>2.36</td>
<td>1.09</td>
<td>1.18</td>
<td>1.06</td>
<td>0.25</td>
<td>0.24</td>
<td>0.37</td>
</tr>
<tr>
<td>850</td>
<td>0.52</td>
<td>0.48</td>
<td>0.64</td>
<td>0.52</td>
<td>0.84</td>
<td>2.41</td>
<td>1.09</td>
<td>1.23</td>
<td>0.82</td>
<td>0.22</td>
<td>0.27</td>
<td>0.35</td>
</tr>
<tr>
<td>900</td>
<td>0.51</td>
<td>0.49</td>
<td>0.55</td>
<td>0.50</td>
<td>0.87</td>
<td>2.42</td>
<td>1.04</td>
<td>1.11</td>
<td>0.92</td>
<td>0.21</td>
<td>0.23</td>
<td>0.36</td>
</tr>
<tr>
<td>950</td>
<td>0.52</td>
<td>0.52</td>
<td>0.54</td>
<td>0.51</td>
<td>0.88</td>
<td>2.45</td>
<td>1.01</td>
<td>1.06</td>
<td>0.95</td>
<td>0.21</td>
<td>0.22</td>
<td>0.36</td>
</tr>
<tr>
<td>1000</td>
<td>0.54</td>
<td>0.45</td>
<td>0.55</td>
<td>0.57</td>
<td>0.96</td>
<td>2.41</td>
<td>1.20</td>
<td>0.96</td>
<td>0.97</td>
<td>0.22</td>
<td>0.22</td>
<td>0.40</td>
</tr>
</tbody>
</table>

This table presents the 1% quantile back-test results for the Swiss government 2-year bond yield data for 1990–2012. The back-test statistic described in Section 1.5.2 is calculated for calibration window sizes varying from 100 to 1,000 days. To make the results comparable across several data sets with different time spans, the test statistic is summed over all daily ABHS values and normalized by the number of ABHS values. Results for the displaced filtered historical simulation (DFHS) model, filtered historical simulation (FHS) model, displaced GARCH (DG) model, standard GARCH (G) model, the displaced relative change model (DRC), and a standard historical simulation (HS) model are presented in Columns 2–7. Columns 8–13 present percentage relationships between the test statistics of the different models. We also test whether the average daily ABHS of the model in the numerator and that of the model in the denominator are significantly different. The stars (**, ***), denote significantly different values at the 10%, 5%, 1% level obtained from a standard two-sample t-test between the ABHS time series.

shows a smaller average daily summarized ABHS than the model without the displaced feature. This means that the model with the displaced feature performs better than the alternative model. This result can also be seen in Columns 8 and 9, which present the calculated ratios of summarized ABHS values. When comparing the DFHS model with the FHS model in Column 8, the DFHS model shows, on average, between 13 to 57% less daily deviation from the theoretical quantile for calibration window sizes up to 500 days. Comparing the displaced GARCH model with the standard GARCH model brings
This table presents the 1% quantile back-test results for the U.S. government bond spreads (10Y-3M) for 1992–2012. The back-test statistic described in Section I.5.2 is calculated for calibration window sizes varying from 100 to 1,000 days. To make the results comparable across several data sets with different time spans, the test statistic is summed over all daily ABHS values and normalized by the number of ABHS values. Results for the displaced filtered historical simulation (DFHS) model, filtered historical simulation (FHS) model, displaced GARCH (DG) model, standard GARCH (G) model, the displaced relative change model (DRC), and a standard historical simulation (HS) model are presented in Columns 2–7. Columns 8–13 present percentage relationships between the test statistics of the different models. We also test whether the average daily ABHS of the model in the numerator and that of the model in the denominator are significantly different. The stars (\(^\ast\), \(^{\ast\ast}\), \(^{\ast\ast\ast}\)) denote significantly different values at the 10%, 5%, 1% level obtained from a standard two-sample t-test between the ABHS time series.

This result increases the attractiveness of the models with the displacement feature. Needing less data to calibrate the models makes them easier to handle. Furthermore, the models are able to produce predictions for the quantiles that are better than or at least equal to the theoretical quantile for calibration window sizes up to 500 days.

However, for longer calibration window sizes, the values of displacement and alternative models converge and the relationship even reverses slightly. From a practical perspective, this result increases the attractiveness of the models with the displacement feature. Needing less data to calibrate the models makes them easier to handle. Furthermore, the models are able to produce predictions for the quantiles that are better than or at least equal to the

<table>
<thead>
<tr>
<th>Wind</th>
<th>DFHS</th>
<th>FHS</th>
<th>DG</th>
<th>G</th>
<th>DRC</th>
<th>HS</th>
<th>DFHS</th>
<th>DG</th>
<th>FHS</th>
<th>DG</th>
<th>HS</th>
<th>DRC</th>
<th>HS</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.49</td>
<td>3.37</td>
<td>0.43</td>
<td>3.07</td>
<td>0.45</td>
<td>7.94</td>
<td>0.15</td>
<td>0.14</td>
<td>1.15</td>
<td>0.06</td>
<td>0.05</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>150</td>
<td>0.58</td>
<td>2.02</td>
<td>0.46</td>
<td>1.99</td>
<td>0.48</td>
<td>8.26</td>
<td>0.29</td>
<td>0.23</td>
<td>1.27</td>
<td>0.07</td>
<td>0.06</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0.69</td>
<td>1.32</td>
<td>0.52</td>
<td>1.13</td>
<td>0.63</td>
<td>8.62</td>
<td>0.53</td>
<td>0.46</td>
<td>1.33</td>
<td>0.08</td>
<td>0.06</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>0.65</td>
<td>1.35</td>
<td>0.64</td>
<td>1.24</td>
<td>0.72</td>
<td>8.94</td>
<td>0.48</td>
<td>0.51</td>
<td>1.02</td>
<td>0.07</td>
<td>0.07</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.73</td>
<td>1.42</td>
<td>0.59</td>
<td>0.90</td>
<td>0.75</td>
<td>9.13</td>
<td>0.51</td>
<td>0.65</td>
<td>1.24</td>
<td>0.08</td>
<td>0.06</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>350</td>
<td>0.63</td>
<td>0.79</td>
<td>0.64</td>
<td>1.12</td>
<td>0.83</td>
<td>9.35</td>
<td>0.80</td>
<td>0.57</td>
<td>0.98</td>
<td>0.07</td>
<td>0.07</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>0.65</td>
<td>0.85</td>
<td>0.77</td>
<td>0.94</td>
<td>0.83</td>
<td>9.59</td>
<td>0.76</td>
<td>0.82</td>
<td>0.84</td>
<td>0.07</td>
<td>0.08</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>450</td>
<td>0.71</td>
<td>1.28</td>
<td>0.85</td>
<td>0.99</td>
<td>0.92</td>
<td>9.80</td>
<td>0.56</td>
<td>0.86</td>
<td>0.83</td>
<td>0.07</td>
<td>0.09</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.67</td>
<td>0.97</td>
<td>0.65</td>
<td>0.97</td>
<td>0.90</td>
<td>9.85</td>
<td>0.70</td>
<td>0.67</td>
<td>1.05</td>
<td>0.07</td>
<td>0.07</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>550</td>
<td>0.67</td>
<td>0.91</td>
<td>0.59</td>
<td>0.94</td>
<td>0.91</td>
<td>10.04</td>
<td>0.74</td>
<td>0.63</td>
<td>1.13</td>
<td>0.07</td>
<td>0.06</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>600</td>
<td>0.62</td>
<td>0.92</td>
<td>0.62</td>
<td>0.94</td>
<td>0.92</td>
<td>10.15</td>
<td>0.67</td>
<td>0.67</td>
<td>0.99</td>
<td>0.06</td>
<td>0.06</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>650</td>
<td>0.63</td>
<td>0.93</td>
<td>0.61</td>
<td>0.92</td>
<td>0.92</td>
<td>10.26</td>
<td>0.68</td>
<td>0.67</td>
<td>1.03</td>
<td>0.06</td>
<td>0.06</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>700</td>
<td>0.63</td>
<td>0.92</td>
<td>0.60</td>
<td>0.94</td>
<td>0.94</td>
<td>10.33</td>
<td>0.69</td>
<td>0.64</td>
<td>1.05</td>
<td>0.06</td>
<td>0.06</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>750</td>
<td>0.61</td>
<td>0.90</td>
<td>0.58</td>
<td>0.92</td>
<td>0.95</td>
<td>10.40</td>
<td>0.68</td>
<td>0.64</td>
<td>1.05</td>
<td>0.06</td>
<td>0.06</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>0.57</td>
<td>0.94</td>
<td>0.56</td>
<td>0.94</td>
<td>0.96</td>
<td>10.55</td>
<td>0.61</td>
<td>0.60</td>
<td>1.01</td>
<td>0.05</td>
<td>0.05</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>850</td>
<td>0.58</td>
<td>0.87</td>
<td>0.55</td>
<td>0.85</td>
<td>0.98</td>
<td>10.67</td>
<td>0.67</td>
<td>0.66</td>
<td>1.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>900</td>
<td>0.56</td>
<td>0.86</td>
<td>0.48</td>
<td>0.82</td>
<td>1.01</td>
<td>10.81</td>
<td>0.65</td>
<td>0.56</td>
<td>1.17</td>
<td>0.05</td>
<td>0.05</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>950</td>
<td>0.52</td>
<td>0.84</td>
<td>0.48</td>
<td>0.85</td>
<td>0.98</td>
<td>10.94</td>
<td>0.62</td>
<td>0.57</td>
<td>1.08</td>
<td>0.05</td>
<td>0.05</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>0.51</td>
<td>0.92</td>
<td>0.48</td>
<td>0.87</td>
<td>0.99</td>
<td>11.08</td>
<td>0.56</td>
<td>0.55</td>
<td>1.08</td>
<td>0.05</td>
<td>0.04</td>
<td>0.09</td>
<td></td>
</tr>
</tbody>
</table>

This table presents the 1% quantile back-test results for the U.S. government bond spreads (10Y-3M) for 1992–2012. The back-test statistic described in Section I.5.2 is calculated for calibration window sizes varying from 100 to 1,000 days. To make the results comparable across several data sets with different time spans, the test statistic is summed over all daily ABHS values and normalized by the number of ABHS values. Results for the displaced filtered historical simulation (DFHS) model, filtered historical simulation (FHS) model, displaced GARCH (DG) model, standard GARCH (G) model, the displaced relative change model (DRC), and a standard historical simulation (HS) model are presented in Columns 2–7. Columns 8–13 present percentage relationships between the test statistics of the different models. We also test whether the average daily ABHS of the model in the numerator and that of the model in the denominator are significantly different. The stars (\(^\ast\), \(^{\ast\ast}\), \(^{\ast\ast\ast}\)) denote significantly different values at the 10%, 5%, 1% level obtained from a standard two-sample t-test between the ABHS time series.

This result increases the attractiveness of the models with the displacement feature. Needing less data to calibrate the models makes them easier to handle. Furthermore, the models are able to produce predictions for the quantiles that are better than or at least equal to the theoretical quantile for calibration window sizes up to 500 days. However, for longer calibration window sizes, the values of displacement and alternative models converge and the relationship even reverses slightly. From a practical perspective, this result increases the attractiveness of the models with the displacement feature. Needing less data to calibrate the models makes them easier to handle. Furthermore, the models are able to produce predictions for the quantiles that are better than or at least equal to the theoretical quantile for calibration window sizes up to 500 days.
predictions from models without the displacement feature, but they need less historical data to do so. This is helpful in situations when there is no long historical time series for a risk factor. Also, from a practical point of view, it is desirable to use only the latest historical observations to calibrate the model rather than outdated market events.

Comparing the DFHS and displaced GARCH models (see Column 10) does not clearly point to one model outperforming the other. For very short calibration window sizes, the displaced GARCH model does better than the DFHS model, but the situation reverses for \( N = 350 - 450 \) and reverses again for \( N = 500 - 800 \) and then reverses again for \( N = 850 - 1000 \). However, we note that the displaced GARCH model exhibits better performance for a greater number of calibration window sizes than does the DFHS model.

In Columns 11, 12, and 13, the performance of the DFHS, displaced GARCH, and DRC models is compared to the standard HS model. As already established by the graphical results in Section I.6.2, on average the HS model is strongly outperformed by models having the displacement feature. However, this result is mainly driven by the fact that the HS model fails in handling the close-to- or below-zero observations for the Swiss government bond yields around the years 2011 and 2012.

Furthermore, we test whether the results in Table 1 are statistically significantly different from each other. We perform a two-sample t-test on the difference of average daily ABHS values for two models. The results are indicated by the stars in Columns 8-13, where *, **, *** denote significantly different means at the 10%, 5%, 1% level, respectively. The tests reveal that, overall, the average daily ABHS values for the different models are statistically significantly different from each other. Indeed, in most cases, we even obtain statistical difference at the 1% level.

The results for the U.S. government spread data in Table 2 are in general very similar to those for the Swiss government bond yield data; the only noticeable difference is that variation in model performance is even more robust.

For example, looking at the DFHS and FHS models in Columns 2, 3, and 7 shows that the DFHS model consistently outperforms the FHS model for all calibration window sizes by having, on average, between 20 to 85% less daily deviation from the theoretical quantile. As for the displaced GARCH and GARCH models, the displaced GARCH model shows, on average, between 14 to 86% less daily deviation from the theoretical quantile. As we might by now expect, the comparison between the DFHS model and the displaced GARCH model is one of mixed results, with the displaced GARCH model showing now less daily deviation from the theoretical quantile for 16 out of 20 calibration size windows. And just as equally expected, the DFHS, the DRC, and the displaced GARCH model strongly and consistently outperform the standard HS model.
I.7 Conclusion

In this paper we propose the displaced relative change model in conjunction with a historical simulation approach as a solution for generating historical simulations for possibly non-positive-valued risk variables, such as interest rate spreads and government bond yields. Additionally, the displaced relative change model has desirable features from a regulatory perspective. When conducting historical simulations changing from the concept of relative changes to absolute changes due to a regime switch in the data poses a operational change between from one model to another which in general has to be approved by a regulator. The displaced relative change model automatically interpolates between relative and absolute changes driven by the needs of the data it and based on a sound methodological approach. We also extend the filtered historical simulation model originally proposed by Barone-Adesi et al. (1999) by adding our displacement feature, naming the result the "displaced filter historical simulation model." In an empirical study based on a VaR analysis we investigate the performance of the models that include this new displacement feature by applying them to Swiss government bond yield data and U.S. government spread (10Y-3M) data. We find that models with our displacement feature strongly outperform models without the displacement feature, particularly in situations when the risk factors are close to zero or even negative. Furthermore, the displaced models perform better for shorter calibration windows, a strong practical advantage. The standard historical simulation model totally fails to handle close-to-zero and non-positive risk factors in our empirical analysis.
I.8 References


I.9 Maximum Likelihood Estimation for Itô Processes

Below is a short derivation of the maximum likelihood method used to estimate the displacement parameter.

