On the efficacy of stop-loss rules in the presence of overnight gaps

Argimiro Arratia*
Computer Science,
Barcelona Tech (UPC),
Barcelona, Spain
argimiro@cs.upc.edu

Albert Dorador
European Central Bank†
Frankfurt am Main, Germany
albert.dorador@ecb.europa.eu

Abstract

A stop-loss rule is a risk management tool whereby the investor predefines some condition that, upon being triggered by market dynamics, implies the liquidation of her outstanding position. Such a tool is widely used by practitioners in financial markets with the hope of improving their investment performance by cutting losses and consolidating gains. We analyse in this work the performance of four popular implementations of stop-loss rules applied to asset prices whose returns are modelled with consideration of overnight gaps, that is, jumps from the closing price of one day to the open price of the next trading day. In addition, our models include acute momentary price drops (*flash crashes*), which are often believed to erode the performance gains that might be derived from stop-loss rules. For this analysis we consider different models of asset returns: random walk, autoregressive and regime-switching models. In addition, we test the performance of the considered stop-loss rules in a non-parametric, data-driven framework based on the stationary bootstrap. As a general conclusion we find that, even when including overnight gaps and flash crashes in our price models, in rising markets stop-loss rules improve the expected risk-adjusted return according to most metrics, while improving absolute expected return in falling markets. Furthermore, we find that in general the simple fixed percentage stop-loss rule may be, in risk-adjusted terms, the most powerful among the popular rules that this work considers.

Keywords: Stop-loss; Risk Management; Financial modeling; overnight gap; flash crash; bootstrap.

1 Introduction

A *stop-loss* order is an order an investor may place so that her position is liquidated the moment a certain pre-specified condition (set according to a *stop-loss rule*) is met by market dynamics. Whether the position the investor holds is short or long, the purpose of setting a stop-loss is to cut losses and consolidate gains, being this a most basic yet popular tool for risk management.

*Research supported by the Spanish Ministry of Science and Innovation (Contract TIN2017-89244-R) and the recognition 2017SGR-856 (MACDA) from AGAUR (Generalitat de Catalunya)

†Disclaimer: the views expressed in this document reflect those of the authors and are not necessarily shared by the European Central Bank.

The importance of placing stop-loss orders cannot be overestimated. The fall of Banco Popular in Spain (June 2017) is a very recent example of the dangers of not incorporating stop-loss orders into one's investing strategy: investors that bought-and-held hoping for an eventual trend reversal lost 100% of their investment, as Banco Santander bought Banco Popular's four billion shares for EUR 1, driving the individual share price to zero. Additionally it has been argued that the setting of stop-loss orders guards investors from the disposition effect [Silva and Da Silva, 2015], which refers to the observed tendency of investors to hold on to losing positions much longer than they hold to winning positions.

Several researchers have looked at the problem of discerning the effectiveness of stop-loss orders in various contexts. We mention a few that are of particular interest to our work. The paper by Acar and Toffel [2001] studies how a stop-loss rule affects the returns distribution, by assuming the asset follows a Brownian Random Walk with drift, and then evaluating the financial profitability of a simple stop-loss strategy under the previous assumption. Kaminski and Lo [2014] follow a similar approach: they developed a rigorous analytical framework for measuring the impact of simple 0/1 stop-loss-re-entry rules on the expected return and volatility of an arbitrary portfolio strategy (again assuming that assets follow a Random Walk) and provide an empirical analysis of performance of a stop-loss strategy against buy-and-hold in U.S. equities. Lo and Remorov [2017] extends the previous work of Kaminski and Lo [2014] by analysing the efficacy of stop-loss trading strategies on serially correlated asset returns that follow a Markov regime-switching process, and subject to transaction costs. They conclude that the stop-loss strategy may outperform buy-and-hold, provided there is sufficient serial correlation in returns with some impact on downside risk, although this can be overturned due to the high trading costs that a stop-loss-re-entry strategy (unlike ours) entails. James and Yang [2010] on the other hand, do not assume that financial assets obey a particular model, but instead base their analysis on the use of the stationary bootstrap as a tool to replicate financial time series adequately.

None of the previously mentioned works, and others that we surveyed related to the assessment of the value of stop-loss rules, considers the oftentimes observed large variation in the price across non trading hours. This clearly affects the correct triggering of a stop-loss rule due to the by-passing of the established stopping-time.

The early works of French [1980], Gibbons and Hess [1981] and Keim and Stambaugh [1984], showed that returns on non trading days present different distribution than returns on trading days. In particular, they showed that in the U.S. market the expected returns of stocks are significantly lower or even negative on Mondays compared to returns on preceding Fridays, a phenomenon that has been coined the *weekend effect*.

Cooper et al. [2008] found that in the more recent price history in U.S. markets overnight returns (including weekends) are consistently positive while daytime returns are close to zero or negative; thus, concluding that the U.S. equity premium on the first decade of the 21st century is solely due to overnight returns. Wiener and Tompkins [2008] extended the study of return differences between trading and non-trading hours to higher moments of the returns distribution and for European and Japanese markets, and found that in general the distribution of non-trading periods (or overnight) returns displayed a higher degree of non-normality compared to that of the trading periods returns. This lead these researchers to conclude that while trading periods returns may follow some diffusion process, the non-trading periods returns follow a jump process. This fact was also argued by Geman et al. [2001] who studied asset prices arising from market clearing conditions.

The present work contributes to the analysis of the value of stop-loss rules in various novel aspects, considering previous studies on the subject and on the different distribution displayed by non-trading periods returns as opposed to trading periods returns. First, and most important, given that an accurate assessment of the value of stop-loss rules requires full consideration of the real behavior of asset returns, we take into consideration two important anomalies. On the one hand, we consider overnight gaps in prices (jumps from the Closing price of one day to the Open price of next trading day) as an additional feature that we include in the implementation of various well-accepted models for the behavior of asset prices. On the other hand, we also consider flash crashes, an extreme market event that consists in a sharp price drop caused by a variety of non-fundamental reasons (in the financial sense of the word "fundamental") that is offset (in part or, usually, in full) in the next minutes or hours. These rare but possible events could potentially have a significant impact on the value of stop-loss rules, and hence we decided to also duly analyze them in this work. Second, as a complement to our asset model-based analysis, we consider a non-parametric approach based on the stationary bootstrap, thus providing an alternative point of view for assessing the advantages of stop-loss rules. Additionally we contribute with two new performance measures, the Return-VaR (RVaR) and the Return-ES (RES) ratios, based on the VaR and Expected Shortfall risk measures. For all our experiments we use high-frequency financial data (hourly prices). It should be noted that implementing the overnight gap or the flash crash in any price model requires in both cases the use of financial data with a higher frequency than daily, for example hourly, as we do in this paper.

Our ultimate goal is to provide solid evidence for the conjecture that good stop-loss strategies may provide higher risk-adjusted returns, with respect to a passive buy-and-hold strategy, and for this reason we take the necessary steps to ensure our models have a sufficiently high degree of realism, bringing theoretical innovations on several fronts along the way, namely, in financial modeling and risk-adjusted performance measures.

The rest of this paper is structured as follows: Section 2 presents the general framework for the analysis of stop-loss rules, and the four specific stop-loss rules that we deal with in our simulations. Section 2.1 contains our two new performance metrics, RVaR and RES ratios, plus three other popular measures to assess the impact of applying to a simulated trading period each of the four stop-loss policies considered in our study. Section 3 presents three financial price models, somewhat more sophisticated than (but representative of) the traditional random walk, the ARMA and Regime-Switching models, but without including yet overnight gaps. The intention of this intermediate step, between the traditional models and our proposed models that include overnight gaps, is to provide a comparative analysis of the impact on the performance measures of including some quantification of overnight gaps in the modeling of stock's time series. We make the comparisons by running the same trading simulations on the models with and without consideration of overnight gaps. Section 4 presents the three financial price models that consider overnight gaps, and subsequently include flash crashes. Section 5 contains our experiments and out-of-sample results based on our models for simulating return behavior, using high-frequency financial data. Section 6 briefly describes the stationary bootstrap resampling, and the complementary experiments with this data-driven method for evaluating performance of the four different stop-loss rules considered here. Section 7 concludes.

2 Stop-loss rule and stop-loss policy

In this section we set up the general framework for the analysis of stop-loss rules. We will focus on long-only strategies, and will not take into consideration two major types of market impact costs, namely slippage (difference between order price and execution price) and transaction costs. Regarding slippage, although it is true that in practice there might be a significant negative correlation between equity performance and its bid-ask spread, in our study this is not an issue because we have deliberately selected 30 highly transacted stocks in 2017, which makes them some of the most liquid too, and so it is safe to assume that the bid-ask spread at the time when the stop-loss is triggered will be similar to the bid-ask spread that the investor would encounter at the end of the investment horizon when the position is liquidated under a buy-and-hold passive strategy. Furthermore, in any case, slippage is a complex function of the chosen transaction volume and stochastic market conditions. As for transaction costs, these are irrelevant to our study, because the transaction costs associated with the stop-loss strategies that we consider are exactly the same as those associated with a buy-and-hold strategy, for two reasons: first, because we only consider pure stop-loss policies, i.e. we disregard re-entry rules and therefore both under buy-and-hold and under the stop-loss policy the trading position is opened and closed exactly once; second, because placing a stop-loss order is free, regardless of which broker the investor operates with.

