Crash Risk in Individual Stocks

- INTERNET APPENDIX -

A ETF ownership Data Construction

Data collection and database merging.

The list of ETFs is taken from ETF.com, which is a subsidiary of Cboe Global Markets specialized in ETF analysis, databases and news. I consider the ETFs based in US and with an equity focus. The historical quarterly holdings of the ETFs are taken from the S12 fund file of Thompson Reuters, which contains the mutual fund holdings data filed to the SEC (Security Exchange Commission) every quarter. The matching of the ETFs taken from ETF.com with the funds in Thompson Reuters is primarily done via the ticker of the fund. However, because many funds have a missing ticker in Thompson Reuters, I match the funds by names with the fuzzy logic Levenstein measure of string distances. I finally confirm all the matchings by comparing the asset under management (AUM) of the funds. With this procedure, 329 ETFs are matched, which covers 85% of the full ETF scenario as reported by ETF.com.

Some of the ETFs are share classes inside a portfolio (e.g. Vanguard ETFs), and Thompson Reuters reports the holdings at the portfolio level. For these ETFs I adjust the holdings with the weight that the ETF share class has in the portfolio.

The data on the share prices and shares outstanding are taken from Compustat and the matching with the holdings of Thompson Reuters is done via the CUSIP. The data on the shares oustanding of the ETF funds are taken from Bloomberg because of better data coverage.

ETF ownership measure.

For every stock, the daily ETF ownership measure is constructed in two steps, following Ben-David et al. (2018).

First, for each stock i and ETF j, I extract the quarterly weight that the stock has in the ETF fund:

$$w_{i,j,q} = \frac{Sh_{i,j,q} * S_{i,q}}{AUM_{j,q}}, \tag{1}$$

where $Sh_{i,j,q}$ is the number of shares of stock i which are held by the ETF fund j at the end of quarter q, $S_{i,q}$ is the price of stock i at the end of quarter q, and $AUM_{j,q}$ is the amount of asset under management of the fund j.

Then, I compute a daily measure of ETF ownership according to the following formula:

$$\text{ETF Ownership}_{i,t} = \frac{\sum_{j=1}^{J} w_{i,j,t} AUM_{j,t}}{MktCap_{i,t}}, \tag{2}$$

where J is the set of ETFs that hold stock i; $w_{i,j,t}$ is the weight of the stock in the ETF j which is extracted with Equation 1 from the most recent quarterly report; $AUM_{j,t}$ is the amount of asset under management of the fund j at time t calculated as the product of the number of shares outstanding of the ETF fund times the price of the ETF share; and $MktCap_{i,t}$ is the market capitalization of stock i at time t.

B Numerical analysis

This section studies the precision of the skewness swap in measuring the third moment of the asset returns in the Merton jump-diffusion model. The choice of the Merton model is by no means restrictive, because, as shown in Hagan et al. (2002), for time horizons less than one year the implied volatility smile generated by the Merton model can satisfactory reproduce the empirical one. The dynamics of the Merton model is the following:

$$ds_t = \left(r - \lambda\kappa - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t + \log(\psi) dq_t, \tag{3}$$

where s_t is the logarithm of the stock price, r is the risk-free interest rate, and σ is the instantaneous variance of the return conditional on no jump arrivals. The Poisson process, q_t, is independent of W_t, and such that there is a probability λdt that a jump occurs in dt, and $(1-\lambda)dt$ probability that no jump occurs:

$$q_t = \begin{cases} 1, & \text{with probability } \lambda dt \\ 0, & \text{with probability } (1-\lambda)dt \end{cases} \tag{4}$$

The parameter λ represents the mean number of jumps per unit of time. The random variable ψ is such that $\psi - 1$ describes the percentage change in the stock price if the Poisson event occurs, and $\kappa = E[\psi - 1]$ is the mean jump size. I further make the standard assumption (for instance, see Bakshi et al. (1997)) that $\log(\psi) \sim N(\mu, \delta^2)$.