Consider a numerical scheme

$$X(t_{i+1}) = F(\Delta W(t_i); X(t_i); \alpha_1, \ldots, \alpha_l),$$

where $\Delta W(t_i)$ is a Brownian increment over the period $[t_i, t_{i+1}]$, $F : \mathbb{R}^{1 \times 1 \times l} \to \mathbb{R}$ constitutes the model, and $\alpha_1, \ldots, \alpha_l$ are model parameters. Given observed realizations $X(t_{i+1}) = x_{i+1}$, we want to determine the model parameters $\alpha_1, \ldots, \alpha_l$ such that the conditional likelihood of the observations is maximized. That is, for each $i$ we maximize the likelihood of $X(t_{i+1}) = x_{i+1}$ given $X(t_i) = x_i$. In other words, we maximize

$$\sum_i \log(\psi(x_{i+1}|x_i)) \quad (I.12)$$

where $\psi(x_{i+1}|x_i)$ is the probability density of $F(\Delta W(t_i); X(t_i); \alpha_1, \ldots, \alpha_l)$.

In general, if $Y$ has a probability density $y \to \phi(y)$ and $Z = f(Y)$ we find from the substitution

$$\phi(y)dy = \phi(f^{-1}(z)) \frac{df^{-1}(y)}{dz} dz,$$

that $Z$ has a probability density

$$\psi(z) := \phi(f^{-1}(z)) \frac{df^{-1}(z)}{dz}.$$

This allows us to write Equation (I.12) in terms of $F$ and the density of $\Delta W(t_i, t_{i+1})$, which is known to be $\phi(y) = \frac{1}{\sqrt{2\pi \Delta t}} \exp\left(-\frac{y^2}{2\Delta t}\right)$. 

40
II Empirical Corporate Finance

Calls of Convertible Debt Securities: No Bad News at All

Abstract
In this paper, I examine the impact of in-the-money convertible bond calls on stock prices, employing a sample of U.S. convertible bond calls over the period 1994-2011. In contrast to previous literature, I find that conversion forcing convertible bond calls do not significantly influence stock prices. I posit that the discrepancy between my results and those in the literature is caused by amplified screening criteria, especially strong news cleaning. Companies tend to announce calls as side notes to other major corporate news, resulting in an event study bias. Further, convertible bond design, moneyness of the conversion option at the announcement date, and convertible-arbitrage strategies cast doubt on previous literature’s negative abnormal returns.
II.1 Introduction

In 2009, convertible bonds accounted for more than 5% of all fixed income securities issued, or more than US$23 billion in absolute amounts. In the literature, it is reported that more than 50% of all issued convertible bonds are called and that the announcement generates abnormal returns depending on the moneyness of the conversion option. If a rational investor decides on equity instead of accepting the company’s cash offer, the call announcement is called in-the-money or a forced conversion. In contrast to out-of-the-money calls, announcements of forced conversions are found to decrease the company’s market value. Bechmann (2004) and Brick et al. (2007), among others, find significant negative impacts of about 1% on the announcement date. However, the existing empirical literature cannot conclusively explain the drivers behind the abnormal returns of forced conversions. Although several theories have been proposed during the last few decades, none proves to be robust.

Mikkelson (1981) was the first to report a negative impact of forced conversions on share prices. He argues that the company’s decreased value is due to loss of a tax shield. Despite their equity feature, according to U.S. GAAP accounting standards, a convertible bond is allocated to debt. Since the bonds are exchanged for new shares, not only does debt vanish but, also, equity increases, which has a significant impact on the company’s capital structure. However, later studies (e.g., Byrd and Moore, 1996) cast doubt on this explanation. Preferred dividends provide no tax shield and the same effect is found in convertible preferred stock calls.

In a theoretical asymmetric information model, Harris and Raviv (1985) argue that managers with negative information call the bonds to preserve cash flows. Managers with positive information wait until the news is priced and bondholders elect to convert voluntarily. To assess the information content, Ofer and Natarajan (1987) examine the long-term evolution of share prices after the call. They find abnormal negative returns after the call, supporting the asymmetric information theory. However, Cowan et al. (1990) and others link the abnormal returns to a bias in the estimation procedure. To overcome the drawbacks of long-term event studies, Byrd and Moore (1996) investigate earnings forecasts after the call event. Hypothetically, earnings forecasts should be revised downward after a call, but the authors find the opposite; earnings forecasts rise after a forced-conversion call. Analogously, Ederington and Goh (2001) report that insiders buy shares before a call rather than selling and that analysts revise their forecasts upward instead of downward. The authors thus reject the hypothesis of negative information.

Mazzeo and Moore (1992) report an increase in trading volume of 44% in their sample.

---

9 Mergent FISD database, U.S. headquartered companies in non-regulated industries.
10 Moneyness of the conversion option is defined as the ratio of the equity value received by conversion (conversion value) divided by the amount of cash obtained from immediate redemption (nominal plus accrued interest) minus one.
11 For a short discussion, see Section II.4.
during the week of the call announcement compared to weeks before the announcement. Together with a recovery of the share prices during the conversion period, the authors suggest the effect may be caused by market micro-structure effects of liquidity. Following Mazzeo and Moore (1992), Bechmann (2004) argues that the immediate drop and then recovery of stock prices can be attributed to short selling by convertible arbitrageurs hedging their equity risk during the conversion period. Using monthly short-selling data, the authors conclude that hedging-induced price pressure at least partly drives the announcement returns. Brick et al. (2007) test the hypothesis with different proxies for liquidity to circumvent short-selling data and conclude that liquidity does not drive abnormal returns, nor do liquidity effects in combination with dividend dilution. Shareholders might expect their dividends to be diluted by the newly issued shares. Also, the recovery period is not related to the abnormal returns and neither are proxies for asymmetric information or tax shields.

To summarize, the question of what causes the abnormal returns remains unresolved. The objective of this paper is to employ robust event study methodology to assess the pure impact of in-the-money convertible bond calls on share prices. Event studies are designed to estimate the impact of new information on a company by relating the real return to a “should be” return without the new information. However, for an event study to be accurate, the released news has to contain information new to market participants and no other news can be released. I argue that an event study on the call of convertible bonds violates these two assumptions.

The news content of forced conversions is doubtful, for two reasons. First, due to market restrictions and financing problems, some convertible bonds are issued with the purpose of being converted into equity in the following years (Lewis et al., 1999). Their covenants underline this backdoor equity purpose. The second reason stems from the moneyness of the conversion option at the official call. Empirical literature on call policies shows that bonds are called extremely deep in-the-money with a strong out performance of the company before the call. On average, stock prices exceed the break-even point, meaning that an investor is indifferent between cash and stock, by about 50% (e.g., Bechmann, 2004). Any informed market participant would expect the bonds to already be called in an environment of optimal call policies.

Multiple corporate news releases during the event period can distort the effect to be investigated. Indeed, the information technology now generates near constant news coverage of companies. It is often assumed in the literature that the effects of other contemporaneous corporate events will wash out in the whole sample. This can be problematic and lead to a bias when the contemporaneous effects are supposed to have a larger impact or they all force a similar market reaction. King and Mauer (2012) report that about 31% of their convertible sample drops due to contemporaneous news. Indeed, several papers, (e.g., García-Feijóo et al., 2010) find evidence for Mayers’s (1998) sequential financing
hypothesis: companies issue callable bonds to overcome agency conflicts of debt and call their bonds to exercise investment options. The announcement of investment options can be simultaneous with the call event. Despite a potentially low information content, hedging pressure due to arbitrage strategies can introduce a liquidity depending share price reaction. However, reasonable doubts for market micro-structure effects can be found in the convertible arbitrage literature. Among others, Brown et al. (2012) argue that the largest share of a convertible bond issue is sold to hedge funds. Hedge fund’s arbitrage strategies can be replicated with buy-and-hold portfolios (Agarwal et al., 2011), leaving only minor “free float” that might be hedged at the announcement day.

In this paper, I employ an extensive and comprehensive news cleaning because convertible calls are often side notes to major corporate announcements that might overlay the call effect. I also control for the rolling of convertible bonds. Companies tend to ensure their financing by issuing a new bond before calling the outstanding bond. As a consequence of this cleaning, I find no abnormal stock returns when forced conversions are announced. In contrast to the cleaned sample, a sample with less restrictive selection criteria, comparable to those in previous research, reveals negative abnormal returns and differs significantly from the cleaned sample. I thus conclude that conversion-forcing calls have no impact on stock prices, which stands in sharp contrast to existing literature.

The rest of the paper is organized as follows. Section II.2 sets out my hypothesis in detail. Section II.3 describes my sample and compares it to that used by others. Based on the sample, Section II.4 reports abnormal returns around the announcement date and the impact of explanatory variables on these. Section II.5 concludes.

II.2 Hypothesis Development

This section sets out my hypothesis regarding the announcement effect of in-the-money calls of convertible bonds. During the trading days prior to the call announcement, companies tend to strongly out-perform the market. Bechmann (2004) reports 8% out-performance in the period 50 days before the announcement and Cowan et al. (1990) find significant alphas of 5% in the year before the announcement. Due to this out performance in a relative short time, simultaneous with the call, the average conversion option is extremely deep in the money. Brick et al. (2007), Bajo and Barbi (2012), and Bechmann (2004) report moneyness of 64%, 88%, and 47%, respectively. In contrast, the theoretical models of Ingersoll (1977) and Brennan and Schwartz (1977) implicate that it is optimal to call a bond as soon as the conversion value exceeds the value of immediate redemption. In an environment of optimal call policies market participants should have expected the convertibles to be called within a reasonable time frame. This expectation should already be reflected in the stock price and not lead to
abnormal returns at the announcement date.\textsuperscript{12}

Convertible bonds are highly flexible financing instruments and there is an extensive literature on their design. Firms can utilize convertible bonds to issue new shares indirectly by emphasizing the equity feature of the bond (Stein, 1992). Following Stein (1992), Lewis et al. (1999) find that companies design their new convertible issuances either more debt-like or more equity-like, depending on their specific financing problems. Testing this hypothesis, King and Mauer (2012) utilize the delta of the conversion option, the probability that the bond will be converted to equity at maturity, in their study. They find that the probability of called bonds to be converted is significantly higher than that of those not called at the date of issuance. Moreover, in the sample of Bechmann (2001), in-the-money calls are called significantly earlier than out-of-the-money calls. Hence, an indication for a seasoned equity offering through the back door can be obtained as data on moneyness or the delta of the implied conversion option are publicly available.

Mazzeo and Moore (1992) discover an increase in trading volume of 44\% in the week around the announcement relative to an average week before the announcement. On this basis, Bechmann (2004) attributes the abnormal returns to convertible-arbitrage hedge funds that hedge the equity risk during the conversion period by shorting shares at the time of announcement. However, there is considerable evidence that convertible-arbitrage hedge funds already hold most of the convertible bonds during maturity and have hedged their equity risk: Brown et al. (2012) report that about \textit{73\%} of all privately placed convertibles were sold to convertible-arbitrage hedge funds at issuance between 2000 and 2008 with a maximum of 80\% in 2008. Supportingly, Mitchell et al. (2007) find that convertible arbitrage funds account for 75\% of the convertible market. Most convertible-arbitrage strategies involve a short selling of the underlying stock. In line with a direct hedging demand of the hedge funds is the negative announcement effect on the issuance day, see de Jong et al. (2013). Agarwal et al. (2011) find that a simple buy(short)-and-hold strategy explains most of the hedge fund variation. Consequently, it can be assumed that hedge funds do not trade their convertibles and their equity risk is hedged at convertible issuance. Based on this literature review I formulate my two, potentially interacting, hypotheses which I am going to test in this paper. Since the hedging demand of convertible-arbitrage hedge funds is low and the negligible information content of a call due to the moneyness of the conversion option and bond design, I implicate that in-the-money convertible calls do not lead to abnormal trading volume in equity shares and in-the-money convertible calls do not generate abnormal stock returns.

\textsuperscript{12}All three market efficiencies according to Fama (1970) imply that all past information is reflected in the current stock price.
II.3 Data

II.3.1 Sources and Screening Criteria

I start by collecting convertible calls from the Bloomberg database. Bloomberg has similar coverage as Datastream, but is slightly more reliable. The sample covers the period 1994 to 2011, is restricted to the U.S. market, and limited to convertible bonds denoted in local currency and exchangeable for common stock. Also, I include call events only when the conversion option is in-the-money, meaning that a rational bondholder would decide to convert the bond into new shares. I employ nine other screening criteria, detailed below. I merge the sample with the Mergent FISD database to obtain call and issuance characteristics of the bonds and (1) require each bond to be covered by the database. Equity related data including SIC codes are taken from the CRSP tapes. In the case of (2) missing coverage or missing returns, the event is dropped. Further, redemptions resulting from other corporate decisions (delisting, bankruptcy) are excluded (3) and I exclude calls if the company announced that it would issue new convertibles one month before the call (4). A call event is excluded if (5) the underlying company is in the regulated financial or utility industry.\(^\text{13}\) I follow García-Feijóo et al. (2010) and King and Mauer (2012) and drop (6) potentially disturbing covenants such as variable coupon bonds and liquid yield option notes because the latter are putable for cash, which can give rise to different motives for calling the bond. Issuers sometimes clean up their convertibles after most of the holders have already converted, to avoid small outstanding amounts. I thus use (7) only calls where the call amount is larger than one-half of the offering amount. Contemporaneous events can bias the results. I use the Dow Jones FACTIVA database to check for public, company-related information released +/- two days around the announcement. An event is (8) removed if there was another major corporate event during this period that might disturb stock market reactions to convertible calls. Using publicly available call announcements, I double-checked all call features and announcement dates. If I (9) could not find an official announcement or not all call features could be discerned by looking at all the databases, I drop the event.

Returns of value-weighted industry indices and the Fama-French risk factors are publicly available through Kenneth French’s database.\(^\text{14}\)

II.3.2 Sampling and Sample Size

The initial sample from the Bloomberg database yields 522 call announcements. After controlling for regulated industries and official announcements in the FACTIVA database

\(^{13}\)SIC codes 6000-6999 and 4800-4999. If there is disagreement in the codes assigned between the Bloomberg and CRSP database, and one code falls in the regulated industries, I exclude the event.

\(^{14}\)http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/
the sample drops to 320. In the same database I find that 114 of the 320 events are free of contemporaneous news, which include merger activity.\textsuperscript{15}

To assess the representativeness of my sample, I create and employ a Dirty sample, based on selection criteria (1), (2), (5), (6) and (9), in addition to the Clean sample, which meets all the criteria specified above. The Dirty sample is intended to reveal differences that arise due to omitted screening criteria. The Clean sample consists of 47 calls; there are 150 in the Dirty sample.\textsuperscript{16} My sample is smaller than those used in other convertible call literature. For example, Brick et al. (2007) employ 392 calls made over 19 years (1985-2002), an extension of Ederington and Goh’s (2001) sample. García-Feijóo et al. (2010) employ 165 non-financial calls between 1986 and 2001. In contrast to this literature, King and Mauer (2012) use a sample of non-financial bonds issued between 1980 and 2002. They find 175 calls between 1994 and 2010 before news filtering and Bajo and Barbi (2012) find 92 calls for 2000 to 2009, a subset of my time period. The general pattern in these other samples is consistent with reported decreasing issuances of convertible bonds and thus goes some way to explaining the size of my sample. Additionally, my screening criteria are considerably stricter than in the earlier literature, but are necessary to avoid distorted abnormal returns. Research on convertible bonds is challenging since the flexibility in contracting yields a strongly heterogeneous set of convertible bonds. Certain covenants can generate incentives to call the bond that are special among the set of call events. I follow García-Feijóo et al. (2010) and exclude putable and variable coupon bonds. The opportunity to call a convertible bond generates financial flexibility; for example, it allows to respond to cheaper refinancing possibilities due to decreased credit risk. Many issue announcements of convertibles explicitly state that the bonds to be sold will replace existing convertible bonds. Even if this is not explicitly stated, depending on characteristics of the new bonds, market participants might expect the outstanding bond to be exchanged. Thus, an informed market participant foresees the call of the outstanding bonds. Also, there is considerable evidence that firms perform poorly around the offering (see Dann and Mikkelson, 1984).\textsuperscript{17}

This type of stringent screening is not done in earlier literature.