Our formal definition of a stop-loss policy is a simplified version of the stop-loss-re-entry trading scheme of Kaminski and Lo [2014] because we are interested in analyzing how different rules for exiting the market can better prevent us from losing money than just staying in for a period of time irrespective of market movements. However, our definition allows for different criteria for stopping losses, as opposed to Kaminski and Lo [2014] where the only criterion considered by the authors is that cumulative returns reach a threshold, a form of the *fixed percentage barrier* criterion as considered below.

Given an asset A, we consider a *stop-loss rule* as a decision criterion SL_t on the price history of A such that at every time t it takes the value 0 if the criterion is not satisfied or 1 otherwise. We denote by $\nu(SL_t)$ the value (0 or 1) of SL_t. We can formalize SL_t as a boolean formula on the real numbers, arithmetic symbols, order and other relations as needed. However, we prefer to skip this formality and show instead what is meant with an example. Consider the criterion of cumulative returns of a certain asset reaching a given barrier. Let R_t be the cumulative returns at time t, and β the value of the barrier at any time. Then the stop-loss rule in this case is described by the boolean formula:

$$SL_t := R_t \geq \beta$$

This expression takes value $\nu(SL_t) = 1$ if SL_t holds (i.e. at time t, R_t has reached or surpassed the barrier β), or value $\nu(SL_t) = 0$ if SL_t does not hold.

Then, given a time horizon T and an initial time instant $t_o < T$ at which we buy a risky asset A, a *stop-loss policy* is a risk management measure that depending on the value of a stop-loss rule it signals to exit the investment at a certain time t, $t_o \leq t \leq T$, where then we shall sell all of our holdings in A and use the proceeds to go long on a risk-free asset F up until time T. We formalize this idea as follows.

Definition 2.1 *Given a stop-loss rule SL_t, a* ***stop-loss policy*** $\mathbf{s}(\vec{x}_t, \vec{\gamma})$ *for a risky asset A with returns $\{r_t\}$ is a binary wealth-allocation scheme $\{s_t\}$ between A and a risk-free asset F with return $\{r_f\}$, in such a way that either we are fully invested in A ($s_t = 1$) or totally withdrawn from A ($s_t = 0$) and invested in F. Formally, at initial time t_o, $s_{t_o} = 1$, and for $t : t_o < t \leq T$,*

$$s_t = s(\vec{x}_t, \vec{\gamma}) = \begin{cases} 1, & \text{if } \nu(SL_t) = 0 \text{ and } s_{t-1} = 1 \text{ (stay in)} \\ 0, & \text{if } \nu(SL_t) = 1 \text{ and } s_{t-1} = 1 \text{ (exit)} \\ 0, & \text{if } s_{t-1} = 0 \text{ (stay out)} \end{cases} \tag{1}$$

where $\vec{x}_t$ is a vector of known information about A up until time t, and $\vec{\gamma}$ is a vector of parameters that guide the stop-loss rule SL_t.

In our experiments the risk-free asset F is a AAA government bond, for example, a 1-year U.S. government bond. As in [Kaminski and Lo, 2014], in our parametric models we assume that the returns $\{r_t\}$ of the risky asset A under consideration satisfy the following property:

> The expected return μ of A is greater than the risk-free rate r_f, and let $\pi := \mu - r_f > 0$ denote the risk premium of A.

This property simply excludes the perverse case where the stop-loss policy adds value just because the risk-free asset that the investor transfers the capital to has a higher expected return than A. In our parametric models we ensure this property holds by setting the abovementioned rates to appropriate values. The risk-free rate is set to the historical mean return of a 1-year U.S. government bond, which, during the period 1990-2017 has been 3.171%. On the other hand, the expected return of our hypothetical asset is set to the average return of the NYSE during the same period (1990-2017), which is 6.14%. Assuming compounding, this implies an hourly return of $3.942924 * 10^{-5}$.

In our study we implement and analyze the performance of the following four stop-loss rules (SL_t):

Fixed percentage barrier (%): $SL_t := P_t < P_{t-1}^{max}(1 - a)$, where P_{t-1}^{max} is the highest price achieved until time $t - 1$ and the positive quantity $a \cdot 100\%$ is the maximum percentage of that price P_{t-1}^{max} that the investor is willing to lose in its position. Note that this corresponds to a mobile support barrier common among practitioners. The value of a can be adjusted from historical data or be given a constant value based on a certain risk profile. We tune a according to each of the five performance metrics to be considered (see Section 2.1).

Average True Range (ATR): $SL_t := P_t < P_{t-1} - \alpha ATR_d$, where usually $\alpha \in [1.5, 3]$, $d \leq t$ and ATR_d is a crude estimation of daily historical volatility in the recent past

$$ATR_d := \frac{1}{N}\sum_{i=0}^{N} TR_{d-i} \tag{2}$$

with $TR_d := max\{High_d - Low_d, High_d - Close_{d-1}, Close_{d-1} - Low_d\}$, $High_d$, Low_d and $Close_d$ are the maximum, minimum and closing price at time d, and a typical value for N is 14 days. Note that whilst ATR_d is measured on a daily basis, SL_t can be hourly or follow other higher time frequency.

Relative Strength Index (RSI): The RSI is a momentum indicator that compares the magnitude of an asset's recent gains and losses over a specified time period. It is defined as

$$RSI_t(w) := 100 - \frac{100}{1 + RS(w)_t} \in [0, 100] \tag{3}$$

where $RS(w)_t$ is the average gain of up periods divided by the average loss of down periods, during the specified time window of w (usually in between 7 to 14 trading sessions); so if w contains u up (U) and $w - u$ down (D) periods,

$$RS(w)_t := \frac{\frac{1}{u}\sum_{i\in U}(P_i - P_{i-1})}{\frac{1}{w-u}\sum_{i\in D}(P_{i-1} - P_i)}$$

RSI values of 70 or above indicate that a security is becoming overbought or overvalued, and therefore may be primed for a trend reversal or corrective pullback in price. The rationale for this strategy may lie in the belief that too much euphoria can in fact anticipate a change of regime. This behavior is in fact what one observes in major market corrections [Sornette, 2004]. In this case the stop-loss rule is: $SL_t := RSI_t(w) \geq 70$.

Triple Moving Average crossover (MA): A popular exit signal is the triple crossing of short, medium and long-term moving averages (MA). Practitioners use many different MA periods, and one that is common and seems reasonable based on popular experience would be the triplet 5-20-70. Assuming the starting scenario is that the shorter-period MAs are above the longer-period ones, the exit signal is given by the cumulative event of the 5-period MA crossing the 20 and the 70-period MAs, together with the 20-period MA crossing the 70-period MA. If those three events happen, there is a high probability of a continued fall in price. This method's rationale is that longer-term MAs show more of an asset's historical price trend, whereas shorter-term MAs show more the asset's recent price trend, and so if the recent trend crosses the historical trend, we may be witnessing a change of regime. A MA of p periods is defined as

$$MA(p) := \frac{1}{p}\sum_{i=0}^{p-1} C_{t-i}$$

where C_t is the asset's closing price at session t. The stop-loss rule in this case reads, for $a = 5$, $b = 20$ and $c = 70$ days,

$$SL_t := MA(a) < MA(b) \wedge MA(b) < MA(c)$$

That is, the stop-loss rule at any time t is given by two events that must take place simultaneously: the 5-day MA is below the 20-day MA, and the 20-day MA is in turn below the 70-day MA.

2.1 Performance metrics

We will assess the impact of a stop-loss policy (based on each of the four stop-loss rules presented) on investment performance with five metrics. The first one is simply the expected return, and therefore there is no penalty for the risk the investment strategy involves. The remaining four metrics do penalize for higher risk. The first two are well-known risk-adjusted performance measures, and the last two are new proposals we are contributing:

Expected return: The expected return of a risky asset A with returns $\{r_{s_t}\}$ or $\{r_t\}$, depending whether a stop-loss policy is in place or not respectively, is given as usual by $\mathbb{E}[r_{s_t}]$, $\mathbb{E}[r_t]$ respectively. Note that,

$$\begin{aligned}\mathbb{E}[r_{s_t}] &= \mathbb{E}[r_{s_t}|s_t=0]\cdot\mathbb{P}(s_t=0)+\mathbb{E}[r_{s_t}|s_t=1]\cdot\mathbb{P}(s_t=1)\\ &= r_f\cdot\mathbb{P}(s_t=0)+\mathbb{E}[r_{s_t}|s_t=1]\cdot\mathbb{P}(s_t=1)\end{aligned}$$

Because **s** is a dynamic binary wealth-allocation rule, the probabilities above are best understood as the fraction of time in which the investor is either stopped out or invested in the risky asset. When the former occurs, the investor switches to the risk-free asset until the end of the investment horizon.

Sharpe Ratio (SR): The Sharpe Ratio of a stop-loss policy [Sharpe, 1994], is the expected return of the risky asset in excess of the risk-free rate divided by the standard deviation of that risky return, for a given time horizon (usually one year):

$$SR := \frac{\mathbb{E}[r_t]-r_f}{sd(r_t)} \tag{4}$$

The Sharpe ratio is a popular metric in finance to compare asset performance, but it is far from being ideal. Due to the fact that it uses the standard deviation as a measure of risk, this metric treats downside risk the same as upside risk, which in general is not desirable. The following three alternative measures address the issue of penalizing only for downside risk.