The characteristic function is given by:

$$\phi(u) = E[e^{iu(s_t - s_0)}] = e^{t(i(r-0.5\sigma^2-\lambda\kappa)u-0.5\sigma^2u^2+\lambda(e^{i\mu u-0.5\delta^2u^2}-1))}$$

and the moments can be recovered from the property of the characteristic function:

$$E[(s_t - s_0)^k] = (-i)^k \frac{d^k}{du^k}\phi(u)_{|u=0}.$$

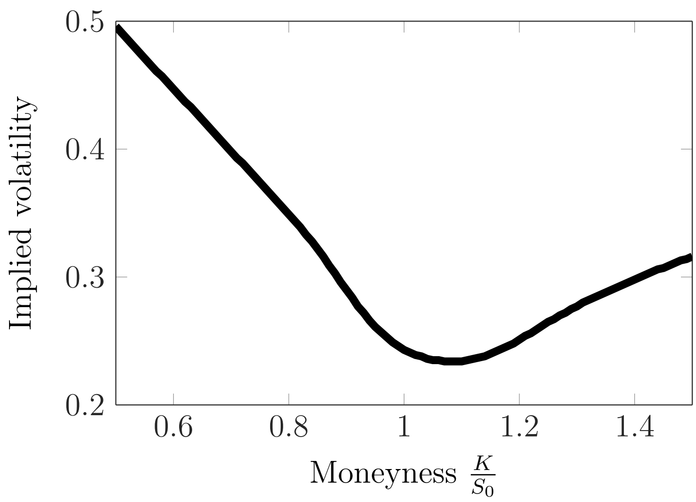

Figure 1. Merton implied volatility smile. The figure displays the one month implied volatility smile generated by the Merton model with the following set of representative parameters: $r = 0, \mu = -0.05, \delta = 0.08, \sigma = 0.2, \lambda = 3$.

As a result, the third moment can be expressed in closed form by the following formula:

$$E[(s_t - s_0)^3] = 3\delta^2\lambda\mu t + \lambda\mu^3 t - 3(-\delta^2\lambda - \lambda\mu^2 - \sigma^2)(r + (-\kappa\lambda + \lambda\mu - 0.5\sigma^2))t^2 + (r - \kappa\lambda + \lambda\mu - 0.5\sigma^2)^3 t^3.$$

I chose the following standard parameter values for the simulation study: $r = 0$, $\mu = -0.05$, $\delta = 0.08$, $\sigma = 0.2$, $\lambda = 3$, $t = 30/365$. With these parameters the Merton implied volatility smile is left skewed as shown in Figure 1. The fixed leg of the swap is defined by Equation 1 in the main text with the discrete approximation given by Equation 4 in the main text; the function Φ is set equal to Φ_S defined in Equation 16 of Appendix A in the main text. I implement the fixed leg of the swap with the option prices implied by the Merton jump-diffusion model.

Table 1 shows the convergence of the fixed leg of the swap to the true model-based third moment when the number of options used increases from 8 to 100. I compare the performance of the swap defined by Φ_S, i.e. the swap used in the empirical application of this paper, with the swap of Schneider and Trojani (2015) defined by Φ_3 (see Equation 15 in Appendix A). The precision of the swap defined by Φ_S is superior because it can reach an error of less than 1%, while the swap defined by Φ_3 has always an error around 20%.

Table 1. Convergence in the number of options.

This table shows the convergence of the fixed leg of the trading strategy to the third moment of the asset returns when the number of options increases. The error is computed as $\frac{|\text{True moment}-\text{Strategy fixed leg}|}{\text{True moment}}$ and it is displayed in percentage. The returns are assumed to follow a Merton jump-diffusion process with standard parameters, i.e. $\mu = -0.05$, $\delta = 0.08$, $\sigma = 0.2$, $\lambda = 3$, $r = 0$, $t = 30/365$. The true moment is computed in closed form and is equal to $-3.12 \cdot 10^{-4}$. In the second column I consider the skewness swap of Schneider and Trojani (2015) with $\Phi = \Phi_3$ while in the third column I consider the skewness swap introduced in this paper with $\Phi = \Phi_S$.

Number of options	Error (%)	
	Φ_3	Φ_S
8	-18.81%	-5.16%
10	-21.10%	-1.70%
20	-22.70%	-0.72%
50	-22.80%	-0.86%
100	-22.80%	-0.87%

This shows that the isolation of the third moment from the fourth is important in order to have a precise measure. It is also interesting to notice that in the swap defined by Φ_S even with only 8 strikes available (which is the average number of options that I have empirically) the error is only of the order of 5%. The results show that the measurement error and the early exercise error of the fixed leg of the swap (see Equation 2 in the main text) are in total less than 1% when the number of options grows to infinity.

Second, I check how the precision of the methodology depends on the range of moneyness available. In this contest the moneyness of an option is defined in standard deviations (SD) as $\log(K/F_{0,T})/(\sigma\sqrt{T})$, where K is the strike price, $F_{0,T}$ is the forward price, σ is the at-the-money implied volatility and T is the time to maturity. I fix a number of options equal to 10 and I calculate the error of the swap if the ten options span the moneyness range $\pm 1SD$, $\pm 2SD$, $\pm 3SD$ and $\pm 4SD$. Figure 2 shows that the precision changes a lot depending on the moneyness range available. At least three standard deviations are needed in order to have an error around 10% and four standard deviations to have an error around 1%.