Another issue is mergers and acquisitions. Market participants might expect the bonds to be called due to contractual fixing of the underlying stock (target) or covenants (acquiror) after a merger announcement: Maquieira et al. (1998) find large and significant increases in convertible bond prices at announcements of stock-for-stock mergers. In addition, the largest share of influential events in King and Mauer (2012) involves merger activity. Lastly, a large share drops out of the sample due to contemporaneous news. In an earlier study, King and Mauer (2012) report that about 31% of their sample comprises influential

\textsuperscript{15}Note that the order of the nine screening criteria is irrelevant and, depending on the order, the decrease of events in every criteria is very volatile. I find only the news filter to have a significantly higher impact on the sample size than the other criteria. If all criteria are fulfilled except (4), (7) and (8), I find that 9 events drop out due to new convertibles, 17 call amounts are too low and 51 incorporate other news.

\textsuperscript{16}In the CAPM-like event study the sample diminishes due to the estimation period.

\textsuperscript{17}I refer to this effect in the hypothesis section.
events, but their focus is on call policies. The strong news cleaning conducted in this paper results in about one-half of the sample being dropped. Companies tend to release call announcements as a side note to other information on the same day. The extension to the five-day period thus does not strongly restrict the sample. Examples of such corporate events include CEO changes, earnings or dividend announcements, major law-suits and announcements of new products or joint ventures. Although the remaining Clean sample is small, 47 call events should yield enough evidence since abnormal returns are found to be large (about 1%) and highly significant in earlier literature.

II.3.3 Descriptive Statistics

Figure 6 reports the distribution of the calls across the years. The differences in time-series pattern between the Clean and Dirty sample are negligible. Interestingly, in the years following a recession, such as burst of the dot-com bubble in 2000/2001 and the recent financial crisis in 2008/2009, only very few convertibles are called. One possible explanation for this could be that managers push forward a call originally planned for later to exploit the potential over valuation during the boom period.

![Figure 6: Yearly distribution of the call announcements.](image)

Ederington and Goh (2001) report an industry clustering of convertible calls in SIC codes 35\(^{18}\) and 36\(^{19}\). Literature on convertible issues also finds higher weights in these industries: the sample of convertible debt offerings employed by Lewis et al. (2002, 2001) also has a large share in SIC codes 35 and 36, but cannot be compared directly since those authors use a different time period and exclude convertibles when another convertible was issued within the last five years. In comparison to the CRSP average, my sample confirms the higher weights in certain industries (Table 3). Since industry clustering can distort the

\(^{18}\)Industrial and Commercial Machinery and Computer Equipment

\(^{19}\)Electronic and Other Electrical Equipment and Components (Except Computer)
results of an event study, I use Fama French 12-industry indices in addition to a broad equity-market index.

**Table 3:** Descriptive statistics regarding the industry clustering/dispersion of the companies at the call announcement. The underlying companies are assigned to the corresponding 2-digit SIC code. Percentages are reported if they exceed 5% for either the Clean sample or the Dirty sample. The CRSP Average reports the average yearly representation of the corresponding 2-digit SIC code in the whole CRSP database with screening (5) for 1990 to 2011. N denotes number of events.

<table>
<thead>
<tr>
<th>2-digit SIC code</th>
<th>13</th>
<th>28</th>
<th>35</th>
<th>36</th>
<th>73</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clean Sample (%)</td>
<td>2.1</td>
<td>12.8</td>
<td>17.0</td>
<td>12.8</td>
<td>14.9</td>
<td>47</td>
</tr>
<tr>
<td>Dirty Sample (%)</td>
<td>7.3</td>
<td>17.3</td>
<td>10.7</td>
<td>12.7</td>
<td>10.0</td>
<td>150</td>
</tr>
<tr>
<td>CRSP Average (%)</td>
<td>4.9</td>
<td>8.7</td>
<td>6.9</td>
<td>10.2</td>
<td>12.7</td>
<td></td>
</tr>
</tbody>
</table>

Table 4 summarizes the characteristics of the calls for the both samples. The studied companies have a medium market capitalization compared to those listed on the NYSE. I assess the relative size of the firms by means of the Fama French 20 market equity breakpoints, which are readily obtainable from Kenneth French’s homepage. This is necessary since the sample covers a long time period in which the distribution of market capitalizations varies strongly and the size relative to the market is of interest. Although news coverage of larger companies is supposed to be greater, I find only a very small size difference between both samples. The outstanding amount called back is smaller for the Dirty sample since the Clean sample does not include clean-up calls and both distributions are highly negatively skewed. Influenced by the call amount, the Clean sample reports a higher mean dilution of about 13%. Convertible bonds are called back in-the-money after they have existed for about 60% of their maturity. Most calls are deeply in-the-money, a finding that conflicts with the optimal call polices derived by Ingersoll (1977) and Brennan and Schwartz (1977).

My sample period is unique in the literature on convertible calls, since it includes more recent years and it stems from a different data provider. To rule out potential biases induced by the period or screening criteria, I relate the call characteristics, presented in Table 4, to other literature. Since dilution is calculated relative to shares outstanding, it can measure the call’s impact on the company’s financing structure. My results from my sample are similar to those found in the literature: Brick et al. (2007) reports a slightly lower (8.9%), and Bechmann (2004) a similar, mean dilution (12.3%). Most research on convertible bonds does not provide direct information on the market capitalization of the calling or issuing companies. Converting the equity breakpoints to market capitalization, the average size of the companies in my sample is in between the sized in King and Mauer (2012) and Brick et al. (2007).

20See Bajo and Barbi (2012) for an excellent review of optimal call policies.
Table 4: Call characteristics of the convertible call events. Size denotes the market capitalization twentieth (20 highest, 1 lowest) of the calling companies at the announcement date. The breakpoints are obtained from Kenneth French’s homepage and include companies listed on the NYSE. Call Amount is denoted in million U.S. Dollar. Dilution reports the number of new shares if all bondholder convert divided by shares outstanding at the call announcement. Maturity denotes the remaining years to maturity divided by contractual maturity. Moneyness presents the conversion premium divided by call premium (nominal plus accrued interest) minus one. N denotes number of events.

<table>
<thead>
<tr>
<th></th>
<th>Clean Sample</th>
<th>Dirty Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>Size</td>
<td>12.0</td>
<td>12</td>
</tr>
<tr>
<td>Call amount</td>
<td>221</td>
<td>115</td>
</tr>
<tr>
<td>Dilution (%)</td>
<td>13.1</td>
<td>12.1</td>
</tr>
<tr>
<td>Maturity (%)</td>
<td>0.61</td>
<td>0.62</td>
</tr>
<tr>
<td>Moneyness</td>
<td>0.80</td>
<td>0.49</td>
</tr>
<tr>
<td>N</td>
<td>47</td>
<td>47</td>
</tr>
</tbody>
</table>

One important aspect of this paper is the moneyness of the conversion option. For the period 1963-1995, the moneyness in Bechmann’s (2004) sample is slightly lower (0.47) compared to Brick et al.’s (2007) sample (0.64). Bajo and Barbi (2012) employ a large sample with a similar time frame and find their in-the-money calls also exercised far from at-the-money (0.37). Although my sample is slightly deeper in-the-money, all bonds are far from at-the-money. The maximum moneyness of 4.75 is extreme in the sense that the company’s stock price must have gained about 500% in the last years. However, the maximum is still lower than the maximum reported by Brick et al. (2007) (10.0) and similar to that reported by Bechmann (2004) (4.1). The Clean sample does not incorporate clean-up calls and still exhibits a high moneyness. I thus suspect that bondholders remain optimistic about the company’s future and postpone conversion since the bond value is supposed to move equity-like.

Figure 7: Indicative plot of the average cumulative abnormal equity return in the run-up period for both samples (y-axis). $t = 0$ denotes the announcement day (x-axis). The abnormal returns are based on a market model with parameters estimated in $t \in (60, 160)$ for the CRSP value-weighted equity market index.
During the trading days prior to the call announcement, companies tend to strongly outperform the market. Figure 7 shows that the average abnormal return of my sample over the run-up period of 100 trading days is higher than 18%. The size of the abnormal return is comparable to that found by Bechmann (2004). He reports 8% in the period 50 days before the announcement; my sample exhibits an abnormal return of around 10%.

Following King and Mauer (2012), the called convertibles in my sample should be issued more equity-like than all convertibles. Managers wanting to issue equity though the backdoor set the conversion price close to the actual stock price. Indeed, the mean issuance moneyness of the Clean sample (-0.11) and the Dirty sample (-0.13) are similar and close to the -0.12 reported in Bajo and Barbi (2012). Both are significantly bigger than the mean moneyness of all convertible issuances in the whole Mergent FISD database (-0.22). To summarize, my sample is representative with regard to all publicly called convertible bonds and to most parameters and characteristics found in the literature on convertible calls. The next section presents the results of the event studies.

II.4 Event Study

II.4.1 Trading Volume

In this section I assess the validity of Hypothesis 1: in-the-money calls do not lead to abnormal trading volume. Mazzeo and Moore (1992) discover an increase in trading volume of 44% in the week around the announcement relative to an average week before the announcement. On this basis, Bechmann (2004) ascribes the abnormal returns to convertible-arbitrage hedge funds who hedge the equity risk during the conversion period by shorting shares at the time of announcement. Table 5 reports mean normalized abnormal trading volume (NATV) of this paper’s samples according to Bajo (2010):

\[ \text{NATV} = \frac{TV - \mu}{\sigma}, \]  

where \( TV \) is the natural logarithm of 1 plus traded shares on each event day, \( \mu \) the mean of a pre-event period, and \( \sigma \) the standard deviation. Compared to the trading volume during the 66 days prior to the event day, in both samples, market participants trade heavily during all three event days. In contrast, the Clean sample exhibits only borderline significant abnormal trading volume when volume measured two weeks before the call event is considered representative. For this estimation period, the shares of the companies in the Dirty sample are still comparably substantially traded. It can be inferred from Table 5 that forced conversions of convertible bonds occur during high trading volume periods and cleaning for contemporaneous news diminishes most abnormal trading volume. Indeed, the results are in line with Bechmann (2004), whose Figure 3 also reveals high trading

\[ \text{Screening criteria (1), (5) and (6) apply. See Table 9 in the Appendix.} \]
Table 5: Mean normalized abnormal trading volume around the call announcement \((t = 0)\). \(\mu\) and \(\sigma\) are calculated during 2 trading weeks or 66 trading days before the announcement. P-values in parentheses are calculated with a 2000-repetitions bootstrap routine. *, **, *** denote significance at the 10, 5, 1 percent level, respectively. \(N\) denotes the number of observations.

<table>
<thead>
<tr>
<th></th>
<th>(-3,-12)</th>
<th>(-3,-68)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Clean Sample</td>
<td>Dirty Sample</td>
</tr>
<tr>
<td>(-1)</td>
<td>0.29*</td>
<td>0.33***</td>
</tr>
<tr>
<td>(0)</td>
<td>0.41*</td>
<td>0.55***</td>
</tr>
<tr>
<td>(1)</td>
<td>0.04</td>
<td>0.32**</td>
</tr>
<tr>
<td></td>
<td>(0.84)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>(N)</td>
<td>47</td>
<td>150</td>
</tr>
</tbody>
</table>

volume in the weeks before the announcement. A counter-intuitive finding is the higher trading volume on the day before than on the day after the announcement date. If calls have information content and there are announcements after the stock markets have closed, day \(t = 1\) should yield abnormal trading volume. I attribute this result to either previous notice to bondholders or spillover effects from other news. This underlines the importance of the extended news cleaning period. Thus, price pressure has only small influence on the share price of an in-the-money-calling company, which supports my first hypothesis.

II.4.2 Stock Returns

As a first step I use a standard event-study approach to assess the validity of my second hypothesis. Several authors document that an event study on convertible calls is sensitive to estimation parameters. I address this issue by employing various event study designs. There is no consensus in the literature as to the beginning of the event period and I later show that the results differ across them. Let \(t = 0\) be the news wire announcement date of the call. The most frequently used event period is \((0, 1)\). The day after is included because information might be released after the stock market has closed. Additionally, I use \((-1, 1)\), first because there might be leakage of information and, second, sometimes bondholders receive notice of the call (by mail or other means) prior to the official announcement. Indeed, some official announcements state explicitly that a notice has been mailed to bondholders. Due to the high abnormal trading volume on the day before the announcement, I also include \((-1, 0)\) as an event window.

I estimate a market model by regressing returns \(r\) of each company \(i\) on index returns \(r_{index}\) and obtain \(\alpha\) and \(\beta\) from Equation II.2.

\[
r_{i,t} = \alpha_i + \beta_i r_{index,t} + \epsilon_{i,t} \quad \epsilon \sim N(0, 1)
\]  

(II.2)
The value-weighted CRSP (NYSE/NASDAQ/AMEX) index serves as the equity market index. To capture potential industry effects, I substitute the company-specific Fama French 12-industry indices for the market index. Because of the average market capitalization, the industry indices remain large enough not to be influenced by potential abnormal returns of the calling firms.

Cowan et al. (1990) mention that bias can arise from the choice of a pre-event estimation period. The conversion option implied in convertible bonds is mostly issued far out-of-the-money (Dann and Mikkelson, 1984; Bajo and Barbi, 2012). Thus, companies forcing conversions experience largely (abnormal) positive returns over the pre-event period. This can bias the results. Indeed, my sample reinforces this finding (indicative plot, Figure (7)).