Sortino Ratio (SOR): The Sortino Ratio of a stop-loss policy [Sortino, 1994], is the expected return in excess of the minimum return rate acceptable (e.g. the risk-free rate) T divided by the so-called Downside Deviation (DD), for a given time horizon (usually 1 year).

$$SOR := \frac{\mathbb{E}[r_t]-T}{DD} \tag{5}$$

The Downside Deviation is a measure that is conceptually similar to the standard deviation of the returns (as a random variable), although it is not exactly the same because the reference value is not an expected return, but a target return, and its support is restricted to values below target. Mathematically,

$$DD := \sqrt{\int_{-\infty}^{T} (r_t - T)^2 f(r_t) dr_t}$$

where $f(r_t)$ is the density function of the returns in the buy-and-hold strategy.

Return-VaR ratio (*RVaR*): The *RVaR* of a stop-loss policy is the median return (in excess of the risk-free rate) divided by the median return minus the 1-year 5% VaR:

$$RVaR := \frac{median(r_t) - r_f}{median(r_t) - VaR_{5\%}} \tag{6}$$

where $VaR_{5\%} = CDF_{r_t}^{-1}(5\%)$, CDF being the cumulative distribution function of the returns generated by the stop-loss strategy.

Return-ES ratio (*RES*): The *RES* of a stop-loss policy is the median return, in excess of the risk-free rate, divided by the median return minus the 1-year 5% Expected Shortfall:

$$RES := \frac{median(r_t) - r_f}{median(r_t) - ES_{5\%}} \tag{7}$$

where $ES_{5\%} := \frac{1}{5\%} \int_0^{5\%} VaR_\gamma \, d\gamma$, which effectively computes the average VaR_γ with γ ranging from 0 to 5%.

Naturally, under the stop-loss policy $\mathbf{s}$, r_{s_t} replaces r_t in the above definitions.

In both the two new performance metrics that we propose (*RVaR* and *RES*), we choose the median return instead of the expected return for two reasons: first, because the median is a metric that is more robust to outliers than the mean; second, because a denominator involving percentiles (VaR or ES) clearly favors the use of the median instead of the mean; then, the fact that the ratio as a whole becomes easier to interpret if both the numerator and denominator have the same metric justifies the use of the median in the numerator too. The ratio can then be interpreted as follows: the numerator informs about how larger is the median return of the risky investment or strategy compared to the return of a risk-free asset, while the denominator takes some of those merits off by penalizing for downside risk.

One may argue that this numerator is harder to interpret, as the expected return is a concept that most investors are perfectly acquainted with and may be easier to interpret from a utility-theoretic perspective. However, by using the median return we provide a different perspective, which has a powerful interpretation: by using the median return, the numerator of our ratio is informing about the minimum amount by which the return of the risky investment surpasses that of the risk-free asset 50% of the time.

Similarly to the Sortino ratio, our *RVaR* and *RES* ratios exclusively penalize for downside risk, which is an important advantage over the Sharpe ratio. Finally, our *RVaR* and *RES* ratios have at least two advantages over the Sortino ratio. On the one hand, there is less arbitrariness involved in the choice of the value of the parameter involved: in the case of the *RVaR* and *RES* ratios, the parameter involved is the percentile of the distribution of returns to be chosen, and reasonable values are essentially just three (5, 2.5, 1), which contrasts with the much higher discretion involved in the choice of the target return in the case of the Sortino ratio. On the other hand, the denominator of the *RVaR* and *RES* ratios is nonzero with probability one, which is something not guaranteed in the Sortino ratio, as this heavily depends on the choice of the target return, among other factors.

3 Models of asset prices without overnight gap

Traditional models for equity prices or returns, especially in the literature about stop-loss rules, are of the form of a random walk, a pure ARMA model or a regime-switching model (see, for example Acar and Toffel [2001], Kaminski and Lo [2014] and Lo and Remorov [2017]). Therefore, we will base our models on these three. The modeling process consists of two parts conceptually: (1) modeling the standard deviation of returns, including white noise; (2) modeling the actual return.

Our analysis is based on 30 NYSE stocks among the most liquid in 2017. The selection of these stocks has been somewhat arbitrary, subjected mostly to the availability of data in a common period, presenting the highest liquidity possible, and representing different industries The list of tickers of our selected set of stocks is {BAC, GE, PFE, S, F, C, T, JPM, WFC, HPQ, KO, AMD, MRK, XOM, JCP, ABX, RAD, GLW, VZ, JNJ, AIG, PG, DIS, HAL, XRX, KEY, BMY, SCHW, ABT, MO}.

On the one hand, we find that 30 stocks provides enough diversity so that the results of the analysis are reliable. On the other hand, extending the analysis to considerably more stocks, e.g. 200 stocks, might require stronger assumptions about how narrow the bid-ask spread can be expected to be at the moment the stop-loss is triggered, because those 200 stocks would necessarily include some stocks that are not as liquid as the 30 stocks considered in this paper. For these simulations, we consider hourly quotes from March 13 to March 24 2017 (both included), for each of the 30 stocks, which equates to 2070 hourly returns. Although a longer period might a priori be desirable, we still believe the quantity of data (2070 returns) is sufficient given that the frequency of the observations is hourly. Last but not least, the length of the period has also been influenced by the availability of high-frequency data that are both reliable and affordable to obtain.

3.1 Modeling standard deviation and white noise

We modeled the standard deviation of the hourly returns with a GARCH(1,1) model. This requires that we model the white noise component as well. The GARCH(1, 1) model for the standard deviation σ_t, at each time t, is defined by the following equation

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2\sigma_{t-1}^2 + \beta\sigma_{t-1}^2$$

where ϵ_t is the random shock or innovation, occurring at time t, and which itself follows a certain conditional distribution that we must specify and calibrate with the observed data.

In order to fit the best GARCH(1,1) possible, we try each of the implemented conditional distributions in the package **fGarch** [Wuertz et al., 2016] in R, fit the model, record its corrected Akaike Information Criterion (AIC), and eventually choose the conditional distribution (for the white noise) that provides the best (i.e. lowest) AIC on average over our 30 selected stocks. We find that the best conditional distribution is the Generalized Error Distribution with or without skewness parameter (SGED or GED, respectively), depending on whether we consider a model with generalized normally distributed returns or a model that considers ARMA returns instead (see next subsections).

A point worth advancing is that in both models the shape coefficient of the above conditional distribution is around 1.45, which is almost exactly half way between that of a Laplace distribution (shape=1) and a Gaussian distribution (shape=2), the former having a much larger kurtosis than the latter.

3.2 Generalized Normally Distributed returns (GED)

Our first model for the hourly price of a hypothetical NYSE stock considers that the corresponding returns follow a Generalized Normal distribution. Recall that the PDF of the Generalized Normal (or Error) distribution is given by

$$GED(\mu, \sigma, \beta) = \frac{\beta}{2\sigma\Gamma(1/\beta)} e^{-(|x-\mu|/\sigma)^{\beta}}$$

where μ, σ, β are the location, scale, and shape parameters respectively, and Γ denotes the Gamma function. As Nadarajah [2005] showed, the PDF of a Generalized Normal distribution is just the same as the PDF of a Normal distribution but with a general β coefficient that is not necessarily equal to 2 (and hence the absolute value function is needed), as well as a different regularization term premultiplying the exponential function, so that the PDF integrates to unity.

After a model selection procedure to fit the parameters of the distributions involved, we obtain the following model to simulate the hourly price of our hypothetical NYSE stock.

$$P_{t-1} * (1 + r_t)$$

where

$$\begin{aligned} r_t &\sim GED(3.943 * 10^{-5}, 0.0042, 1.326), \\ \sigma_t^2 &\sim GARCH(1,1) \text{ with } \epsilon_t \sim SGED(\mu = 0, \sigma = 1, 1.4, 0.928) \end{aligned} \tag{8}$$

3.3 ARMA returns

Our second model for the price of a hypothetical stock that trades in the NYSE considers that the returns follow an ARMA(p, q) process, with GARCH volatility. After the model selection procedure we obtain the following model for the hourly price of our stock:

$$P_{t-1} * (1 + r_t)$$

where

$$r_t \sim ARMA(3,0),\ \sigma_t^2 \sim GARCH(1,1) \text{ with } \epsilon_t \sim GED(\mu = 0, \sigma = 1, 1.47) \tag{9}$$

3.4 Regime-Switching

The relevance of a Regime-Switching model for analysing stop-loss rules has been discussed extensively in Kaminski and Lo [2014], and we refer the reader to that paper, from which we borrow the essential characteristics of this model. In brief, the importance of this model to the problem we have at hand is its capability to capture changes in regime, which is a common underlying motivation for setting up stop-loss rules. A regime-switching model for return $\{r_t\}$ has the form:

$$r_t = I_t r_{1t} + (1 - I_t) r_{2t}, \quad r_{it} \sim N(\mu_i, \sigma_i^2),\ i = 1, 2 \tag{10}$$

$$A := \begin{array}{c|cc|} & I_{t+1} = 1 & I_{t+1} = 0 \\ \hline I_t = 1 & p_{11} & p_{12} \\ I_t = 0 & p_{21} & p_{22} \\ \hline \end{array}$$

where

$$I_t = \begin{cases} 1 & \text{if state 1 prevails,} \\ 0 & \text{if state 2 prevails} \end{cases}$$

and A is the transition probability matrix that governs the transitions between the two states. The model has six parameters: the means and variances plus transition probabilities $(\mu_1, \mu_2, \sigma_1, \sigma_2, p_{11}, p_{22})$, and we fit it to data using the Baum-Welch algorithm to fit Hidden Markov Models [Welch, 2003].