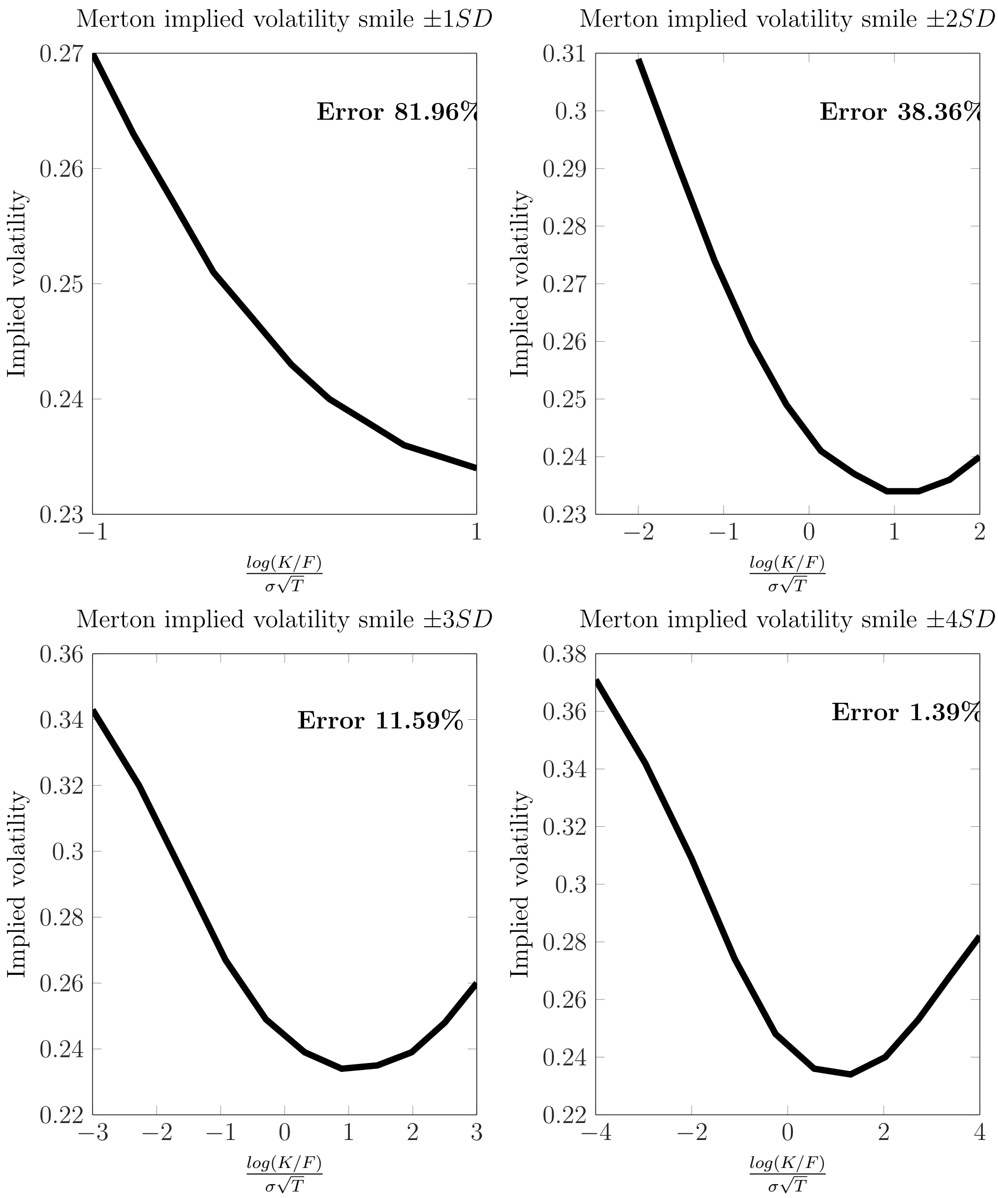

Figure 2. Convergence in the moneyness range. This figure shows the implied volatility smile implied by the Merton jump-diffusion process with parameters $\mu = -0.05$, $\delta = 0.08$, $\sigma = 0.2$, $\lambda = 3$, $r = 0$, $t = 30/365$, for increasing moneyness ranges. In the first plot I consider a moneyness range of $[-1SD, 1SD]$, in the second one $[-2SD, 2SD]$, the third one $[-3SD, 3SD]$ and the fourth one $[-4SD, 4SD]$, where SD is defined as $SD = \log(K/F_{0,T})/(\sigma\sqrt{T})$, in which K is the strike price, $F_{0,T}$ is the forward price, σ is the at-the-money implied volatility and T is the time to maturity. Each plot is a zoom out of the previous one by $1SD$. For each moneyness range the risk-neutral third moment of the returns is computed with the skewness swap defined by Equation 1 in the main text, and the result is compared with the benchmark model-based risk-neutral third moment which is recovered in closed form. At the top right of each graph is displayed the error in percentage.