For robustness reasons, I use two estimation windows suggested in the literature for the market model (e.g., Brick et al., 2007)): after the announcement and the conversion period: \( t \in (60, 160) \) and one year after the announcement \( t \in (251, 506) \). The selection criterion (2) from Section II.3 already covers the availability of CRSP return data.\(^{22}\)

Reported average cumulative abnormal returns (ACAR) are the cross-sectional mean of all cumulative abnormal returns over \( m \) announcement days for every company/event \( i \) of sample size \( N \) (Equation (II.3)). Since I have a small sample, I choose a bootstrap routine (see for example Davison and Hinkley, 1997) to assess the significance of the abnormal returns. Bootstrapping the abnormal returns can overcome the strict normality assumption, which can bias the significances (McWilliams and Siegel, 1997) and is especially crucial in the case of few events. I choose 2,000 repetitions, which is a conservative number that suits the methodology and reduces errors (Andrews and Buchinsky, 2001).

\[
\text{ACAR} = \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{m} r_{it} - \alpha_i - \beta_i r_{\text{index}, t}
\]  

(II.3)

Due to the post-announcement regressions for normal performance, late events drop out of the sample, particularly those occurring in the year 2011. To address this issue, I use an alternative event-study methodology based on Biktimirov et al. (2004). The methodology trades the advantage of a circumvented estimation period for missing event-specific abnormal returns. Since the methodology does not need an estimation period, recent events can be included. Rather than time-series data, a cross-sectional Fama and French (1993) factor estimation for every event day across all companies is used. For event periods longer than one day, I pool the data and create a panel. Again, I use a 2,000-repetitions bootstrap for the significances.

\[
r_{it} - RF = \alpha + \beta_1 (\text{Market} - RF) + \beta_2 \text{SMB} + \beta_3 \text{HML} + \epsilon_i
\]  

(II.4)

\(^{22}\)For the later estimation window I already include an event if the company exists in \( t \in (251, 351) \), a subset of 100 days of the full estimation period \( t \in (251, 506) \). However, results do not change for only full periods.
where $RF$ denotes the risk-free rate, $Market$ the market index, $SMB$ the small-cap factor, and $HML$ the growth factor. Furthermore, the procedure accounts for size and growth effects that might play a role in my sample: Nash et al. (2003) find that convertibles have a high share in firms with greater growth opportunities. Table 6 reports the abnormal returns

**Table 6:** Cumulative abnormal returns for the Dirty sample (Panel A) and the Clean sample (Panel B) according to Equation (II.3) and with the CRSP value-weighted index or the Fama French 12-industry classification as independent variables in Equation (II.2). Results are reported in %, for different estimation periods (columns) and different event periods (rows). Winsor reports ACARs for a winsorized (5% top/bottom) sample. *,**,*** denote significance at the 10, 5, 1 percent level obtained from a 2000-repetitions bootstrap procedure according to the p-values in parentheses. The p-values for the median are generated by a signed-rank test. $N$ denotes the number of call events.

### Panel A: Dirty Sample

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Winsor</th>
<th>Median</th>
<th>Mean</th>
<th>Winsor</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRSP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1,0)</td>
<td>-0.4</td>
<td>-0.2</td>
<td>-0.9</td>
<td>-0.3</td>
<td>-0.2</td>
<td>-0.3</td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(0.32)</td>
<td>(0.34)</td>
<td>(0.37)</td>
<td>(0.57)</td>
<td>(0.29)</td>
</tr>
<tr>
<td>(0,1)</td>
<td>-0.6*</td>
<td>-0.6**</td>
<td>-0.7***</td>
<td>-0.6</td>
<td>-0.6*</td>
<td>-0.8**</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.04)</td>
<td>(0.01)</td>
<td>(0.15)</td>
<td>(0.05)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>(1,1)</td>
<td>-0.7*</td>
<td>-0.6**</td>
<td>-0.7</td>
<td>-0.7</td>
<td>-0.6</td>
<td>-0.8**</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.06)</td>
<td>(0.03)</td>
<td>(0.14)</td>
<td>(0.12)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>FF12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1,0)</td>
<td>-0.3</td>
<td>-0.2</td>
<td>0.0</td>
<td>-0.2</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>(0.35)</td>
<td>(0.30)</td>
<td>(0.53)</td>
<td>(0.56)</td>
<td>(0.93)</td>
<td>(0.91)</td>
</tr>
<tr>
<td>(0,1)</td>
<td>-0.6</td>
<td>-0.6**</td>
<td>-0.4**</td>
<td>-0.6*</td>
<td>-0.6**</td>
<td>-0.6**</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.04)</td>
<td>(0.02)</td>
<td>(0.10)</td>
<td>(0.04)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>(1,1)</td>
<td>-0.7</td>
<td>-0.6**</td>
<td>-0.7</td>
<td>-0.7</td>
<td>-0.5</td>
<td>-0.7</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.15)</td>
<td>(0.18)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>N</td>
<td>133</td>
<td>133</td>
<td>133</td>
<td>133</td>
<td>133</td>
<td>133</td>
</tr>
</tbody>
</table>

### Panel B: Clean Sample

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Winsor</th>
<th>Median</th>
<th>Mean</th>
<th>Winsor</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRSP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1,0)</td>
<td>0.0</td>
<td>0.0</td>
<td>-0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>(0.99)</td>
<td>(0.98)</td>
<td>(0.55)</td>
<td>(0.89)</td>
<td>(0.84)</td>
<td>(0.84)</td>
</tr>
<tr>
<td>(0,1)</td>
<td>-0.3</td>
<td>-0.2</td>
<td>-0.2</td>
<td>0.0</td>
<td>0.1</td>
<td>-0.1</td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td>(0.61)</td>
<td>(0.94)</td>
<td>(0.99)</td>
<td>(0.88)</td>
<td>(0.95)</td>
</tr>
<tr>
<td>(1,1)</td>
<td>-0.6</td>
<td>-0.6</td>
<td>-0.7</td>
<td>-0.3</td>
<td>-0.2</td>
<td>-0.2</td>
</tr>
<tr>
<td></td>
<td>(0.32)</td>
<td>(0.25)</td>
<td>(0.20)</td>
<td>(0.68)</td>
<td>(0.66)</td>
<td>(0.57)</td>
</tr>
<tr>
<td>FF12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1,0)</td>
<td>0.3</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>(0.53)</td>
<td>(0.69)</td>
<td>(0.72)</td>
<td>(0.49)</td>
<td>(0.57)</td>
<td>(0.55)</td>
</tr>
<tr>
<td>(0,1)</td>
<td>0.0</td>
<td>-0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>-0.1</td>
</tr>
<tr>
<td></td>
<td>(0.97)</td>
<td>(0.85)</td>
<td>(0.84)</td>
<td>(0.74)</td>
<td>(0.84)</td>
<td>(0.91)</td>
</tr>
<tr>
<td>(1,1)</td>
<td>-0.2</td>
<td>-0.3</td>
<td>-0.3</td>
<td>0.0</td>
<td>0.1</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>(0.77)</td>
<td>(0.60)</td>
<td>(0.72)</td>
<td>(0.98)</td>
<td>(0.82)</td>
<td>(0.83)</td>
</tr>
<tr>
<td>N</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>45</td>
</tr>
</tbody>
</table>
for both samples, different event windows, and for varied estimation periods using the CRSP value-weighted index and the Fama French 12-industry classification.

The mean abnormal returns of the Dirty sample in Panel A of Table 6 are consistent with the literature. For example, Brick et al. (2007) report mean abnormal returns of -0.75%, Bechmann (2004) -1.75%, and García-Feijóo et al. (2010) -1.07%. The results for the median and the winsorized sample reveal no large outliers as well as a symmetric distribution of the abnormal returns. The three-day event period yields the largest and, despite the high trading volume, (−1, 0) the lowest abnormal returns. While most medians remain significantly negative, the industry indices increase the variance of the abnormal returns and lead to borderline significant means. The significances are lower than those found in the literature. All three papers discussed above find their mean abnormal returns significant at the 1% level. I attribute the lower significances of the Dirty sample to different contemporaneous news announcements. As I point out in Section II.3, many call announcements are side notes to other important company news, often earnings reports. As my sample covers a different period than does previous literature, the news is also different. For example, earnings reports contemporaneously announced with the calls could be more positive or announced dividends smaller.

The main contribution of this paper can be found in Panel B of Table 6, the abnormal returns of the Clean sample. Pure, non-disturbed call announcements of in-the-money convertible bonds do not lead to significant abnormal returns. Mean, median, and winsorized samples are far from any common significance level. Although my sample size of 45 calls is small, it should reveal at least a tendency toward negative abnormal returns if the results in previous literature are accurate. However, some mean abnormal returns are actually positive and two specifications reveal more positive than negative abnormal returns. Again the industry indices increase most p-values even more.

In line with previous literature, I find that the abnormal returns vary with estimation specification. Estimation period, event period, and index all have an influence on the abnormal returns. Unreported estimations with non-bootstrapped standard errors come to the same results. Table 7, the cross-sectional event study results, supports the previous findings. For the Clean sample, all alphas are far from any common significance level. The fit of the regressions, measured by the adjusted $R^2$, is negative for $t = 0$ compared to 20% in $t = −1$ and $t = 1$. This could mean that the call announcement slightly influences the share price, but these influences are positive rather than negative and far from any common significance level. The Dirty sample shows negative alphas not only for the pooled regressions but also for the announcement day regression. Except for the market factor, all Fama-French factors are insignificant and the signs of the loadings change. García-Feijóo et al. (2010) employ a similar methodology and find significantly negative abnormal returns. Combining the results of the different event studies, I conclude that my hypothesis is confirmed for this data set; there are no significant abnormal returns for forced-conversion
convertible calls. In the next section, I test whether the abnormal returns are statistically different and comment on the screening criteria.

II.4.3 Explanatory Regressions

I assess the statistical difference between the samples by regression of abnormal returns (CAR) including several literature-based explanatory variables:

\[
CAR = \alpha + \beta_1 \text{Dirty} + \beta_2 \text{Size} + \beta_3 \text{Liquidity} + \beta_4 \text{Overvaluation} + \epsilon \quad \text{(II.5)}
\]

The dependent variable is the Dirty sample which includes the Clean sample as a subset. Dirty is a dummy variable that takes the value 1 if the event is not in the Clean sample. If the abnormal returns differ between the Clean and Dirty samples, I expect this variable to

Table 7: This table reports regression results based on Equation (II.4) for the Dirty sample (Panel A) and the Clean sample (Panel B). Results report the factor loadings and p-values. The columns denote the different days or periods around the announcement \( t = 0 \). Column 5 and 6 report results for pooled data. P-values in parentheses are derived through a 2000-repetitions bootstrap routine. *, **, *** denote significance at the 10, 5, 1 percent level, respectively. N denotes the number of data points in the regression.

<table>
<thead>
<tr>
<th></th>
<th>Panel A: Dirty Sample</th>
<th></th>
<th>Panel B: Clean Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(-1)</td>
<td>(0)</td>
<td>(-1)</td>
</tr>
<tr>
<td>( \alpha ) (%)</td>
<td>-0.3</td>
<td>-0.5*</td>
<td>-0.1</td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
<td>(0.06)</td>
<td>(0.59)</td>
</tr>
<tr>
<td>Market-RF</td>
<td>1.79***</td>
<td>0.90***</td>
<td>1.11***</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>SMB</td>
<td>1.00**</td>
<td>0.25</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.72)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>HML</td>
<td>0.68</td>
<td>0.40</td>
<td>-1.01**</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.50)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>adj. ( R^2 )</td>
<td>0.20</td>
<td>0.06</td>
<td>0.30</td>
</tr>
<tr>
<td>N</td>
<td>150</td>
<td>150</td>
<td>150</td>
</tr>
</tbody>
</table>

adj. \( R^2 \)   | 0.21                  | -0.01            | 0.20                  | 0.11                 |
| N                | 47                    | 47               | 47                    | 94                   |

The dependent variable is the Dirty sample which includes the Clean sample as a subset. Dirty is a dummy variable that takes the value 1 if the event is not in the Clean sample. If the abnormal returns differ between the Clean and Dirty samples, I expect this variable to...
be significantly negative. I follow Brick et al. (2007) and assess the Size of the company by the logarithmic market capitalization at the announcement date $t = 0$. However, the results are similar when call amount divided by market capitalization is used instead. Size is intended to capture analyst coverage; thus abnormal returns for larger companies are expected to be closer to zero. To assess potential market micro-structure effects I employ the liquidity measure of Amihud (2002):

$$\text{Liquidity} = 10^6 \cdot \sum_{i=-2}^{-251} \frac{|\text{Return}_i|}{\text{Closing Price}_i \cdot \text{Trading Volume}_i}$$  \hspace{1cm} (II.6)

If the price pressure hypothesis is correct, I expect the variable Liquidity to have a negative impact on the abnormal returns. If the asymmetric information hypothesis holds true, firms call their convertible bonds when managers consider the stock overvalued. Overvaluation is the only significant and partly robust independent variable in Brick et al. (2007). I thus take my lead from those authors and estimate cumulative abnormal returns between $t \in (-60, -2)$ according to Equation (II.2) and the Fama French 12-industry indices estimated between $t \in (60, 160)$. I expect a negative sign.

Variables for moneyness of the conversion option and dividend dilution are excluded from the regressions. Since nearly all calls are made deeply in-the-money, I drop the moneyness variable. Nash et al. (2003) report that firms with high investment opportunities are more likely to issue convertible bonds. If firms are engaged in valuable investment projects, they are less likely to pay shareholder dividends. Indeed, very few companies paid a dividend during the one year before the call announcement. For this reason, dividend dilution cannot be of concern to current shareholders.
Table 8: This table reports regression results based on Equation (II.5) for the Clean sample (Panel A) and the Dirty sample (Panel B). Results report the factor loadings and p-values. The columns denote the different days or periods around the announcement \( t = 0 \). Independent variables are a dummy for the Dirty sample (Dirty), a liquidity measure (Liquidity), logarithmic market capitalization (Size) and a proxy for Overvaluation. P-values in parentheses are derived through a 2000-repetitions bootstrap routine. *, **, *** denote significance at the 10, 5, 1 percent level, respectively. \( N \) denotes the number of data points in the regression.