4 Models of asset prices with overnight gap and flash crash

Because we are introducing overnight gaps, the modeling process consists now of three parts: (1) modeling the overnight gap; (2) modeling the standard deviation of returns, including white noise; (3) modeling the actual return.

Part (2) of the modeling process will be the same as for the models without overnight gaps, since we want to consider the same modeling for the noise, and assess the impact of considering overnight gaps. Thus, we only describe the modeling of overnight gaps and their inclusion in the model for the returns.

4.1 Modeling overnight gaps

An overnight gap is defined as the difference between the Open price at day t, and the Close price at day $t-1$. Using our empirical data on 204,120 gaps (27 years worth of data on 30 NYSE stocks), we find that the average gap is about 1.000463 (i.e. +0.0463%). Furthermore, on average, gaps of any size above a minimum threshold of 0.1% happen around 77.93% of the days, and this is the frequency we will use in our simulations.

Including overnight gaps into the analysis of stop-loss rules is of critical importance, for two main reasons. First, overnight gaps can severely affect the profitability of an investing strategy that uses stop-loss orders, because they render the stop-loss barrier completely unable to protect the investor from a loss that is substantially larger than expected. In particular, this implies that by disregarding the phenomenon of overnight gaps one might inflate the usefulness of a stop-loss rule, and this is something we wish to avoid. Second, overnight gaps are a very frequent phenomenon in the financial markets, and their magnitude can be non-negligible. For example, Figure 1 shows several overnight gaps in the candlestick plot of the price history of the American International Group (AIG:NYSE) from 28th of August to 22nd of September, 2017. In this plot on 28th August the Close price is 60.66 and next session Open price (on 29th August) is 60.20, a fall of 0.7%; on 8th Sept. (a Friday) the Close is 59.78 and next session on 11th Sept. the Open is 60.77, a jump of 1.6%. All in all, by considering

Figure 1: AIG candlestick plot of price from 28/08 to 22/09 showing overnight gaps (source: *yahoo finance*).

overnight gaps we obtain a more reliable assessment of stop-loss performance.

We modeled overnight gaps as a Weibull random variable for the following reasons. We think of overnight gaps as extreme price movements, and the Weibull is one of the three families of extreme value distributions. Moreover, a convenient way to reproduce the impact of overnight gaps in the price of a stock is to see the gap as a scaling factor of the price, and so we need a distribution with non-negative support, as it is the Weibull, because the scaled stock price must naturally remain non-negative. Finally, the Weibull distribution is very flexible, capable of showing a completely different density plot depending on the value of its two parameters: scale (λ) and shape (κ).

Recall that the probability density function (PDF) of a Weibull random variable is

$$Weibull(x;\lambda,\kappa) = \begin{cases} \frac{\kappa}{\lambda}\left(\frac{x}{\lambda}\right)^{\kappa-1} e^{-(x/\lambda)^{\kappa}}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

where $\lambda > 0$ is the scale parameter and $\kappa > 0$ is the shape parameter of the distribution. When the random variable is clear from context it is customary to write $Weibull(\lambda, \kappa)$. We follow a 3-step procedure to optimize the Weibull parameters:

1. Explore seven estimation methods to find a first approximation: Method of Moments (MoM), Maximum Likelihood Estimation (MLE), Median and Quartiles (MQ), Mean and Standard Deviation (MSD), Mean Rank (MR), Median Rank (MDR), and Symmetric CDF (SCDF).

2. Select the best estimation method and do a grid search around the estimated parameters to improve *Stress*. *Stress* is a measure of goodness of fit that we define as a convex combination of the D statistic from the Kolmogorov-Smirnov test and the Mean Squared Error, in proportion 3:7 as the latter metric is more robust.

3. Correction for Survivorship Bias: the 30 stocks we have picked are the ones that have "survived" from 1990 to 2017, and so it is advisable to be more conservative in asserting the efficacy of stop-loss rules. This is achieved by increasing the downside risk of the overnight gap, which in the case of a Weibull random variable this may be achieved by reducing the shape parameter of this distribution. We perform this step with care: we require that the new calibration delivers an average gap that is more negative than the average gap observed empirically, but with the least increase in Stress possible with respect to Step 2.

Table 1: Weibull parameter calibration

Method	Step 1 λ, κ, Stress	Step 2 λ, κ, Stress	Step 3 λ, κ, Stress
MoM	1.0063, 98.56, 0.0677	–	–
MLE	1.1743, 20.16, 0.4790	–	–
MQ	1.0018, 203.62, **0.0385**	1.0028, 189.32, 0.0244	1.0033, 170.72, 0.0363
MSD	1.0055, 112.53, 0.0586	–	–
MR	1.0075, 81.35, 0.0812	–	–
MDR	1.0075, 81.47, 0.0811	–	–
SCDF	1.0075, 81.49, 0.0811	–	–

As it can be observed in Table 1, the value of the optimized parameters according to the three steps above are $\lambda = 1.0033$, and $\kappa = 170.7193$, delivering a Stress level of 0.0363.

4.2 Generalized Normally Distributed returns with gap (GEDgap)

In this subsection we introduce overnight gaps into the first parametric model presented earlier. Again, after a model selection procedure to fit the parameters of the distributions involved, we obtain the following model to simulate the hourly price of our hypothetical NYSE stock.

$$P_t = \begin{cases} P_{t-1} * Gap_t, & \text{for } t \bmod 7 = 1 \\ P_{t-1} * (1 + r_t), & \text{otherwise} \end{cases} \tag{12}$$

where the Gap_t occurs daily (t mod 7 = 1) with a certain probability (again, in our case, 77.93%)

$$Gap_t = \begin{cases} 1, & \text{if no gap occurs} \\ Weibull(1.0033, 170.7193), & \text{if a gap occurs} \end{cases} \tag{13}$$

and

$$\begin{aligned} r_t &\sim GED(3.943 * 10^{-5}, 0.0042, 1.326), \\ \sigma_t^2 &\sim GARCH(1,1) \text{ with } \epsilon_t \sim SGED(\mu = 0, \sigma = 1, 1.4, 0.928) \end{aligned} \tag{14}$$

As it can be seen, the value of the parameters is the same compared to the model without gap. This is consistent for two reasons. First, because the Weibull distribution is approximately centered at unity. Second, because the impact of overnight gaps on return volatility is carefully taken into account at a later stage (see Section 4.5).

4.3 ARMA returns with gap (ARMAgap)

In this case we are introducing the overnight gap into our second parametric model.

After the model selection procedure we obtain the following model for the hourly price of our stock:

$$P_t = \begin{cases} P_{t-1} * Gap_t, & \text{for } t \bmod 7 = 1 \\ P_{t-1} * (1 + r_t), & \text{otherwise} \end{cases} \tag{15}$$

where Gap_t is defined as in Equation 13, and

$$r_t \sim ARMA(3,0),\ \sigma_t^2 \sim GARCH(1,1) \text{ with } \epsilon_t \sim GED(\mu = 0, \sigma = 1, 1.47) \tag{16}$$

4.4 Regime-Switching returns with gap (RSgap)

Similar to the previous two cases, we model prices also by the process defined in (15) to account for the overnight gaps, while the market hours returns r_t are modelled by the regime-switching model (Equation 11).

4.5 Incorporating post-gap effect on volatility

In addition, we consider that a stock's volatility at time t is a function of the magnitude of the most recent gap, and so we attempt to model this relationship in order to obtain an even more realistic model.

With this in mind, we fit a polynomial in the gap magnitude to the *relative range* (a proxy for volatility to be made precise below). We observed that terms of order higher than three of the variable "gap magnitude", although significantly different from 0 in a statistical sense were practically insignificant in magnitude (in the order of 10^{-6}). For this reason, we propose the following model for the impact of the gap magnitude on the relative range:

$$RelRange = \beta_0 + \beta_1 gm + \beta_2 gm^2 + \beta_3 gm^3 + u \tag{17}$$

where gm is the gap magnitude, and the relative range of the stock on a given day, and the relative range, referred to as $RelRange$, is defined as $\frac{Range}{Low} = \frac{High - Low}{Low}$.