C Bakshi-Kapadia-Madan (2003) risk-neutral skewness

The risk-neutral skewness proposed by Bakshi et al. (2003) has the following expression:

$$SKEW(t,\tau) = \frac{e^{r\tau}W(t,\tau) - 3\mu(t,\tau)e^{r\tau}V(t,\tau) + 2\mu(t,\tau)^3}{\left[e^{r\tau}V(t,\tau) - \mu(t,\tau)^2\right]^{3/2}}$$

where $\mu(t,\tau) = e^{r\tau} - 1 - \frac{e^{r\tau}}{2}V(t,\tau) - \frac{e^{r\tau}}{6}W(t,\tau) - \frac{e^{r\tau}}{24}X(t,\tau)$, and $V(t,\tau)$, $W(t,\tau)$, $X(t,\tau)$ are respectively the prices of a volatility, cubic and quartic contracts which can be computed with option prices. The price of a volatility contract is given by

$$\begin{aligned} V(t,\tau) = &\int_{S(t)}^{\infty} \frac{2\left(1 - \log\left[\frac{K}{S(t)}\right]\right)}{K^2} C(t,\tau,K)dK \\ &+ \int_0^{S(t)} \frac{2\left(1 + \log\left[\frac{S(t)}{K}\right]\right)}{K^2} P(t,\tau,K)dK, \end{aligned}$$

where $C(t,\tau,K)$ ($P(t,\tau,K)$) is the price of an European call (put) option quoted on day t with time to maturity τ and strike K. $S(t)$ is the stock price on day t. Similarly, the prices of a cubic and quartic contract are given by the following portfolios of options:

$$\begin{aligned} W(t,\tau) = &\int_{S(t)}^{\infty} \frac{6\log\left[\frac{K}{S(t)}\right] - 3\left(\log\left[\frac{K}{S(t)}\right]\right)^2}{K^2} C(t,\tau,K)dK \\ &- \int_0^{S(t)} \frac{6\log\left[\frac{S(t)}{K}\right] + 3\left(\log\left[\frac{S(t)}{K}\right]\right)^2}{K^2} P(t,\tau,K)dK, \end{aligned}$$

$$\begin{aligned} X(t,\tau) = &\int_{S(t)}^{\infty} \frac{12\left(\log\left[\frac{K}{S(t)}\right]\right)^2 - 4\left(\log\left[\frac{K}{S(t)}\right]\right)^3}{K^2} C(t,\tau,K)dK \\ &+ \int_0^{S(t)} \frac{12\left(\log\left[\frac{S(t)}{K}\right]\right)^2 + 4\left(\log\left[\frac{S(t)}{K}\right]\right)^3}{K^2} P(t,\tau,K)dK. \end{aligned}$$

I calculate the risk-neutral skewness of Bakshi et al. (2003) for each day of my sample period (01/01/2003 - 31/12/2017) and for each stock of my sample. I use the daily options quotes given by the optionmetrics interpolated implied volatility surface with 30 days to maturity. I convert the implied volatilities given by optionmetrics in European Black-Scholes prices and I use these prices in the algorithm ($C(t, \tau, K)$ and $P(t, \tau, K)$). The above expressions are written for a continuum of options, but in practice only a finite number of options is available. I therefore apply the following discrete approximation. Suppose that on day t, I have N_c out-of-the-money call options available with strikes $S(t) < K_{c,1} < \cdots < K_{c,N_c}$ and N_p out-of-the-money put options available with strikes $K_{p,1} < \cdots < K_{p,N_p} < S(t)$. The price of a volatility contract is discretized as follows

$$V(t,\tau) = \sum_{i=1}^{N_c} \frac{2\left(1 - \log\left[\frac{K_{c,i}}{S(t)}\right]\right)}{K_{c,i}^2} C(t,\tau,K_{c,i}) \Delta K_i$$
$$\sum_{j=1}^{N_p} \frac{2\left(1 + \log\left[\frac{S(t)}{K_{p,j}}\right]\right)}{K_{p,j}^2} P(t,\tau,K_{p,j}) \Delta K_j,$$

where

$$\Delta K_i = \begin{cases} K_1 - S(t) & \text{if } i = 1, \\ (K_i - K_{i-1}) & \text{if } i > 1. \end{cases}$$

and

$$\Delta K_j = \begin{cases} (K_{j+1} - K_j) & \text{if } j < N_p, \\ S(t) - K_{N_p} & \text{if } j = N_p. \end{cases}$$

The cubic and quartic contracts are discretized in the same way.