**Panel A: FF12, (60,160)**

<table>
<thead>
<tr>
<th></th>
<th>((-1))</th>
<th>((0))</th>
<th>((1))</th>
<th>((-1,0))</th>
<th>((0,1))</th>
<th>((-1,1))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>-0.010</td>
<td>0.014</td>
<td>-0.021</td>
<td>0.004</td>
<td>-0.007</td>
<td>-0.017</td>
</tr>
<tr>
<td>(0.76)</td>
<td>(0.70)</td>
<td>(0.70)</td>
<td>(0.93)</td>
<td>(0.92)</td>
<td>(0.83)</td>
<td></td>
</tr>
<tr>
<td>Dirty</td>
<td>0.001</td>
<td>-0.012**</td>
<td>0.003</td>
<td>-0.011</td>
<td>-0.009</td>
<td>-0.008</td>
</tr>
<tr>
<td>(0.87)</td>
<td>(0.04)</td>
<td>(0.23)</td>
<td>(0.15)</td>
<td>(0.27)</td>
<td>(0.40)</td>
<td></td>
</tr>
<tr>
<td>Liquidity</td>
<td>0.004</td>
<td>-0.067</td>
<td>0.015</td>
<td>-0.063</td>
<td>-0.052</td>
<td>-0.048</td>
</tr>
<tr>
<td>(0.99)</td>
<td>(0.84)</td>
<td>(0.99)</td>
<td>(0.92)</td>
<td>(0.97)</td>
<td>(0.97)</td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>(0.77)</td>
<td>(0.76)</td>
<td>(0.77)</td>
<td>(0.97)</td>
<td>(0.94)</td>
<td>(0.86)</td>
<td></td>
</tr>
<tr>
<td>Overvaluation</td>
<td>-0.006</td>
<td>0.017</td>
<td>0.002</td>
<td>0.012</td>
<td>0.019</td>
<td>0.013</td>
</tr>
<tr>
<td>(0.50)</td>
<td>(0.24)</td>
<td>(0.83)</td>
<td>(0.43)</td>
<td>(0.31)</td>
<td>(0.48)</td>
<td></td>
</tr>
<tr>
<td>adj. ( R^2 )</td>
<td>-0.03</td>
<td>0.03</td>
<td>-0.03</td>
<td>-0.01</td>
<td>-0.00</td>
<td>-0.02</td>
</tr>
<tr>
<td>( N )</td>
<td>133</td>
<td>133</td>
<td>133</td>
<td>133</td>
<td>133</td>
<td>133</td>
</tr>
</tbody>
</table>

**Panel B: FF12, (251,506)**

<table>
<thead>
<tr>
<th></th>
<th>((-1))</th>
<th>((0))</th>
<th>((1))</th>
<th>((-1,0))</th>
<th>((0,1))</th>
<th>((-1,1))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>-0.001</td>
<td>-0.003</td>
<td>0.003</td>
<td>-0.004</td>
<td>0.000</td>
<td>-0.001</td>
</tr>
<tr>
<td>(0.98)</td>
<td>(0.94)</td>
<td>(0.95)</td>
<td>(0.95)</td>
<td>(0.95)</td>
<td>(0.99)</td>
<td></td>
</tr>
<tr>
<td>Dirty</td>
<td>0.002</td>
<td>-0.011*</td>
<td>0.000</td>
<td>-0.009</td>
<td>-0.011</td>
<td>-0.009</td>
</tr>
<tr>
<td>(0.78)</td>
<td>(0.07)</td>
<td>(0.94)</td>
<td>(0.22)</td>
<td>(0.14)</td>
<td>(0.28)</td>
<td></td>
</tr>
<tr>
<td>Liquidity</td>
<td>-0.002</td>
<td>-0.000</td>
<td>0.024</td>
<td>-0.049</td>
<td>-0.024</td>
<td>-0.025</td>
</tr>
<tr>
<td>(0.99)</td>
<td>(0.86)</td>
<td>(0.98)</td>
<td>(0.95)</td>
<td>(0.98)</td>
<td>(0.98)</td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>(0.99)</td>
<td>(0.76)</td>
<td>(0.89)</td>
<td>(0.90)</td>
<td>(0.94)</td>
<td>(0.99)</td>
<td></td>
</tr>
<tr>
<td>Overvaluation</td>
<td>-0.008</td>
<td>0.013</td>
<td>-0.004</td>
<td>0.005</td>
<td>0.010</td>
<td>0.002</td>
</tr>
<tr>
<td>(0.40)</td>
<td>(0.36)</td>
<td>(0.52)</td>
<td>(0.73)</td>
<td>(0.57)</td>
<td>(0.93)</td>
<td></td>
</tr>
<tr>
<td>adj. ( R^2 )</td>
<td>-0.03</td>
<td>0.01</td>
<td>-0.03</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-0.02</td>
</tr>
<tr>
<td>( N )</td>
<td>133</td>
<td>133</td>
<td>133</td>
<td>133</td>
<td>133</td>
<td>133</td>
</tr>
</tbody>
</table>

Table 8 reports results for the industry-indices based CARs; the results are similar for CRSP-index based CARs. The only significant variable is the Dirty sample dummy, yielding a difference between the samples due to amplified screening criteria at the announcement day. This seems to drive the multi-day event periods Dirty variable. All other independent variables are either small or change the sign, and are far from any common significance level, which is consistent with earlier literature. Also, the low fit of the regressions is in line with literature containing explanatory regressions of cumulative abnormal returns. I now try to answer the question which screening criteria are responsible for the difference
between the samples? Since I find the 30 strongest negative abnormal returns to be robust to estimation event window and setup, I evaluate the specific screening criteria that resulted in them being put into one or the other of the samples. If the event is screened, it is either due to delisting/bankruptcy, exchanged bonds, clean-up calls, or other news (criteria (3), (4), (7), or (8)). Out of these 30 abnormal returns, ten can be attributed to other news such as CEO changes or joint ventures and five are contemporaneous to earnings announcements. Five bonds are exchanged, seven events are clean-up calls, and two calls are merger related. I conclude that it is not one specific event that is driving the differences between the samples, but my empirical evidence suggests that it is rather the complete set of contemporaneous events.

II.5 Conclusion

This paper sheds new light on the announcement effects of forced conversions of convertible bonds. In stark contrast with previous work, I find no significant impact from in-the-money calls on stock prices, a finding that is robust to different event study techniques and parameters. Previous literature provides information- and liquidity-based explanations for negative abnormal returns. The result of this paper make me respectfully disagree. First, convertible bonds that undergo a forced conversion are mostly issued more equity-like than debt-like. A conversion is in line with convertible bond’s backdoor-equity purpose. Second, the average conversion option at the announcement date is deeply in-the-money. Market participants expect the bonds to be converted and thus all relevant information is already priced. Third, hedging demand at announcement date should be low due to hedge funds’ buy-and-hold arbitrage strategies. These funds buy and hedge large shares of the bonds at issuance. I find only low abnormal trading volume in a screened sample, in contrast to a sample with less restrictive selection criteria.

The sample with less restrictive screening criteria has (partly) significantly negative abnormal returns on the announcement day in contrast to a cleaned sample where abnormal returns are not distinguishable from zero. The difference between both samples is partly significant. In assessing what is behind the differences in the two samples, I find the occurrence of corporate news at the same time as the call announcement to be very relevant. Companies tend to announce a call of a convertible and also other firm-specific news on the same day. Such news often involves earnings results or bond exchanges. These are high-impact and well-cited events and overshadow any potential convertible call effect.

\[23\text{ Another regression setup would be to introduce a dummy variable for every contemporary news and screening criteria, but this would diminish the degrees of freedom to a large extent since the set of news is very heterogeneous.}\]

\[24\text{ An event can be in more than one category. Merger related calls do not count as other news.}\]
II.6 References


II.7 Further Results

Table 9: Moneyness at issuance of the implied conversion option for different data sets. Mergent FISD stands for the complete database until 2011 with screening criteria (1), (5) and (6). Differences in mean with respect to the Mergent FISD dataset are derived through a standard t-test, t-stats in brackets. (*,**,***) denote significance at the 10, 5, 1 percent level. N denotes the number of observations.

<table>
<thead>
<tr>
<th></th>
<th>Clean Sample</th>
<th>Dirty Sample</th>
<th>Mergent FISD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>Mean</td>
</tr>
<tr>
<td>Moneyness at issuance</td>
<td>−0.11</td>
<td>−0.17</td>
<td>−0.13</td>
</tr>
<tr>
<td>Difference</td>
<td>−0.11***</td>
<td>[2.91]</td>
<td>−0.10***</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>47</td>
<td>175</td>
</tr>
</tbody>
</table>

Figure 8: The figure displays the mean of the company-specific daily trading volumes scaled by the company-specific mean trading volumes during the plotted period $t \in (-50, 50)$ for the screened sample (black line) and the complete sample (grey line). $t = 0$ denotes the announcement date. With constant trading volume the lines would be flat at 1. Figure according to Bechmann (2004).
Abstract
This paper investigates the contractual risk sharing as it can be observed in project finance. I argue that lenders have a strong influence on the shape of the firm because the financing contribution of the sponsoring firm is low and debt ownership is concentrated among few parties. However, this does not fully explain the large wealth expropriation from equity holders as implicated by a decrease in asset volatility in standard models of the firm. I propose a simple, but powerful extension of standard real option models to failure risk that can explain the low residual risks in project finance. Failure risk can lead to a total loss of a company’s assets-in-place, a situation that most of contingent claim models ignore and which cannot be captured by, for example, jumps in the cash-flow process. The concept of failure risk is regularly observable in practice, but is especially pronounced in project finance. Abandoned projects, catastrophes, and political instability are just a few examples of how the claims of lenders and equity holders can be rendered worthless. Calibrated to reasonable parameters, this model shows that it can be beneficial for equity holders to contract away asset volatility.
III.1 Introduction

Boundaries of the firm, especially highly leveraged firms, and their determinants attract a great deal of attention from researchers. Outstanding in this regard are project firms, whose average leverage ratios\(^{25}\) are 70% and above (Esty, 2002; Byoun et al., 2012) and whose syndicated loans account for almost 10% of the worldwide syndicated loan market (Bonetti et al., 2010). Theoretical research in this field aims at understanding corporate decisions to separate the growth options from the parent company into a project firm, or to explain the origin of the distinct characteristic of strong debt financing. Yet, one important question that is usually ignored is why managers of project firms contract away asset risks when they should be acting in the interests of equity holders. This paper proposes a simple, but powerful extension of classical contingent claims models that introduces another source of risk—failure risk—and evaluates it from the project finance perspective.

A ”project finance” firm was originally defined to mean a new, legally independent company with one sole purpose, one corporate activity.\(^{26}\) However, project finance is often referred to as contractual finance because these firms employ a net of contracts to delegate risk. Typically, a separate project company negotiates with all prospective contractual partners to delegate risk to institutions with the potential to bear it. Examples include investment banks insuring exchange-rate risks and development banks carrying political risks. Separating the individual risks is often not possible, however, which leads to a multidimensional optimization problem for the project firm. Case studies provide evidence that managers do indeed try to contract away almost all risks. For example, Bonetti et al. (2010) find that the credit risk of the project firm is closely related to the off-taker’s equity volatility.

Insuring risks is costly. The project company either has to make monetary considerations or transfer control rights in return. Kleimeier and Megginson (1999) provide considerable evidence that even though project firms are exposed to potentially higher bankruptcy risk due to low financing contributions by the parent firm, they do not compensate the lenders with increased spreads. Thus, control benefits have to be transferred to lenders. Indeed, Byoun et al. (2012) argue that project firms that use high leverage ratios give up control to the concentrated debt holders. Project company debt is mostly nonpublic and divided among just a few creditors. This situation concentrates a high amount of liability in a few parties who have a strong interest in closely monitoring the firm and gaining control over project accounts, (Subramanian et al., 2008). The authors further find that cash flow in project firms is verifiable through unique firm characteristics and that lenders gain control over project cash flow via a net of project accounts.

Consequently, classical theories of the firm do not hold for project finance firms. In project

\(^{25}\)Debt to total capitalization.

\(^{26}\)Note that project finance cannot be directly compared to leveraged buyouts since it has a low impact on managers due to the fact that the project finance company has no history, past profitability, or previous strategic commitments. (Esty, 2004)
finance, lenders still have bankruptcy priority, but managers act not only in the interests of shareholders, but also take lenders’ priorities into consideration. Lenders gain control, influence the company in their favor, and negotiate with the managers over firm structure. However, this does not explain the low residual risks found in project finance. Standard theoretical models imply that equity holders are interested in increasing asset volatility because they hold a call option on the asset value of the company (Merton, 1974). Although equity holders may be willing to partially surrender risk for the lenders, their losses from average project finance contracting would be extensive. I extend a classical real option model by including failure risk and show that equity holders do indeed have an incentive to contract away standard cash-flow risks when it interferes with failure risk. Failure risk is the probability that the firm will lose all of its assets, a consequence of which is that all claims of stakeholders will be worthless.

The loss of a total project company or a branch of a corporation can regularly be observed in practice but predominates in project finance firms. The failure of a firm is different from bankruptcy risk, since no residual claim, shared among the bondholders is left and failure can occur independent of the cash-flow stream. Below are four examples of this sort of failure:

1. Products can suddenly become obsolete due to innovation or new technology. This will not only affect the cash-flow stream, but can also render the assets-in-place worthless if, for example, specialized production machinery becomes outdated. 2. A firm’s assets can be lost due to political instability or the impossibility of investors withdrawing or liquidating them. Consider, for example, the outbreak of a war or, for a real-life example, the recent expropriation of private mining and electrical supply firms in Bolivia. A government decision to nationalize a firm results in an immediate and total loss for all investors. This is an especial concern for project finance firms since this type of firm is relatively frequently established in high-risk countries (Byoun et al., 2012). 3. A project may be aborted during construction before generating the expected final assets-in-place. In their sample of large engineering projects, Lessard and Miller (2001) find that about 10% are abandoned. The authors claim that projects are like high-stakes games and go on to note that the likelihood of abortion during the development period is extremely high. Lessard and Miller (2001) also claim that certain types of assets, especially those involved in engineering projects, are not re-deployable and that, moreover, infrastructure project assets generally cannot be broken up into separate parts and used in other projects, e.g., a half-done roads can hardly be moved to another economic environment. 4. A catastrophe such as a hurricane, earthquake, or tornado can destroy an entire production facility. The importance of planning for catastrophes in corporate financing is evidenced by the increasing volume of catastrophe bonds (CAT) being issued. For example, the East Japan Railway Cooperation has issued CAT bonds to hedge part of the high earthquake risk predominant in Japan.

Empirical evidence shows that contracting reduces cash-flow as well as failure risk, mostly
with the same contract. Corielli et al. (2010) highlight that the nexus of contracts serves as a type of risk management. In about 19% of projects, at least one contract ensures that a large quantity of the new firm’s output will be bought by another firm or by the public sector for at least a limited time (Byoun et al., 2012). These unconditional and obligatory contracts of sale dramatically decrease a firm’s short-term cash-flow volatility and can lower the financing costs of generating firm assets. Empirical literature and case studies show that project companies use many more risk-decreasing and thus typically ex ante value-enhancing contracts. Banks, for example, are used to mitigate exchange risk; they guarantee proceeds and loan payments in the same currency (Bonetti et al., 2010), which smooths the revenues and increases the odds for the survival of the firm. These project loans are found to be more long-term than standard corporate loans (Sorge and Gadanez, 2008). Additionally, project companies use banks for inflation indexing (Bonetti et al., 2010), which insures against political risk and decreases bankruptcy risk. If the location of the project is politically unstable, political risk is reduced by development bank and government intervention or guarantees (Hainz and Kleimeier, 2012). Moreover, although projects are generally large, they use approved, reliable technology (Brealey et al., 1996), which helps ensure the project’s completion and a risk-free production process. Valuing these effects requires agents to recognize the structure of the project. Indeed, in a case study with a publicly traded debt instrument, Dailami and Hauswald (2007) find that market participants are aware of the contracting firm’s high dependency on the off-takers credit worthiness.

Given the size of the project finance market and appealing features of project financing compared to standard corporate finance, there is an abundant theoretical literature on this topic. These appealing features include the separate corporate structure of the growth option, which avoids inefficient internal investment (Scharfstein and Stein, 2000) and cross-subsidization, and, similar to securitization, project finance debt has limited recourse to the option exercising firm’s assets. Indeed, much of corporate finance theory is inapplicable to project finance firms. For example, at exercise of the growth option, the new company has no assets-in-place, hence asymmetric information problems between managers and investors (pecking order theory, Myers and Majluf, 1984) and ex ante shareholder-bondholder conflicts (Jensen and Meckling, 1976) are negligible. Additionally, few growth options of the project firm result in only minor underinvestment problems (Byoun et al., 2012; Myers, 1977).