Next, we compute the average relative range for our stocks and compute the relative betas, which are the original betas divided by the average relative range. Then, the relative betas are averaged across stocks. From here, we obtain two standard deviation multipliers:

$$NGM = \overline{rel.beta_0}$$

and

$$GM = \overline{rel.beta_0} + \overline{rel.beta_1} * gm + \overline{rel.beta_2} * gm^2 + \overline{rel.beta_3} * gm^3$$

where the horizontal bars signal that the *average* relative beta has been considered, while NGM and GM are the standard deviation multipliers when there is no gap and when there is, respectively. Those quantities squared multiply the GARCH-based variance that the models produce at a given time t, displayed in Equation 14 and Equation 16 respectively. In particular, we find that, in expectation

$$NGM = 0.80 \tag{18}$$

and

$$GM = 0.80 + 0.25gm + 0.06gm^2 - 0.01gm^3 \tag{19}$$

Considering that the price gap is a real number generally close to unity, we find that a higher price gap in magnitude is associated with higher volatility, as measured by the relative range.

4.6 An additional layer of realism: introducing flash crashes

A flash crash is a rapid, deep, and volatile fall in the price of a security, followed by a rebound, occurring within an extremely short period of time.

Notable examples include:

- Procter & Gamble, May 6, 2010. A large mutual fund sold an unusually large number of E-Mini S&P futures, which first exhausted available buyers, and then high-frequency traders started aggressively selling those futures, accelerating the effect of the mutual fund's selling and contributing to the sharp price declines that day. After a significant decline in the E-Mini S&P 500 futures, Procter & Gamble's price declined by 37%.

- The Dow Jones Industrial Average, April 23, 2013. The Dow Jones momentarily dropped 1.5% due to a tweet about fictional attacks in the White House that left President Barack Obama injured.

- EUR/CHF, January 15, 2015. The Swiss National Bank announced without prior notice that it would suddenly remove the 1.2000 self-imposed floor on the EUR/CHF exchange rate, causing a flash crash that lead to a historic dysfunction never seen before in the global Forex market.

- EUR/USD, March 18, 2015. The EUR/USD futures plunged 3% in less than four minutes for no clear reason, and most of the loss was recovered shortly after.

- S&P 500, August 24, 2015. A (non flash) crash of the Chinese Shanghai Composite Index (-8.5%) undermined the already weak confidence in the US markets, and led to a rapid 5% decline in the S&P 500, that was recovered almost in full shortly after.

- HSBC, September 18, 2015. A "fat-finger" trade, exacerbated by automatic trading systems, caused HSBC share price to suddenly drop 4.8%, rebounding moments later.

- GBP/USD, October 6, 2016. The GBP/USD suddenly dropped by 6% due to a combination of complex trading positions, inexperienced traders, and algorithmic trading.

Flash crashes are included in the price models with gap based on the few (but important) historical examples witnessed in the last decades. As in the case of the overnight gap, one must distinguish two related random variables. On the one hand, the occurrence of a flash crash is modeled as a Bernoulli random variable with probability 0.05% per hour (recall we simulate high-frequency quotes), which yields an expected number of such crashes of about one per year - enough to see a clear impact on the models, but still in touch with reality. On the other hand, we have the magnitude of the flash crash (conditional on its occurrence), which we model as a uniform random variable in the interval 5% - 35%, signifying the price drop percentage caused by the flash crash. Given the few historical examples and the lack of a well-accepted theoretical work on the figure of flash crashes, we opted for a simple, flat prior with bounds given, approximately, by the second smallest and largest flash crashes witnessed (respectively: HSBC, September 2015, -4.8%, and Procter & Gamble, May 2010, -37%).

Then, based on the above parameters, flash crashes randomly appear in our high-frequency simulations, and then the price is completely recovered in the next hour, embodying the very definition of a flash crash. This inclusion, therefore, has no impact on the expected return (under buy-and-hold) but volatility is naturally bound to increase.

5 Model-based simulation

We present in this section our experimental results for the different model-based simulations of the four stop-loss rules considered: the fixed percentage, ATR, RSI and triple MA crossover rules. First we specify the choice of parameters for each of the stop-loss rules.

5.1 Parameter choice for the fixed percentage rule

As seen in Section 2, we implement the following stop-loss:

$$SL_t = P_{t-1}^{max}(1-a)$$

where $a > 0$ is the maximum percentage of the highest price achieved until time $t-1$ that the investor is willing to lose in that operation.

There seems not to be a consensus among practitioners regarding the most suitable choice of value for the parameter of this stop-loss rule. Hence, instead of arbitrarily choosing the parameter's value that determines how tight the exit barrier is placed around a reference price, we decided to tune the parameter to a value that delivers a good out-of-sample stop-loss performance. In order to accomplish this, we have a range in the interval $[0.03, 0.1]$ and select the value that improves the performance of the rule the most on average, in terms of expected return, Sharpe Ratio, Sortino Ratio, Return-VaR ratio, and Return-ES ratio.

5.2 Parameter choice for the ATR, RSI and triple MA crossover rules

On the other hand, there seems to be a consensus among practitioners regarding suitable parameter values for these three other stop-loss rules, and for that reason we have chosen not to tune those parameters and instead use some of the standard values used in practice. We then consider:

- ATR

$$SL_t = P_{t-1} - 2.5ATR_d$$

where ATR_d, for $d \leq t$, is the estimation of daily volatility given in Equation 2, with a time frame of 14 days, as is standard practice.

- RSI

$$SL_t := RSI_t(7) \geq 70$$

where $RSI_t(7)$ is given by Equation 3. As shown above, we decided to use the standard value of 70 as a threshold signalling that the stock has become overbought and a price reversal may follow soon. The chosen time frame of 7 sessions is also common among practitioners.

- Triple MA crossover

$$SL_t := MA(5) < MA(20) \wedge MA(20) < MA(70)$$

We use standard time frames of 5, 20, and 70 trading days. These figures are, however, modified in our non-parametric model to accommodate for the shorter period considered (due to constraints on high-frequency data availability), and they become instead 3, 12, and 30 trading days.

5.3 Simulation results

Next we present the results of our Monte Carlo simulations, based on 5,000 repetitions for each stop-loss rule. Because we repeat each set of simulations 50 times, we are able to use the Welch's t-test to assess the statistical significance of the difference in mean values that the stop-loss rules provide compared to Buy-and-Hold. We use the following notation to denote statistical significance:
ns: p-value $\geq$ 0.05, *: 0.01 $\leq$ p-value $<$ 0.05, **: 0.001 $\leq$ p-value $<$ 0.01, ***: p-value $<$ 0.001

Tables 2, 3, 5, and 6 show the out-of-sample results for two of the four stop-loss rules, comparing stop-loss performance both without and with overnight gaps. Results for the other two rules are presented in the Appendix (Tables 11, 12, 14 and 15).

As a general conclusion, both in the models without and with gap, we observe that all stop-loss rules provide a higher risk-adjusted return than Buy-and-Hold (B& H). The results are particularly remarkable for the fixed percentage rule with returns following a Regime-Switching (RS) model with gap, but no flash crash (Table 3): Sharpe Ratio is approximately multiplied by a factor of three, Sortino is multiplied by eight, while RVaR and RES are roughly multiplied by seven.

Note that in a market with a positive risk premium (with expected annual return of around 5% in our models including overnight gaps), in the case without flash crashes, stop-loss rules seem to provide a slightly lower expected return in absolute terms, but this difference is not statistically significant and, in any case, it is on average small (usually less than 100 bps). Once flash crashes are included in the models, expected

Table 2: Fixed percentage SL, without gap: Simulation results

Strategy-Model	$\mathbb{E}[r]$	Sharpe	Sortino	RVaR	RES
SL - GED	4.57%	0.81	2.02	0.85	0.75
B&H - GED	7.09%	0.32	0.16	0.16	0.14
Difference	-2.52% ns	+0.49***	+1.86***	+0.69***	+0.61***
SL - ARMA	4.51%	0.82	2.09	0.86	0.76
B&H - ARMA	6.82%	0.34	0.17	0.18	0.15
Difference	-2.31% ns	+0.48***	+1.92***	+0.68***	+0.61***
SL - RS	3.66%	0.82	3.20	0.88	0.80
B&H - RS	6.26%	0.29	0.18	0.14	0.12
Difference	-2.6% ns	+0.53***	+3.02***	+0.74***	+0.68***

Table 3: Fixed percentage SL, with gap: Simulation results

Strategy-Model	$\mathbb{E}[r]$	Sharpe	Sortino	RVaR	RES
SL - GEDgap	4.04%	0.81	1.23	0.84	0.69
B&H - GEDgap	5.01%	0.29	0.13	0.15	0.12
Difference	-0.97% ns	+0.52***	+1.1***	+0.69***	+0.57***
SL - ARMAgap	4.06%	0.81	1.21	0.83	0.69
B&H - ARMAgap	4.93%	0.30	0.13	0.16	0.13
Difference	-0.87% ns	+0.51***	+1.08***	+0.67***	+0.56***
SL - RSgap	3.99%	0.80	1.02	0.84	0.72
B&H - RSgap	5.93%	0.26	0.13	0.11	0.09
Difference	-1.94% ns	+0.54***	+0.89***	+0.73***	+0.63***

Table 4: Fixed percentage SL, with gap and flash crash: Simulation results

Strategy-Model	$\mathbb{E}[r]$	Sharpe	Sortino	RVaR	RES
SL - GEDcrash	2.34%	0.41	0.07	0.57	0.12
B&H - GEDcrash	5.03%	0.29	0.12	0.15	0.12
Difference	-2.69% *	+0.12*	-0.05 ns	+0.42***	+0.00 ns
SL - ARMAcrash	2.27%	0.39	0.07	0.55	0.11
B&H - ARMAcrash	4.90%	0.29	0.13	0.16	0.12
Difference	-2.63% ns	+0.1 ns	-0.06 ns	+0.39***	-0.01 ns
SL - RScrash	2.70%	0.52	0.07	0.74	0.18
B&H - RScrash	5.54%	0.24	0.11	0.10	0.09
Difference	-2.84% *	+0.28***	-0.04 ns	+0.64***	+0.09*

Table 5: RSI-based SL, without gap: Simulation results

Strategy-Model	$\mathbb{E}[r]$	Sharpe	Sortino	RVaR	RES
SL - GED	6.27%	0.33	0.14	0.22	0.19
B&H - GED	7.04%	0.32	0.15	0.16	0.13
Difference	-0.77% ns	+0.01 ns	-0.01 ns	+0.06***	+0.06***
SL - ARMA	6.47%	0.35	0.14	0.22	0.18
B&H - ARMA	6.87%	0.34	0.15	0.18	0.15
Difference	-0.40% ns	+0.01**	-0.01*	+0.04**	+0.03***
SL - RS	5.78%	0.30	0.15	0.19	0.17
B&H - RS	6.27%	0.29	0.16	0.14	0.12
Difference	-0.49% ns	+0.01 ns	-0.01*	+0.05***	+0.05***

Table 6: RSI-based SL, with gap: Simulation results

Strategy-Model	$\mathbb{E}[r]$	Sharpe	Sortino	RVaR	RES
SL - GEDgap	4.55%	0.31	0.12	0.24	0.20
B&H - GEDgap	5.03%	0.29	0.13	0.15	0.12
Difference	-0.48% ns	+0.02***	-0.01*	+0.09***	+0.08***
SL - ARMAgap	4.57%	0.32	0.11	0.23	0.19
B&H - ARMAgap	4.94%	0.30	0.12	0.16	0.13
Difference	-0.37% ns	+0.02***	-0.01*	+0.07***	+0.06***
SL - RSgap	5.71%	0.28	0.10	0.19	0.16
B&H - RSgap	6.25%	0.27	0.11	0.12	0.10
Difference	-0.54% ns	+0.01***	-0.01*	+0.07***	+0.06***

Table 7: RSI-based SL, with gap and flash crash: Simulation results

Strategy-Model	$\mathbb{E}[r]$	Sharpe	Sortino	RVaR	RES
SL - GEDcrash	4.50%	0.31	0.11	0.23	0.19
B&H - GEDcrash	4.95%	0.29	0.12	0.14	0.12
Difference	-0.45% ns	+0.02 ns	-0.01 ns	+0.09 ns	+0.07 ns
SL - ARMAcrash	4.63%	0.32	0.12	0.23	0.19
B&H - ARMAcrash	4.99%	0.30	0.12	0.16	0.13
Difference	-0.36% ns	+0.02 ns	-0.00 ns	+0.07 ns	+0.06 ns
SL - RScrash	5.19%	0.25	0.10	0.17	0.14
B&H - RScrash	5.55%	0.24	0.11	0.11	0.09
Difference	-0.36% ns	+0.01 ns	-0.01 ns	+0.06 ns	+0.05 ns

return under most {stop-loss rules, price model} pairs drops significantly as the flash crash triggers the stop. Important exceptions to this general behaviour are observed in the case of the MA crossover (Table 16) and RSI (Table 7) - based stop-loss rules, under all price models. In other words, these two stop-loss rules have shown to be robust to flash crashes, which is an intuitive result, perhaps especially so for the MA crossover stop-loss, as a momentary sharp (yet realistic) price drop is unlikely to trigger the crossing of the short, medium and long-term moving averages, provided a sufficiently large time window, as it seems to be our case (5, 20 and 70 days). Further note that the usefulness of each stop-loss rule is apparent in all the considered models for returns, being the fixed percentage stop-loss the generally most useful strategy in risk-adjusted terms, even in the presence of flash crashes.

Nonetheless, it is worth pointing out that stop-loss performance – with perhaps the exception of the fixed percentage rule – is in general somewhat worse under the no-gap models. This fact may seem counterintuitive at first because one may reasonably think that since large negative overnight gaps are able to bypass stop-loss rules, rendering the investor unable to liquidate his/her position at the desired price level, a model that does not include this kind of phenomena should inflate stop-loss performance. However, this reasoning relies on the inaccurate assumption that large negative overnight gaps are sufficiently frequent. In fact, large negative overnight gaps are much less common than what it might appear; it is just that because when they occur they attract considerable attention, individuals tend to overestimate their frequency or probability (*availability heuristic*). On the other hand, overnight gaps small in magnitude (negative or positive) are relatively frequent, and it seems that some of the stop-loss rules considered are able to avoid the negative ones while benefiting from the positive ones, improving risk-adjusted performance.

By not including overnight gaps when modeling stock returns, it seems that one is not getting the most out of stop-loss rules. This, together with our use of high-frequency data, more complex price models and alternative risk-adjusted performance metrics, might explain why the literature on stop-loss rules, which at present does not consider overnight gaps, has not found a significant performance improvement in the use of stop-loss rules, despite the fact that in practice they are considered an extremely effective tool for loss protection, in particular, and risk management, in general.

Therefore, since real financial markets quite frequently exhibit overnight gaps, and assuming that our overnight gap specification is sufficiently accurate, we believe that the results we obtained with the models that do include gaps are closer to the true stop-loss performance than those obtained under the no-gap models.

For the ATR and MA based stop-loss rules, we obtained again statistically significant risk-adjusted performance improvements, albeit more moderate. See Tables 11, 12, 14, and 15 in the Appendix.

6 Non parametric approach: the stationary bootstrap.

There is no one universally accepted model for asset returns, and whatever this model might be the danger of overfitting is always present. In light of those reasons, we complement our analysis with a non parametric approach to add robustness: the stationary bootstrap. The stationary bootstrap [Politis and Romano, 1994] is one of the existing block bootstrap methods, which is characterized by using random block lengths that are distributed according to the geometric distribution. Provided that the original data is stationary, the use of geometrically distributed block lengths ensures that the resampled series remain stationary. By using an appropriate scenario generator for financial time series such as the stationary bootstrap, our backtesting is more reliable than if we just used the observed prices, as these are just one possible realization of an underlying stochastic process.

The stationary bootstrap algorithm goes as follows (cf. [Davison and Hinkley, 1997, Ch. 8]): Given $X_1, \ldots, X_n$ observations of a time series, wrap this data around a circle; that is, define $Y_i = X_{1+(i-1 \mod n)}$ and $Y_0 = X_n$. Let $I_1, I_2, \ldots \in \{1, \ldots, n\}$ be drawn iid with uniform distribution (these are the starting points of blocks). Construct a bootstrap sample as follows:

- set $Y_1^* = Y_{I_1}$.
- For $i = 2, \ldots, n$, let $Y_i^* = Y_{I_i}$, with probability p, and let $Y_i^* = Y_{j+1}$ with probability $1 - p$, where $Y_{i-1}^* = Y_j$.

The output is a series $Y_1^*, \ldots, Y_n^*$, composed of blocks with mean length $1/p$, since the lengths $L_1, L_2, \ldots$ of attached blocks through the iteration follows a geometric distribution: $\mathbb{P}(L = k) = p(1-p)^{k-1}$, $k = 1, 2, \ldots$.

In our stationary bootstrap sampling of time series data we make use of the automatic block-length selection heuristic laid out in [Politis and White, 2004]. In particular, since $\mathbb{E}[L] = \frac{1}{p} = b$ we immediately have that $p = \frac{1}{b}$, and b is obtained using the heuristic in question. Therefore, p is determined by the automatic block-length selection heuristic mentioned before.

Also, to be rigorous with the application of the stationary bootstrap, we checked for weakly stationarity in the observed data, applying traditional tests for unit root (Augmented Dickey Fuller, KPSS, Phillips-Perron) and a more advanced test, namely, the Priestley-Subba Rao (PSR) test for nonstationarity [Priestley and Rao, 1969], which is based upon examining how homogeneous a set of spectral density function (SDF) estimates are across time, across frequency, or both. Thus, we determine that at the 1% significance level, only 11 stocks (BAC, PFE, F, C, JPM, MRK, XOM, AIG, HAL, KEY, and BMY) out of our original universe of 30 present weakly stationary hourly returns in the period considered: March 13 to May 26, 2017. To be noted that this period is longer than the one considered in our parametric approach. This fact is due to two reasons. On the one hand, we believe that a data-based model – as opposed to a purely parametric model in which having longer periods of data might actually be counter-productive when calibrating the parameters of the model – should include as many observations as possible, which justifies a larger resource expenditure to obtain data until the end of May 2017. On the other hand, since we only consider 11 stocks in this case (as the other 19 are not weakly stationary and thus not suitable for the stationary bootstrap), we would wish to compensate for that potential loss of information by extending the period under study.