My model assumes that the decision to separate the growth option has been made and that the necessary equity funds have been separated from the parent company’s equity. At this point, managers only have to decide when to exercise the growth option and establish the new, leveraged project firm and decide how much of each risk – cash flow and failure – should be contracted away. By calibrating my model with realistic parameters, I can show

---

27In the 2004 U.S. market, project investments were half the size of the total IPO market (Byoun et al., 2012).
effects of failure risk on the firm value and explain the observable high leverage ratios of project companies. However, the model is not intended to provide optimal contracting/risk values since the dependencies between, e.g., leverage and debt holder control, are not known, and the influence of a specific contract on both risks is potentially individual to each project. I find that by contracting out risk in a high-risk project finance environment, a firm’s value can increase by 120% compared to a case of no contracting. Additionally, the model allows accounting for a limited lifetime of the new company without introducing time-dependent parameters.

Previous literature uses discrete and continuous models to assess the nature of project firms. The earliest model is from Shah and Thakor (1987), who find that, under the assumption of costly information about project riskiness, projects are more highly leveraged when information production is more efficient. Berkovitch and Kim (1990) also provide an asymmetric information model for optimal ex ante financial contracting of a new, debt-financed project. Over- and underinvestment incentives are tested in the event the new debt has seniority or there is a dividend policy. Under symmetric information between shareholders and bondholders regarding the payoff of the new project, project finance is the optimal ex ante rule to mitigate over- and under-investment. A complete separation will be optimal if there is no synergy between the new project and existing assets. Agency costs due to debt financing trade off against tax shields in John and John (1991). Project finance is valuable when the two projects are far apart in growth options. Projects using known technology, with low managerial discretion, and tangible assets will be allocated a high proportion of debt, whereas those using new technology, for which managers are given a great deal of discretion, and that have few tangible assets will be allocated a low proportion of debt. Chemmanur and John (1996) propose a model for valuable (management bonus, reputation effects) corporate control. In their model, the equilibrium choices of an entrepreneur depend on organizational form, capital structure, and ownership structure. The authors imply that if the new project has much lower control benefits than the parent firm, the project will have much higher debt ratio. Habib and Johnsen (1999) argue that non-recourse debt ensures optimal allocation of ownership over specific assets. Ex ante, before either the good or bad state of the economy is revealed, non-recourse debt contracts can maximize the overall returns, which is not possible ex post. Gorton and Souleles (2007) employ a moral hazard model where the bank can decide between on- or off-balance sheet financing before and after the quality of the assets has been revealed. The multi-period model assumes that ethical decisions must be made repeatedly and that investors will punish the company by not investing if the company does not subsidize failed special purpose vehicles: securitization is undertaken to avoid bankruptcy costs. The two-period model of Leland (2007) explores the financial benefits of a single or merged capital structure in case of non-synergistic cash flows. The chief benefits of separate financing stem from the possibility of limited liability shelter, and separate financing allows for higher debt. The
The extant model most similar to mine is that of Hackbarth et al. (2012). They employ a similar, continuous, real-option model that answers the question of whether a growth option should be integrated into or separated from the original firm. Their model examines the interaction between debt financing and new product market opportunities. The two opposing factors are market timing (competitors can be quicker) and cannibalizing profits due to the new company, affecting the growth option and the assets-in-place. The model also includes tax effects, bankruptcy costs, and size of the growth option. The authors find that higher cash-flow risk favors non-integration, while higher obsolescence risk favors integration. An increase in tax benefits makes non-integration more likely.

The above-cited project-finance-related models either provide reasons to separate a project based on information asymmetries and/or agency conflicts or answer the question of whether integration or separation is optimal. However, the project-finance-specific contracting or risk sharing that changes the shape of the new firm after a decision to separate has been made has not yet been the subject of a thorough theoretical investigation. This paper is a start in that direction and the remainder of it is organized as follows. Section III.2 presents the model for exploring contracting effects in project finance. Section III.3 discusses how the model is solved. Calibrating the model to observable parameters in Section III.4 yields the results presented in Section III.5. Section III.6 concludes the paper.

### III.2 The Project Finance Model

My model assumes that the parent company has already decided to separate the growth option instead of integrating it into its business. Thus, I omit modeling the parent company, since the separated company has a limited liability shelter and the benefits from diversification should be low because both companies typically operate in the same industry.\(^{28}\) I employ a standard contingent claim approach to dynamically model the real option. Before exercise, the company consists solely of equity provided to execute the growth option. Since I assume risk-neutral agents, the equity must produce capital gains at the risk-free rate \(r > 0\). The equity claim follows a geometric Brownian motion \(dX = \mu_B X dt + \sigma_B X dW_t\), \(\mu_B > 0\), \(\sigma_B > 0\), with initial value \(X_0 > 0\). The company has a growth option that can be exercised for constant investment costs of \(\kappa > 0\) at any time during its infinite lifetime. After exercise, the project company begins to operate.

After exercising the growth option, the company continues in existence infinitely and receives stochastic earnings before interest and tax (EBIT)\(^{29}\) of \(\pi Z_t\), \(\pi > 0\), where \(Z_t\) follows the geometric Brownian motion \(dZ = \mu Z dt + \sigma Z dW_t\), with \(\mu > 0\) and \(\sigma > 0\). This implicitly assumes that the parent company’s equity and the new firm’s cash flow remain

\(^{28}\)Case study evidence by Dailami and Hauswald (2007) and Bonetti et al. (2010).

\(^{29}\)It is standard to model a cash-flow stream. However, for this model, it is important to define the cash-flow stream as earnings since this ensures a link between contracting to reduce the volatility of expenses and the volatility of the cash-flow stream. Nevertheless, I refer to it as “cash flow” in the remainder of the paper.
perfectly correlated (an assumption that easily can be relaxed), but that they operate in two different markets. The initial value is bound to the value of the equity claim process at exercise time $\tau_E$, $X_{\tau_E} = Z_{\tau_E}$.

After exercise, the new company faces two sources of risk. First, earnings volatility can lead to a situation where it is optimal for equity holders to declare bankruptcy. The threshold $Z$ is ex ante set to maximize equity value and, when the threshold is reached, manifests as loss of tax benefits and results in a constant earnings level, $(1 - \alpha)\pi(1 - \tau)Z$, where $\alpha > 0$ is a reduction factor and $\tau > 0$ is the constant, overall tax rate. The bankruptcy shelter allows the parent company’s assets to remain untouched. The second risk, should it come to pass, will render the entire firm worthless: both equity holders and debt holders will lose their claims. I assume that for each time interval $dt$, the probability of failure is constant, with size $\psi$. As noted in the introduction to this paper, this factor is found in the real world, but cannot be captured by standard models via modifications in the diffusion process, such as or jumps in drift or volatility, which typically increase the overall volatility and thus increase the value of the equity claims. Additionally, these modifications do not affect all the assets-in-place. Note that the funds remaining after failure can be set to an arbitrary percentage of the original firm value. Further note that failure risk is not present before the option is exercised. This is reasonable since the separated equity will not lose its value. If it is not used to finance the new company, the parent firm will use it for other purposes such as other investments independent of this particular growth option. Thus, bankruptcy risk is not independent of failure risk, but failure risk is independent of bankruptcy. That is, even if the company is bankrupt, it can fail, but if it has failed, its assets are worthless and cannot be redeployed to fulfill bankruptcy obligations. The construction costs are financed by debt and equity. Leverage is introduced by issuing a perpetual coupon $C$ at time $t = \tau_E$. The capital structure is set to maximize the firm’s value. Further, the model is based on typical assumptions such as that agents are risk neutral, that the risk-free rate is strictly greater than the drift of both state variables ($r > \mu_B, r > \mu$), that all cash flows are taxed by rate $\tau$, and that debt payments can be deducted from taxes. Control rights are valuable. Although the parent company provides the equity and is the sole owner of the newly established company, I explicitly assume that the new firm is self-sufficient. It fixes its capital structure and can independently declare bankruptcy. Further, it does not face any agency effects of debt, since it has no assets-in-place before exercise. I use this model to investigate the effects of reducing two types of risk – cash flow and failure – through contracting, thus further linking the model to project finance.

The following notation is employed in the remainder of the paper: subscript $B$ denotes values before exercise of the growth option, and $V$, $\Xi$, and $D$ denote firm value, equity value, and debt value, respectively. I suppress indication of before or after exercise since the dependency on state variables $X$ and $Z$ reveals that information.

---

30 For literature on jumps in real options, see, e.g., Martzoukos and Trigeorgis (2002).
III.3 Solving the Model

In this section I show how I derive the equations for the model backward through time, starting with the value after exercise.

III.3.1 The Option

Firm value, equity value, and coupon for the new company will depend on state variable $Z$, which starts when the option is exercised at stopping time $\tau_E$. The cash flow follows the diffusion process:

$$dZ = \mu Z dt + \sigma Z dW_t, \quad Z_{\tau_E} = X_{\tau_E} \quad (III.1)$$

Since the agents in the market are risk neutral, the value of equity and of the firm over time evolve as:

$$rV(Z) dt = E[dV(Z)] + [-\psi V + (1 - \tau)\pi Z + \tau C] dt \quad (III.2)$$

$$r\Xi(Z) dt = E[d\Xi(Z)] + [-\psi \Xi + (1 - \tau)\pi Z + (\tau - 1)C] dt \quad (III.3)$$

$$rD(Z) dt = E[dD(Z)] + [-\psi D + C] dt \quad (III.4)$$

The firm receives the taxed, relative fraction of the state variable $Z$ and, in the case of firm value, the tax-deductible coupon, whereas for equity value, coupon payments are subtracted. Expected capital gains account for the behavior of the process $Z$. Now I assume that the company sets its perpetual coupon and bankruptcy threshold to maximize firm value and equity value, respectively. The derivations are fairly standard; for a short description of the individual steps, the reader is referred to the Appendix III.9. As a result of the derivations, I obtain the firm value $V(Z)$, equity value $\Xi(Z)$, debt value $D(Z)$, optimal coupon $C$, and optimal bankruptcy threshold $Z$ after exercise for the initial state variable level $Z$.

$$V(Z) = \pi \Lambda Z + \frac{\tau C}{r + \psi} \left(1 - \left(\frac{Z}{\bar{Z}}\right)^\vartheta\right) - \alpha \pi \Lambda Z \left(\frac{Z}{\bar{Z}}\right)^\vartheta \quad (III.5)$$

$$\Xi(Z) = \pi \Lambda Z - \frac{(1 - \tau)C}{r + \psi} - \left(\pi \Lambda Z - \frac{(1 - \tau)C}{r + \psi}\right)\left(\frac{Z}{\bar{Z}}\right)^\vartheta \quad (III.6)$$

$$D(Z) = \frac{C}{r + \psi} + \left(1 - \alpha\right)\Lambda \pi Z - \frac{C}{r + \psi} \left(\frac{Z}{\bar{Z}}\right)^\vartheta \quad (III.7)$$

where

$$Z = \frac{\vartheta}{\vartheta - 1} \frac{r + \psi - \mu C}{r + \psi}, \quad \vartheta = \frac{1}{2} - \frac{\mu}{\sigma^2} - \sqrt{\left(\frac{\mu}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2(r + \psi)}{\sigma^2}},$$

$$\Lambda = \frac{1 - \tau}{r + \psi - \mu} \quad (III.8)$$
and
\[ C = \pi \frac{\vartheta - 1}{\vartheta} \frac{r + \psi}{r + \psi - \mu} \left( 1 - \vartheta \left( 1 - \alpha + \frac{\alpha}{\tau} \right) \right)^{\frac{1}{\vartheta}}. \] \tag{III.9}

### III.3.2 Survival Properties

The probability of failure for every time interval \( dt \) is \( \psi \).

If I now set the stopping time \( \theta \) to be the time the company fails and set the following indicator variable
\[ Y_t = \begin{cases} 1 & t \geq \theta \\ 0 & t < \theta \end{cases}, \]
then the probability of survival \( Q(t) \) at time \( t \) follows the exponential law:
\[ dQ = -\psi Qdt \quad \Rightarrow \quad Q(t) = \exp(-\psi t) = P(Y_t = 0) \]

I now obtain the following for the expected time to failure, which is independent of the time \( t \).
\[ E[\theta_t] = \int_t^\infty \psi(u - t) \exp(-\psi(u - t))du \]
\[ = \frac{1}{\psi} \left[ -\exp(-\psi(u - t)) - \psi(u - t) \exp(-\psi(u - t)) \right]_t^\infty = \frac{1}{\psi} \]

### III.3.3 Before Exercise

Before exercise, the risk neutrality of agents implies that the equity value, which is at the same time the firm value, generates expected capital gains over time as high as the risk-free rate.

\[ dX = \mu_B X dt + \sigma_B X dW_t, \quad X_0 = X_0 \] \tag{III.10}

\[ rV(X)dt = E[dV(X)] \] \tag{III.11}

Solving the resulting differential equation yields the value of the firm at time \( t = 0 \)
\[ V(X) = \left( \frac{\lambda \pi \bar{X} \left( 1 + \frac{\tau}{1-\tau} \left( 1 - \vartheta \left( 1 - \alpha + \frac{\alpha}{\tau} \right) \right)^{1/\vartheta} \right)}{\bar{X}} \right)^{\xi} \]
with $\Lambda$ and $\theta$ defined in Equations III.8,

$$
\xi = \frac{1}{2} - \frac{\mu_B}{\sigma_B^2} + \sqrt{\left(\frac{\mu_B}{\sigma_B^2} - \frac{1}{2}\right)^2 + \frac{2r}{\sigma_B^2}}
$$

(III.12)

and the optimal exercise threshold

$$
\bar{X} = \frac{\kappa}{\Lambda \pi} \frac{\xi}{1 - \xi} \left(1 + \frac{\tau}{1 - \tau} \left(1 - \theta \left(1 - \alpha + \frac{\alpha}{\tau}\right)^{1/\theta}\right)^{-1}\right).
$$

### III.4 Calibrating the Model

This section describes how parameters were chosen so as to reproduce the real-life decision to contract away risk in a project company. Thus, I need to calibrate the model to empirically observable values if possible and, if not, make reasonable assumptions. All parameters remain time invariant. For example, the company has the same tax rate during its infinite lifetime, regardless of cash flow or revenues. For parameters for which I have no observable proxy, I follow recent literature so as to make my results comparable at least to some extent.

The initial diffusion level and the exercise cost are set arbitrarily and are scalable: $X_0 = 20$, $\kappa = 100$. Naturally, the more contracts that are signed, the higher the project company’s exercise costs will be. Contracting costs are challenging to estimate since data on projects are sparse, project contracting is not very comparable to other corporate events, and the company has no history. Consequently, I refrain from increasing the costs as number of contracts increases.

The revenue stream, a fraction of the diffusion process, is arbitrarily set to $\pi = 100\%$. For the annual risk-free rate, $r$, I use the average return of of 7\% from 10-year U.S. Treasury Bills between 1980 and 2012. The annual growth rate of the cash flows before exercise, $\mu_B$, is fixed at 1\%, based on Morellec and Schürhoff (2011), who calculate it as S&P 500 firm’s payout ratio to shareholders (average dividends) and bondholders (average coupons) weighted by the median leverage ratio. Since the project company has no further growth options, the cash-flow stream should not grow over time ($\mu = 0\%$).