Finally, we note that under this nonparametric approach, our price model is

$$P_t = P_{t-1}(1 + r_t^*)$$

for all $t \geq 1$, where r_t^* is the bootstrapped return at time t. Figure 2 shows real price time series (dashed, black line) and bootstrapped versions (continuous, red line) for Citigroup and American Int. Group.

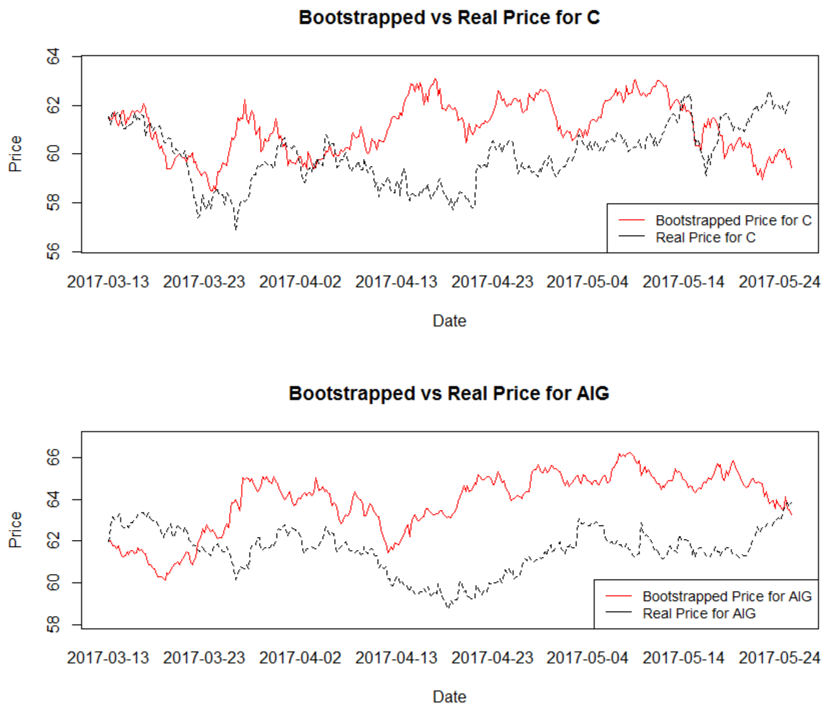

Figure 2: Citigroup (C) and AIG bootstrapped price and real price time series.

Table 8: Bootstrap simulation results under Buy-and-Hold

Stock	$\mathbb{E}[r]$	Sharpe	Sortino	RVaR	RES
BAC	-7.1794	-	-	-	-
PFE	-5.6818	-	-	-	-
F	-12.8653	-	-	-	-
C	1.5171	0.1578	0.0594	0.0741	0.0603
JPM	-6.1557	-	-	-	-
MRK	-0.3745	-	-	-	-
XOM	0.1034	0.0182	0.0074	-	-
AIG	3.3076	0.4296	0.1821	0.2527	0.2053
HAL	-9.5564	-	-	-	-
KEY	-2.8110	-	-	-	-
BMY	-7.1723	-	-	-	-

Table 9: Bootstrap simulation results under fixed % SL

Stock	$\mathbb{E}[r]$	Sharpe	Sortino	RVaR	RES
BAC	-0.6170	-	-	-	-
PFE	-1.9672	-	-	-	-
F	-1.6234	-	-	-	-
C	1.2653	0.1405	0.1205	-	-
JPM	-0.9784	-	-	-	-
MRK	-0.0666	-	-	-	-
XOM	0.1272	0.0373	0.0408	-	-
AIG	3.1447	0.4063	0.3044	0.2378	0.2248
HAL	-0.7096	-	-	-	-
KEY	-0.1206	-	-	-	-
BMY	-1.3588	-	-	-	-

Table 10: Bootstrap simulation results under RSI SL

Stock	$\mathbb{E}[r]$	Sharpe	Sortino	RVAR	RES
BAC	-6.9053	-	-	-	-
PFE	-5.6759	-	-	-	-
F	-12.6731	-	-	-	-
C	1.4081	0.1656	0.0560	0.0899	0.0735
JPM	-5.9739	-	-	-	-
MRK	-0.3390	-	-	-	-
XOM	0.1185	0.0226	0.0064	0.0082	0.0066
AIG	3.0064	0.4306	0.1616	0.2746	0.2244
HAL	-9.2652	-	-	-	-
KEY	-2.6207	-	-	-	-
BMY	-7.1070	-	-	-	-

Importantly, it must be stressed that in the period March 13 to May 26, all the stocks under consideration, except Citigroup (C) and American International Group (AIG), experienced a negative return, which explains why most of the expected returns are negative, especially under Buy-and-Hold. Because the metrics besides expected return that we have considered are meaningless when below zero, we restrict our attention to expected return for those stocks and leave blank those entries below zero. In such a situation in which 9 out of the 11 stocks have experienced a negative return in the period considered, the stop-loss rules have provided a higher expected return, in some cases extremely higher than if no stop-loss had been used, as in the case of the fixed percentage strategy, which has achieved a positive expected return in 5 out of 11 stocks, and near positive in two more stocks. In fact, on average, the fixed percentage rule has increased the expected return in over 4 percentage points compared to Buy-and-Hold, and in some cases (Ford), as much as 11.3 percentage points higher than Buy-and-Hold, on average (in 200,000 repetitions), in *just* two and a half months. In a full year, the improvement might be even more spectacular. This is the opposite case as in our model-based simulation: in the present case, using real (bootstrapped) prices, because mid March to late May has been a bad period for most of the stocks under consideration, it seems indeed natural that a strategy that at some point switches to a risk-free asset will perform better than merely Buy-and-Hold. However, three aspects must be stressed: first, the risk-free rate used in the bootstrap simulations is the 3-month U.S. Treasury bill rate that was in place in March 2017, just 0.74%; second, the fact that the remarkable results shown in Table 9 for the falling stocks are, again, obtained out of sample; lastly, in the case of the fixed percentage rule, note that the expected return premium over Buy-and-Hold in a Bear market is much higher than the expected return premium of Buy-and-Hold over this stop-loss rule in a Bull market, in general.

On the other hand, for C and AIG, since most of the metrics are above 0, we can conduct a similar analysis as in Section 5. In this case the results are favorable too, albeit less spectacular: certain stop-loss rules seem to provide a modest improvement in risk-adjusted returns, at the expense of a very slight drop in absolute expected return. For example, as shown in Table 10 we may highlight the improvement in *all* risk-adjusted metrics that the RSI-based stop-loss provides for the two rising stocks (C and AIG).

The simulation results for the other two stop-loss rules (ATR and MA) can be found in the Appendix at the end of this work (Tables 17 and 18).

7 Conclusions

Combining three different parametric price models, together with a nonparametric approach based on the stationary bootstrap as a scenario generating method, our out-of-sample research shows that irrespective of the market situation (a rising or falling market), stop-loss rules are able to provide a higher expected return than buy-and-hold, at least when adjusting for risk, and this difference is statistically significant at the 0.1% level in most of the cases. On the one hand, in the case of a market with a positive risk premium, certain stop-loss rules provide a better risk-adjusted return than Buy-and-Hold under most risk-adjusted metrics. On the other hand, in the case of a market with a negative risk premium, all stop-loss rules provide a better expected return than Buy-and-Hold. In the case of the fixed percentage rule, the improvement in expected return in a market with a negative risk premium is outstanding in many cases, and the improvement in risk-adjusted return in a market with a positive risk premium is often remarkable as well.

These results show that by incorporating a relatively simple risk management mechanism such as stop-loss rules it is possible sometimes to significantly enhance risk-adjusted investment performance, introducing very few changes to an existing buy-and-hold passive strategy.

We also showed that by not including overnight gaps, stop-loss performance might in fact be reported to be lower than it truly is, which might be one of the factors that explains the general skepticism towards this risk management mechanism present in the academic literature, despite its undisputed popularity among practitioners.

Finally, the inclusion of flash crashes in our parametric price models reveals that the percentage-based and the ATR-based stop-loss rules are triggered by the the flash crash but in the first case risk-adjusted performance is still improved under most metrics. On the other hand, the moving average crossover and RSI-based stop-loss rules are robust to flash crashes and their performance (a slight improvement in risk-adjusted terms) is roughly the same under the three layers of realism considered: no gap, gap and gap plus flash crash.

References

Emmanuel Acar and Robert Toffel. *Stop-loss and Investment Returns*. Financial Times / Prentice Hall, 2001. ISBN 0273650343, pp. 51-68.

M. J. Cooper, M. T. Cliff, and H. Gulen. Return differences between trading and non-trading hours: Like night and day. Available at SSRN: https://ssrn.com/abstract=1004081, 2008.

A. C. Davison and D. V. Hinkley. *Bootstrap Methods and their Application*. Cambridge University Press, 1997.

Kenneth R French. Stock returns and the weekend effect. *Journal of Financial Economics*, 8(1):55–69, 1980.

H. Geman, D. Madan, and M. Yor. Asset prices are brownian motion: Only in business time. In *Quantitative Analysis In Financial Markets: Collected Papers of the New York University Mathematical Finance Seminar (Volume II)*, pages 103–146. World Scientific, 2001.