The effective tax rate varies dramatically in the literature. For example, Hackbarth et al. (2012) use an annual tax rate of 15\%, whereas Strebulaev (2007) uses 35\%. Empirical evidence suggests that tax payments have a highly volatile distribution. Dyreng et al. (2008) extract the real cash payments to the U.S. IRS and divide them by pretax income. They find a mean of 30\% and a volatility of 0.18 for 10-year periods ($\tau = 30\%$). Although project finance is often assumed to be off-balance-sheet financing, it cannot be directly associated with special purpose vehicles (SPV). SPVs are often headquartered in tax havens, whereas empirical literature on project finance finds no such location bias. In robustness tests, I varied the overall tax rate and obtained consistent results.

Davydenko et al. (2012) estimate the total cost of default for a U.S. company, including
direct and indirect costs, to be around 22% of firm’s asset value. However, Esty (2002) points out that project companies are frequently restructured prior to actual default and Byoun et al. (2012) argue that direct and indirect bankruptcy costs in project finance are negligible, which may be due to the high degree of power wielded by lenders in project companies. Consequently, the parameter for $\alpha$ should be significantly lower. Davydenko et al. (2012) estimate the total cost of bond renegotiation to be approximately 14% of bond value. Although using this figure means that some contracts will be disregarded, it serves as a good approximation ($\alpha = 14\%$).

Similar to Strebulaev (2007), annual cash-flow volatility, $\sigma_B$, is set at 0.25. This is the average, leverage-adjusted equity return volatility of S&P 500 firms. If an off-take agreement exists, the cash-flow volatility can tend toward zero, dependent on the credit worthiness of the off-taker. However, this type of agreement is usually for a limited time only, whereas the revenue generating company is not and it may not cover the total output. After the off-take agreement expires, the separated company may face sales volatility, depending on supply and demand of the specific industry. Since the growth option has a hypothetically infinite lifetime that outlasts any off-take agreements, the resulting average cash-flow volatility should be comparably low. Further, Kleimeier and Megginson (1999) point out that project finance is primarily used for projects with transparent cash flows. To set the lower boundary, I follow Dyreng et al. (2008), who estimate a volatility ($\sigma$) of 7% for the lowest 5% quantile of the cash-flow volatility distribution for U.S. firms.

![Figure 9: Influence of the factor $\psi$ on the probability of failure and expected time to failure of the company.](image)
Parameter $\psi$ is fixed to reproduce real-life, annual project company failure risk. I find a maximum of $\psi = 0.1$ and a minimum of $\psi = 0.01$ to be representative. Figure 9 displays the influence of the factor $\psi$ on the survival of the firm. The left figure shows that at the age of 10 years, the probability of failure for the highest and lowest values of $\psi$ differ significantly. For the highest value of $\psi$, the probability of failure after 10 years is slightly higher than 60%; the lowest value yields a probability of less than 10%. The setup shows that default probability is concave and, indeed, this makes sense because the longer the company exists, the more settled it becomes in the market and its environment, and the more its marginal probability of failure decreases. The right figure shows the expected lifetime of the company, which is independent of the point in time. Again, the influence of $\psi$ on firm value differs significantly. For the highest value of this factor, the company has an expected lifetime of 10 years, while for the lowest value, the expected lifetime is 100 years. Since the project is presumably large, the highest probability of failure, $\psi = 0.1$, introduces an expected time to failure after the construction phase. An expected lifetime of 100 years renders failure risk negligible for the company, bankruptcy risk now is most important. Altogether the figure shows that the firm has an incentive to reduce the factor $\psi$ to increase its overall value and that the parameters do generate reasonable scenarios.

Table 10 summarizes all parameter choices described in this section.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Fixed Parameters</th>
<th>Contracting Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial value of the state variable</td>
<td>$X_0$</td>
<td>20</td>
</tr>
<tr>
<td>Revenue fraction of the state variable</td>
<td>$\pi$</td>
<td>100%</td>
</tr>
<tr>
<td>Annual risk-free rate</td>
<td>$r$</td>
<td>7%</td>
</tr>
<tr>
<td>Drift of the after-exercise state variable</td>
<td>$\mu$</td>
<td>0%</td>
</tr>
<tr>
<td>Drift of the before-exercise state variable</td>
<td>$\mu_B$</td>
<td>1%</td>
</tr>
<tr>
<td>Annual cash flow volatility</td>
<td>$\sigma_B$</td>
<td>25%</td>
</tr>
<tr>
<td>Annual tax rate</td>
<td>$\tau$</td>
<td>30%</td>
</tr>
<tr>
<td>Exercise costs</td>
<td>$\kappa$</td>
<td>100</td>
</tr>
<tr>
<td>Bankruptcy costs</td>
<td>$\alpha$</td>
<td>14%</td>
</tr>
<tr>
<td>Annual failure risk</td>
<td>$\psi$</td>
<td>10%</td>
</tr>
</tbody>
</table>

III.5 Results

III.5.1 The Separated Firm

I assess the values for each participating party in the project firm and assume an initial level of 20 for the diffusion process $Z$ and calibrate the model with values from Table 10.
Figure 10: Influence of the factors $\psi$ and $\sigma$ on equity value, debt value, leverage ratio, and bankruptcy threshold for three values of the other risk factor.
Figure 10 displays the impact of varying parameters of $\psi$ and $\sigma$ on equity and debt value, leverage ratio, and the bankruptcy threshold. The first row of plots indicates that for equity holders, a reduction in the risks has a contrasting effect on their wealth. Reducing the hazard of failure strongly increases the value of the equity claims. The implication of Merton’s (1974) model that cash-flow volatility is favorable for equity holders still holds. For low values of both risks, the effect of reducing the other risk is about the same, just with another sign. In a scenario where $\sigma$ is low, the gain of reducing $\psi$ equals the loss of reducing $\sigma$ in a low $\psi$ environment. For higher values of sigma, the gains of reducing $\psi$ clearly outweigh the losses of $\sigma$. This indicates that equity holders prefer a lower value of $\psi$ even if cash-flow volatility is reduced by the same relative amount. Additionally, this might explain the low, observable cash-flow volatility in project finance. For lenders, reducing both risks is optimal. This is not surprising, as failure risks diminish the value of the perpetual coupon due to a total loss, while cash-flow volatility increases the risk of bankruptcy, a partial loss for lenders, and decreases the optimal coupon. However, the rise in $\psi$ is still greater than the rise in $\sigma$. For lenders, a reduction in $\psi$ is preferred over a reduction of $\sigma$.

![Figure 11](image)

**Figure 11**: Gains in firm value in percent at the time of exercise of the growth option compared to the baseline value of the firm. The value is zero if neither equity holders nor debt holders are better off. The left plot holds three values of $\sigma$ while varying $\psi$ and the right plot holds three values of $\psi$ while varying $\sigma$.

It is natural that leverage ratios increase with lower cash-flow risk. However, I observe a reduction with lower failure risk. Although the impact is small, it shows that failure risk has an influence on leverage. Mathematically speaking, the ratio deteriorates because firm value increases faster than the value of debt. This is clear because equity also gains in value. Since the absolute gain of equity holders is low compared to the gains of bondholders, the leverage ratio responds only minimally. Economically, the leverage ratios correspond
to those empirically observed. Byoun et al. (2012) find a mean leverage ratio of project companies of about 89%, which corresponds to the leverage value when most risks are contracted away. Similarly, Esty (2002) report leverage ratios of 70–90 % for power-producing projects and Gatti et al. (2013) find a mean (median) leverage ratio of 77% (72%). The bankruptcy threshold exhibits the same pattern as the leverage ratio. The maximum impact of contracting away the danger of failure is about 4%, while the average impact for cash-flow risk is more than 10%. Failure risk is only slightly associated with bankruptcy, which is plausible since bankruptcy solely depends on the level of the diffusion process modeling cash flows.

I assume that in the normal state of nature, that is, exercising the growth option without contracting, exposure to both value-decreasing dangers is maximal ($\psi = 0.1, \sigma = 0.25$). I refer to this as the baseline scenario, and it is typical of project finance. Cash-flow volatility is similar to that of an average U.S. firm, depending on the sales market, and the probability of losing all assets is expected to occur within 10 years of operation. I further assume that the risks are interdependent. It is reasonable to assume that the reduction in cash-flow volatility influences the failure risk and vice versa. Consider, for example, the contracting of exchange-rate risk. If a bank guarantees a fixed exchange rate, it hedges part of the project’s political risk in the project country and decreases cash-flow volatility. The same holds for inflation risk. Further, construction guarantees can ensure that assets-in-place are generated and limit the risk of delayed output. Supply contracts smooth the payments for material and hedge a small amount of political risk. Extreme cost overruns can result in a project being abandoned. In their sample of 47 mega-projects, Merrow et al. (1988) find mean cost factors of 1.88, meaning that, on average, large projects need 88% more money than was expected. The authors further find a minimum of 0.97, meaning that most mega-projects barely meet their target costs (only four in their sample); the maximum cost overrun is an impressive 353%. The authors argue that the non-anticipated costs stem from inflation and conflicts between the host government and project management. Thus, for example, inflation indexing of revenues smooths the income stream and insures political risks. It thus makes sense to contract away the risks that influence both factors $\psi$ and $\sigma$.

Projects are large and capital intensive and the parent company’s equity contribution is low. It is thus assumable that both equity holders and lenders have about the same amount of power regarding contracting and that both try to reduce the risk most dangerous to them. Consequently, I strictly assume that one party does not expropriate wealth from the other party compared to the baseline scenario. Lenders want to increase the value of their claims. Equity holders do, too, if their wealth also increases with the contract. In Figure 10, the marginal increase for equity holders in $\psi$, for realistic values, is bigger than the marginal decrease in $\sigma$. Plots in Figure 11 underline this. Assuming that both claim...
holders have equal power, I assess the surplus in firm value when both parties are better off compared to the baseline scenario. The left plot indicates that if $\psi$ has decreased enough, the influence of diminishing cash flow risk is negligible. For low cash-flow volatility, a high amount of failure risk has to be contracted to improve the position of both parties. The right plot shows that for high-failure-risk firms, contracting only cash flow risk does not increase stakeholder wealth and, hypothetically, a project will not be undertaken. The plot further indicates that in the case of average failure risk, managers may refrain off-take agreements. Overall, project firms can increase their value by about 150% compared to a typical project finance scenario if they are able to contract away failure as well as cash-flow risk to practically induced lower boundaries.

I disentangle the gains shown in Figure 11 into gains from equity and gains from debt in Figure 12. All surpluses depend on the equity holders because, first, they profit from reducing only one risk factor, whereas debt holders prefer to reduce both risks, and, second, the gain from reduced failure risk is less than the gain of debt holders. Lenders gain only if equity holders are sufficiently compensated with a reduction in failure risk. In relative terms, equity holders gain more than debt holders in the case of high cash-flow risk, but generally are worse off than lenders. The maximum relative gain after contracting both risks is about 120% for lenders and 55% for equity holders. In unreported, absolute terms I find that debt holders are always better off due to the increase in value for both risk factors.
**Figure 12:** The overall surplus from Figure 11 disentangled by stakeholder for varying failure risk $\psi$ on the left side and varying $\sigma$ on the right. The value is zero if neither equity holders nor debt holders are better off. The surplus is the difference from the baseline scenario for positive values.
III.5.2 Value at Time $t=0$

![Graphs showing the influence of factors $\psi$ and $\sigma$ on the exercise boundary, value of the option, and value of the equity provided by the parent company at time $t=0$.]

Figure 13: Influence of the factor $\psi$ and $\sigma$ on the exercise boundary, value of the option, and value of the equity provided by the parent company at time $t=0$.

I also assess the value of the firm at time $t=0$, which is when the parent company can decide on the timing of the exercise. In contrast to standard real-option models, wealth is transferred between two markets. Before exercise, expected capital gains from equity separated by the parent company follow the stochastic process $X$ with drift $\mu_B$ and volatility $\sigma_B$. At initiation of the new company, the equity is transferred to the project finance firm and expected capital gains follow the drift-less ($\mu = 0$) stochastic process $Z$ with volatility $\sigma$ and failure risk $\psi$, depending on the amount of contracting. However, these results must be viewed with some caution as the net of contracts is woven at the time of exercise. The value is known only when sponsors can assess ex-ante the amount of contracting possible even if lenders are not yet involved. In Figure 13 I plot the values of the equity and the exercise threshold for both risk parameters. The process value $X_{t\xi}$ where exercise is optimal strongly depends on the value of the growth option. Figure 10 revealed that the value of the firm heavily depends on failure risk and only minimally on cash-flow risk. This is reflected in the time of the exercise: the higher the value of the option, the earlier the
option is exercised, and the higher the value of the separated equity. However, while the value of the new firm increases exponentially in failure risk, the exercise value decreases only linearly. The marginal effect of ensuring the firm assets’ survival has a smaller impact on timing than on the value of the newly initiated firm. From the lower two plots I infer that before exercise, a reduction in cash-flow risk $\sigma$ and transfer to the project finance environment only slightly influences equity value. In contrast, the reduction in failure risk exponentially increases the value of equity, although before exercise, equity can be redeployed to other parent company projects.

### III.6 Conclusion

In this paper, I depart from assessing the choice of the organizational form of growth options and instead investigate the effects of contracting away risks after the decision to form a project company has been made. Projects are found to be large and capital intensive and the parent company’s (the sponsor) equity contribution to be low. Further, empirical studies show that lenders are not compensated for the high corporate risk by higher spreads but instead receive control rights. Thus, standard models of the firm are inadequate because managers in project finance not only decide in favor of equity holders but also take into consideration the interests of lenders. However, lender’s power still does not explain why project firms contract away asset risk to a large extent, significantly reducing the wealth of equity holders. Project firms typically have a higher exposure than standard corporations to risk that, if not avoided, can render the claims of all stakeholders worthless, such as political risk and construction risk. I introduce a new risk factor to a standard real-option model: failure risk, which is the probability that the entire firm will fail. With this type of risk, in contrast to bankruptcy induced by low income, debt holders also lose their claims. I show that since cash-flow risk and failure risk typically interact, equity holders have good reason to reduce asset risk. Calibrated to realistic values, I show that the model explains the empirically observable leverage ratios and I find that firm value can increase by 120% if both risks are reduced, compared to a typical project finance scenario. Lastly, this approach can overcome the typical assumption of an infinite lifetime for exercised growth options.
III.7 References


and Nicholas S. Majluf, “Corporate financing and investment decisions when firms have information that investors do not have,” *Journal of Financial Economics*, 1984, 13, 187–221.


III.8 Discussion of the Model and Extensions

In this appendix, I discuss natural extensions of the model with respect to the project finance background. Like every theoretical model, it has the potential to become more realistic, typically as a tradeoff against non-analytic solutions and high computational effort. The model does not contain a financing period, even though it might take more time to leverage a project than a standard firm. A financing period will make exercise occur sooner, since the lag in the revenue stream will force the company to exercise earlier. Thus, reducing the revenue stream $\pi$ can also be interpreted as a cost of difficult financing; however, my results remain unchanged for different parameters of $\pi$.

I exclude from the model a construction phase for the project and a time-to-build for the cash flows for the following reasons. Project finance contracts also influence the costs of exercising the growth option and how long it will be before first revenues are received. There are two opposing aspects regarding the time factor. On the one hand, writing contracts is time consuming. Big projects can involve more than 1,000 contracts (Esty, 2002). Moreover, these contracts are not stand-alones, but involve complex interdependencies with each other, which can best be described as a mathematical multidimensional optimization problem. This takes time to solve and settle. On the other hand, many project contracts specify penalties for delayed completion or supply. For example, construction firms engaged to build a plant may be required by contract to meet a pre-specified deadline or face various forms of monetary penalties. Thus, the threat of penalties has the potential to speed project completion. Although these incentives should force construction companies to fulfill contractual deadlines or limits, 50% of projects experience time overruns (Esty, 2002). Thus, contracting can be either a faster or slower way of completion.