M. R. Gibbons and P. Hess. Day of the week effects and asset returns. *Journal of Business*, pages 579–596, 1981.

J. James and L. Yang. Stop-losses, maximum drawdown-at-risk and replicating financial time series with the stationary bootstrap. *Quantitative Finance*, 10(1):1–12, 2010.

Kathryn M. Kaminski and Andrew Lo. When do stop-loss rules stop losses? *Journal of Financial Markets*, 18(C):234–254, 2014.

Donald B Keim and Robert F Stambaugh. A further investigation of the weekend effect in stock returns. *The Journal of Finance*, 39(3):819–835, 1984.

Andrew W Lo and Alexander Remorov. Stop-loss strategies with serial correlation, regime switching, and transaction costs. *Journal of Financial Markets*, 24:1–15, 2017.

S. Nadarajah. A generalized normal distribution. *Journal of Applied Statistics*, 32(7): 685–694, 2005.

D. N. Politis and J. P. Romano. The stationary bootstrap. *Journal of the American Statistical Association*, 89(428):1303–1313, 1994.

D. N. Politis and H. White. Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, 23(1):53–70, 2004.

M. B. Priestley and S. T. Rao. A test for non-stationarity of time-series. *Journal of the Royal Statistical Society. Series B (Methodological)*, 31(1):140–149, 1969.

W. F. Sharpe. The sharpe ratio. *The Journal of Portfolio Management*, 21(1):49–58, 1994.

E. M. Silva and S. Da Silva. Offsetting the disposition effect with a stop-loss rule. *Mathematical Finance Letters*, 2015, 2015.

D. Sornette. *Why Stock Markets Crash: Critical Events in Complex Financial Systems*. Princeton University Press, 2004. ISBN 0691118507.

F.A. Sortino. Performance measurement in a downside risk framework. *Journal of Investing*, 3:50–58, 1994.

Lloyd R. Welch. Hidden Markov Models and the Baum–Welch algorithm. *IEEE Information Theory Society Newsletter*, 53(4), 2003.

Zvi Wiener and Robert Tompkins. Bad days and good nights: A re-examination of non-traded and traded period returns. Available at SSRN: https://ssrn.com/abstract=1102165, 2008.

Diethelm Wuertz, Yohan Chalabi with contribution from Michal Miklovic, Chris Boudt, Pierre Chausse, and others. *fGarch: Rmetrics - Autoregressive Conditional Heteroskedastic Modelling*, 2016. URL https://CRAN.R-project.org/package=fGarch. R package version 3010.82.1.

Appendix

Below are the model-based simulation results for stop-loss rules ATR (Tables 11, 12 and 13) and MA (Tables 14, 15 and 16), followed by the data-based (i.e. bootstrapped) simulation results for stop-loss rules ATR (Table 17) and MA (Table 18).

Table 11: ATR-based SL, without gap: Simulation results

Strategy-Model	$\mathbb{E}[r]$	Sharpe	Sortino	RVaR	RES
SL - GED	6.49%	0.32	0.13	0.15	0.13
B&H - GED	7.23%	0.32	0.14	0.16	0.14
Difference	-0.74% ns	0.00 ns	-0.01**	-0.01*	-0.01**
SL - ARMA	5.59%	0.33	0.13	0.16	0.13
B&H - ARMA	6.84%	0.33	0.14	0.17	0.14
Difference	-1.25% ns	0.00 ns	-0.01**	-0.01**	-0.01***
SL - RS	6.12%	0.29	0.15	0.14	0.11
B&H - RS	6.28%	0.29	0.16	0.14	0.12
Difference	-0.16 ns% ns	0.00 ns	-0.01 ns	-0.01 ns	-0.01 ns

Table 12: ATR-based SL, with gap: Simulation results

Strategy-Model	$\mathbb{E}[r]$	Sharpe	Sortino	RVaR	RES
SL - GEDgap	4.03%	0.33	0.12	0.16	0.13
B&H - GEDgap	4.99%	0.29	0.12	0.15	0.12
Difference	-0.96% ns	+0.04***	0.00 ns	+0.01***	+0.01*
SL - ARMAgap	3.98%	0.35	0.12	0.18	0.14
B&H - ARMAgap	4.96%	0.30	0.12	0.16	0.13
Difference	-0.98% ns	+0.05***	0.00 ns	+0.02***	+0.01***
SL - RSgap	5.53%	0.28	0.10	0.11	0.09
B&H - RSgap	6.32%	0.27	0.11	0.13	0.11
Difference	-0.79% ns	+0.01 ns	-0.01 ns	-0.02***	-0.02***

Table 13: ATR-based SL, with gap and flash crash: Simulation results

Strategy-Model	$\mathbb{E}[r]$	Sharpe	Sortino	RVaR	RES
SL - GEDcrash	-2.18%	-	-	-	-
B&H - GEDcrash	4.95%	0.29	0.12	0.15	0.12
Difference	-7.13% **	-	-	-	-
SL - ARMAcrash	-1.82%	-	-	-	-
B&H - ARMAcrash	4.98%	0.30	0.13	0.16	0.13
Difference	-6.80% *	-	-	-	-
SL - RScrash	-4.27%	-	-	-	-
B&H - RScrash	5.71%	0.25	0.11	0.11	0.09
Difference	-9.98% *	-	-	-	-

Table 14: MA crossover-based SL, without gap: Simulation results

Strategy-Model	$\mathbb{E}[r]$	Sharpe	Sortino	RVaR	RES
SL - GED	5.73%	0.32	0.13	0.16	0.12
B&H - GED	7.12%	0.32	0.15	0.16	0.13
Difference	-1.39% ns	+0.00 ns	-0.02***	+0.00 ns	-0.01**
SL - ARMA	5.49%	0.34	0.12	0.18	0.14
B&H - ARMA	6.79%	0.33	0.15	0.18	0.15
Difference	-1.30% ns	+0.01 ns	-0.03***	+0.00 ns	-0.01***
SL - RS	5.12%	0.30	0.13	0.15	0.11
B&H - RS	6.19%	0.29	0.16	0.14	0.11
Difference	-1.07% ns	+0.01**	-0.03***	+0.01***	+0.00 ns

Table 15: MA crossover-based SL, with gap: Simulation results

Strategy-Model	$\mathbb{E}[r]$	Sharpe	Sortino	RVaR	RES
SL - GEDgap	4.28%	0.31	0.11	0.16	0.13
B&H - GEDgap	4.91%	0.28	0.12	0.15	0.12
Difference	-0.63% ns	+0.03***	-0.01**	+0.01***	+0.01***
SL - ARMAgap	4.30%	0.32	0.11	0.17	0.13
B&H - ARMAgap	4.95%	0.30	0.12	0.16	0.12
Difference	-0.65% ns	+0.02***	-0.01**	+0.01***	+0.01*
SL - RSgap	5.11%	0.28	0.10	0.13	0.10
B&H - RSgap	6.22%	0.27	0.11	0.12	0.10
Difference	-1.11% ns	+0.01**	-0.01**	+0.01***	+0.00 ns

Table 16: MA crossover-based SL, with gap and flash crash: Simulation results

Strategy-Model	$\mathbb{E}[r]$	Sharpe	Sortino	RVaR	RES
SL - GEDcrash	4.27%	0.31	0.11	0.16	0.13
B&H - GEDcrash	4.95%	0.29	0.12	0.15	0.12
Difference	-0.68% ns	+0.02 ns	-0.01 ns	+0.01 ns	+0.01 ns
SL - ARMAcrash	4.24%	0.32	0.10	0.17	0.13
B&H - ARMAcrash	4.96%	0.30	0.12	0.16	0.13
Difference	-0.72% ns	+0.02 ns	-0.02 ns	+0.01 ns	+0.00 ns
SL - RScrash	4.74%	0.26	0.10	0.12	0.09
B&H - RScrash	5.62%	0.25	0.11	0.11	0.09
Difference	-0.88% ns	+0.01 ns	-0.01 ns	+0.01 ns	+0.00 ns

Table 17: Bootstrap simulation results under ATR SL

Stocks	$\mathbb{E}[r]$	Sharpe	Sortino	RVAR	RES
BAC	-6.5103	-	-	-	-
PFE	-4.9619	-	-	-	-
F	-10.6830	-	-	-	-
C	1.4632	0.1553	0.0606	0.0682	0.0554
JPM	-5.7171	-	-	-	-
MRK	-0.3448	-	-	-	-
XOM	0.1091	0.0206	0.0068	-	-
AIG	3.1674	0.4244	0.1767	0.2512	0.2036
HAL	-9.2054	-	-	-	-
KEY	-2.5792	-	-	-	-
BMY	-6.8442	-	-	-	-

Table 18: Bootstrap simulation results under MA SL

Stocks	$\mathbb{E}[r]$	Sharpe	Sortino	RVAR	RES
BAC	-6.8231	-	-	-	-
PFE	-5.4746	-	-	-	-
F	-12.4149	-	-	-	-
C	1.4312	0.1535	0.0530	0.0652	0.0526
JPM	-5.8740	-	-	-	-
MRK	-0.3471	-	-	-	-
XOM	0.1154	0.0220	0.0080	-	-
AIG	3.1591	0.4191	0.1743	0.2377	0.1904
HAL	-9.1302	-	-	-	-
KEY	-2.6658	-	-	-	-
BMY	-6.8560	-	-	-	-