Ex ante, the effect of contracting on exercise costs, which include the cost of building the option, is straightforward. Marginal costs of contracting remain positive due to the difficulty of coming to agreement and the need to engage the services of law firms. Hence, every decrease in market risk due to every new contract is accompanied by higher exercise cost. Ex post, the effect remains unclear and depends on the market situation, which can change either to advantage of the fixed supply contracts or not. For example, fixed supply agreements yield a discount since they are an off-take agreement for the supplier, but could lead to a loss for the purchaser due to the possibility of new, cheaper, or more efficient construction materials. The same holds for fixed exchange rates. Many projects involve multinational contractors and investment banks tend to fix exchange rates between the parties. Although this eliminates exchange-rate risk, ex-post it is not clear whether the effect is optimal for the project company, depending on the direction of the exchange rate change after contracting. However, modeling this is beyond the scope of my approach. I thus assume constant exercise cost and exclude ex post market situation effects due to contracting since those remain unknown.
III.9 Derivations

III.9.1 After Exercise

III.9.1.1 Firm Value

Investors are risk neutral and thus require a rate of return of $r$ for the company. The company’s owner receives capital gains $\mathbb{E}[dV(Z)]$, taxed cash flows $(1 - \tau)\pi Z dt$, and payments due to tax-deductible debt $\tau C dt$.

$$rV(Z) dt = \mathbb{E}[dV(Z)] - \psi V(Z) dt + [(1 - \tau)\pi Z + \tau C] dt \quad (\text{III.13})$$

Applying Itô’s Lemma to $dV(Z)$ on the right side of the Bellman equation (III.13) and using the martingal property yields

$$\mathbb{E}[dV] = E\left[\frac{\partial V}{\partial Z} dZ + \frac{1}{2} \frac{\partial^2 V}{\partial Z^2} \sigma^2 Z^2 dt\right]$$

$$= E\left[\frac{\partial V}{\partial Z} \mu Z dt\right] + E\left[\frac{1}{2} \frac{\partial^2 V}{\partial Z^2} \sigma^2 Z^2 dt\right]. \quad (\text{III.14})$$

Plugging this back into (III.13) and canceling out the $dt$ term gives a non-homogeneous, second-order, ordinary differential equation.

$$\frac{\sigma^2}{2} Z^2 V''(Z) + \mu Z V'(Z) - (r + \psi)V(Z) + (1 - \tau)\pi Z + \tau C = 0 \quad (\text{III.15})$$

Finding a general solution to the non-homogeneous, second-order differential equations with non-constant factors requires a particular solution as well as two independent solutions for the homogeneous case. Consider $f(Z) = mZ + n$ as the particular solution with $f'(Z) = m$ and $f''(Z) = 0$. Solving for both constants yields

$$-(r + \psi)n + \tau C = 0 \iff n = \frac{\tau C}{r + \psi} \quad \text{and}$$

$$\mu Z m - (r + \psi)m Z + (1 - \tau)\pi Z = 0 \iff m = \frac{(1 - \tau)\pi}{(r + \psi - \mu)} = \Lambda \pi.$$ 

Therefore, I found a particular solution to Equation (III.15), namely

$$V(Z) = \Lambda \pi Z + \frac{\tau C}{r + \psi}. \quad (\text{III.16})$$

Next I am looking for two independent solutions to the homogeneous version of Equation (III.15). Obtain the homogeneous equation by dropping all summands independent of $V$:

$$\frac{\sigma^2}{2} Z^2 V''(Z) + \mu Z V'(Z) - (r + \psi)V(Z) = 0 \quad (\text{III.17})$$
Consider $f(Z) = Z^a$ as a solution with $f'(Z) = aZ^{a-1}$ and $f''(Z) = a(a-1)Z^{a-2}$ as solutions for (III.17).

$$a(a-1)Z^a + \frac{2\mu}{\sigma^2}aZ^a - \frac{2(r + \psi)}{\sigma^2}Z^a = 0 \Leftrightarrow a^2 + \left(\frac{2\mu}{\sigma^2} - 1\right)a - \frac{2(r + \psi)}{\sigma^2} = 0 \quad \text{(III.18)}$$

Now Equation (III.18) can be solved with the standard formula for quadratic equations, yielding two values for $a$: $\theta$ and $\lambda$.

$$\lambda = \frac{1}{2} - \frac{\mu}{\sigma^2} + \sqrt{\left(\frac{\mu}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2(r + \psi)}{\sigma^2}} \quad \text{(III.19)}$$

and

$$\theta = \frac{1}{2} - \frac{\mu}{\sigma^2} - \sqrt{\left(\frac{\mu}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2(r + \psi)}{\sigma^2}}. \quad \text{(III.20)}$$

Thus $Z^\theta$ and $Z^\lambda$ are solutions to the homogeneous differential Equation (III.17). It remains to show that these two are independent. Before doing this, I will show that $\theta < 0$ and $\lambda > 1$ for normal parameter values. This implies that $\mu > 0$, $\sigma > 0$, $Z > 0$ and $r + \psi - \mu > 0$ and an upper estimate of 0 for $\theta$ and a lower estimate for $\xi$ of 1 can easily be found. This allows showing that the Wronskian determinant is strictly lower than zero and thus the solutions are independent:

$$\begin{vmatrix}
Z^\lambda & Z^\theta \\
(Z^\lambda)' & (Z^\theta)'
\end{vmatrix} = \theta Z^\lambda Z^{\theta-1} - \lambda Z^{\lambda-1} Z^\theta < 0$$

The general solution to the original non-homogeneous, second-order, ordinary differential equation with non-constant factors (III.15) is of the form:

$$V(Z) = f_p(Z) + A_1 f_{h,1}(Z) + A_2 f_{h,2}(Z)$$

where $f_p(Z)$ denotes the particular solution, $f_{h,1}(Z)$ and $f_{h,2}(Z)$ the solutions for the homogeneous case, and $A_1$ and $A_2$ are constants to be determined by value-matching conditions. Inserting all three solutions yields:

$$V(Z) = \Lambda \pi Z + \frac{\tau C}{r + \psi} + A_1 Z^\theta + A_2 Z^\lambda$$

Now, determine the two constants $A_1$ and $A_2$. For high values of $Z$, the probability of bankruptcy is low and coupon and tax shield are negligible. Therefore, $\lim_{Z \to \infty} V(Z) = \lim_{Z \to \infty} Z$ and $A_2 = 0$ since $\lambda > 1$. Further, I know that the cash-flow stream will be constant after
bankruptcy at threshold $Z$: 

$$V(Z) = (1 - \alpha)\pi AZ \iff A_1 = \frac{-\tau C}{r + \psi} - \alpha \Lambda \pi Z$$

Rearranging the equation for the value of the firm after exercise yields

$$V(Z) = \pi AZ + \frac{\tau C}{r + \psi} \left(1 - \left(\frac{Z}{\bar{Z}}\right)^{\delta}\right) - \alpha \Lambda \pi Z \left(\frac{Z}{\bar{Z}}\right)^{\delta}. \quad (\text{III.21})$$

### III.9.1.2 Equity Value

Certain steps in the derivations already explained in detail above are omitted here. Investors are risk neutral and thus require a rate of return of $r$ for the company. The company’s equity owner receives capital gains $E[\delta \Xi(Z)]$, taxed cash flows $(1 - \tau)\pi Z dt$, has to pay the perpetual coupon, and receives the tax deduction $(1 - \tau)C dt$.

$$r \overline{\Xi}(Z) dt = E[\delta \Xi(Z)] + [(1 - \tau)\pi Z + (\tau - 1)C] dt$$

Again applying Itô’s Lemma yields the following differential equation:

$$\frac{\sigma^2}{2} Z^2 \overline{\Xi''}(Z) + \mu Z \overline{\Xi'}(Z) - (r + \psi) \overline{\Xi}(Z) + (1 - \tau)\pi Z + (\tau - 1)C = 0 \quad (\text{III.22})$$

The only difference from (III.15) is the coupon involving term. Thus, the particular solution to (III.22) is:

$$\overline{\Xi}(Z) = \Lambda \pi Z + \frac{(\tau - 1)C}{r + \psi}. \quad (\text{III.23})$$

Since the homogeneous ODE for equity value is similar to (III.22), the solution of (III.22) is

$$\overline{\Xi}(Z) = \Lambda \pi Z + \frac{(\tau - 1)C}{r + \psi} + A_1 Z^0 + A_2 Z^\delta \quad (\text{III.24})$$

Again, similar asymptotic behavior arguments yield $A_2 = 0$. If the cash flow $Z$ reaches the bankruptcy threshold, the value of equity is zero.

$$\overline{\Xi}(Z) = 0 \iff A_2 = \frac{-\frac{\tau - 1)C}{r + \psi} - \Lambda \pi Z}{Z^\delta}$$

Rearranging the equation for the equity value after exercise yields

$$\overline{\Xi}(Z) = \pi AZ - \frac{(1 - \tau)C}{r + \psi} - \left(\pi AZ - \frac{(1 - \tau)C}{r + \psi}\right) \left(\frac{Z}{\bar{Z}}\right)^{\delta}. \quad (\text{III.25})$$
III.9.1.3 Bankruptcy Threshold
The bankruptcy threshold is determined by the equity holders to maximize their value: 
\[ \frac{\partial E(Z)}{\partial Z} \bigg|_{Z=Z^*} = 0 \]
\[ Z^* = \frac{\theta}{\theta - 1} \frac{r + \psi - \mu C}{r + \psi - \pi} \]

It can easily be shown that this is indeed a maximum.

III.9.1.4 Coupon
The optimal coupon is determined so as to maximize firm value. First, plug in the optimal default threshold and take the derivative toward the coupon C:
\[ \frac{\partial V(Z)}{\partial C} = 0 \]
\[ \Rightarrow C = Z\pi \frac{\partial}{\partial \theta} \left( \frac{r + \psi}{r + \psi - \mu} \left( 1 - \theta \left( 1 - \alpha + \frac{\alpha}{\tau} \right) \right) \right) \]

This is a maximum for all values of C.

III.9.1.5 Value of Debt
The bondholders’ value \( D(Z) \) evolves over time as:
\[ rD(Z)dt = E[\Delta D(Z)] - \psi D(Z)dt + Cdt. \]

This value depends on the level of cash flows since bondholders are affected when \( Z \) reaches the bankruptcy threshold. This again leads to the following PDE.
\[ \frac{\sigma^2}{2} Z^2 D''(Z) + \mu Z D'(Z) - (r + \psi)D(Z) + C = 0 \] (III.26)

Arguments similar to those made for derivation of the firm value yield
\[ D(Z) = \frac{C}{r + \psi} + G_1 Z^\theta + G_2 Z^\lambda \]
as a solution for the PDE (III.26). Since I have \( D(Z) \to C/r \) as \( Z \to \infty \), bankruptcy is unlikely, and because \( \xi > 0 \), the factor \( G_2 \) must have the value zero. In case of bankruptcy, debt holders receive all the money since equity holders lose all their wealth:
\[ D(Z) = (1 - \alpha)\Lambda \pi Z \]
\[ D(Z) = \frac{C}{r + \psi} + \left( (1 - \alpha)\Lambda \pi Z - \frac{C}{r + \psi} \right) \left( \frac{Z}{Z^*} \right)^\theta \] (III.27)
III.9.2 Before Exercise

Similar arguments yield that before exercise, firm value, which also represents equity value, does not generate a revenue stream:

\[ rV(X)dt = E[dV(X)] \]  
\[ \mu_B XV + \frac{\sigma_B^2}{2} X^2 V_{XX} - rV = 0 \]

Again, using \( f(X) = X^\xi \) yields

\[ V(X) = B_1X^\xi + B_2X^{\lambda_B} \]  
with

\[ \xi = \frac{1}{2} - \frac{\mu_B}{\sigma_B^2} + \sqrt{\left(\frac{\mu_B}{\sigma_B^2} - \frac{1}{2}\right)^2 + \frac{2r}{\sigma_B^2}} \]

and

\[ \lambda_B = \frac{1}{2} - \frac{\mu_B}{\sigma_B^2} - \sqrt{\left(\frac{\mu_B}{\sigma_B^2} - \frac{1}{2}\right)^2 + \frac{2r}{\sigma_B^2}}. \]

Since the equity is worthless when the revenue stream tends to zero, I have \( V(X) \to 0 \) if \( X \to 0 \). Therefore, I obtain \( B_2 = 0 \), since \( \lambda_B < 0 \). For \( B_1 \), I need the value at exercise threshold \( \bar{X} = \bar{Z} \).

\[ V(\bar{X}) = V(\bar{Z}) - \kappa \]

Merging the values for firm value, coupon, and bankruptcy threshold after exercise, I derive a simplified version of firm value:

\[ V(Z) = \Lambda\pi Z \left(1 + \frac{\tau}{1 - \tau} \left(1 - \vartheta \left(1 - \alpha + \frac{\alpha}{\tau}\right)\right)^{1/\theta}\right) \]

This yields the following equation for firm value at time \( t = 0 \):

\[ V(X) = \left(\Lambda\pi \bar{X} \left(1 + \frac{\tau}{1 - \tau} \left(1 - \vartheta \left(1 - \alpha + \frac{\alpha}{\tau}\right)\right)^{1/\theta}\right) - \kappa\right) \left(\frac{X}{\bar{X}}\right)^\xi \]

The only thing I need is the optimal exercise threshold \( \bar{X} \). Therefore, I use the smooth-pasting condition on the boundary of exercise, \( \partial V(X)/\partial X|_{X=\bar{X}} = \partial V(Z)/\partial Z|_{Z=\bar{Z}} \), which yields

\[ \bar{X} = \kappa \frac{\xi}{\Lambda\pi \left(1 - \xi\right)} \left(1 + \frac{\tau}{1 - \tau} \left(1 - \vartheta \left(1 - \alpha + \frac{\alpha}{\tau}\right)\right)^{1/\theta}\right)^{-1} \]
IV Curriculum Vitae

Personal Information
Name: Tobias Nigbur
Date of birth: April 1, 1983
Place of origin: Ibbenbüren, Germany
Citizenship: German

Education
2013 University of British Columbia, Vancouver, Canada
    Visiting Scholar
2009 – 2014 University of St.Gallen, St.Gallen, Switzerland
    Ph.D. Student
2010, 2011, 2012 Advanced Courses for Doctoral Students, Switzerland
2003 – 2009 University of Münster, Münster, Germany
    Diploma in Mathematics
2007 University of California, Santa Barbara, USA

Working and Teaching Experience
2013 – Financial Markets and Portfolio Management
    Referee
2009 – 2012 Swiss Institute of Banking and Finance, University of St.
    Gallen, Switzerland
    Teaching and Research Assistant
2010 – 2012 Fit for Finance, Zurich, Switzerland
    Lecturer
2006 – 2007 Mathematics Department, University of Münster, Germany
    Teaching Assistant