D Model-based skewness swap

As a robustness check, I implement a model-based skewness swap, where instead of using the actual option prices, I use the option prices calculated from a fitted model. This skewness swap is not tradable, because the fitted option prices are not real quotations, but it is nevertheless a useful econometric check to compare the model-based skewness swap returns with the real tradable skewness swap returns.

There are two main steps in the implementation of the model-based skewness swap: i) calibration of a model ii) implementation of the swap according to the calibrated model.

Calibration.

The model-based skewness swaps are implemented monthly, as the tradable skewness swaps, and they start and end on the third Friday of each month. I exclude the months in which the stocks pay dividends, in order to simplify the calculation of the model based option prices. I recalibrate the model at each start date of the swap and for each stock separately. In details, at each start date of the swap t and for each stock S_t, I consider all the out-of-the-money options with maturity 30 days provided by the Optionmetrics implied volatility surface file. This sample constitutes my calibration sample. I choose as benchmark model the Merton jump-diffusion model with gaussian jump-size distribution, whose dynamics is given by Equation 3. The variable $\log(\psi)$ (the size of the jump) is normally distributed with parameters $N(\mu, \delta^2)$. This model is simple and tractable, and many empirical studies (see e.g. Hagan et al. (2002)) show that it provides a good fit for short-term options data. I then calibrate the parameters of the Merton jump-diffusion model by minimizing the implied volatility mean squared error (IVMSE) as

$$IVMSE(\chi) = \sum_{i=1}^{n} (\sigma_i - \sigma_i(\chi))^2,$$

where $\chi = \{\lambda, \mu, \delta, \sigma\}$ is the set of parameters to estimate, $\sigma_i = BS^{-1}(O_i, T_i, K_i, S, r)$ is the market implied volatility provided by Optionmetrics and $\sigma_i(\chi) =$

Table 2. Calibrated parameters of the Merton jump-diffusion model.

This table displays the average calibrated parameters of the Merton jump-diffusion model for the S&P500 and for the cross-section of stocks. The model is calibrated separately for each stock and index and it is recalibrated monthly at each start date of the swap. The calibration sample includes all the out-of-the-money options with maturity 30 days quoted by the Optionmetrics interpolated volatility surface file on the calibration day. The numbers displayed are the average calibrated parameters across time and across stocks.

	λ	μ	δ	σ
S&P500	2.59	-0.08	0.05	0.11
All stocks	3.19	-0.07	0.13	0.20

$BS^{-1}(O_i(\chi), T_i, K_i, S, r)$ is the model implied volatility, where $O_i(\chi)$ is the Merton model price of the option i. The model implied volatility is obtained by inverting the Black-Scholes formula where the option price is given by the Merton model price. In the Merton jump-diffusion model the option prices are available in closed form (see Merton (1976)). The price of a call option is given by:

$$C_{MRT}(t, \tau, K) = \sum_{n=0}^{\infty} e^{-\lambda'\tau + n\log(\lambda'\tau) - \sum_{i=1}^{n}\log n} C(S_t, K, r_n, \sigma_n),$$

where $\lambda' = \lambda(1+k)$, $k = e^{\mu+\frac{1}{2}\delta^2} - 1$, $C(S_t, K, r_n, \sigma_n)$ is the Black-Scholes price of an European call with volatility $\sigma_n = \sqrt{\sigma^2 + \frac{n\delta^2}{\tau}}$ and risk-free rate $r_n = r - \lambda k + \frac{n\log(1+k)}{\tau}$. The price of a put option is defined analogously. The choice of the implied volatility mean squared error (IVMSE) loss function follows the argumentation of Christoffersen and Jacobs (2004), where they show that the calibration made on implied volatilities is more stable out of sample. Table 2 displays the average calibrated parameters for the S&P500 index and for the cross-section of stocks.

Implementation of the model-based swap.

With the calibration of the Merton jump-diffusion model in the previous step, I estimate the parameters $\{\widehat{\lambda}, \widehat{\mu}, \widehat{\delta}, \widehat{\sigma}\}$ on day t for stock S_t. I then compute the Merton option prices of an equispaced grid of strikes covering the moneyness range $[-4SD, +4SD]$, where $SD = \frac{\log(K/S_t)}{\sigma\sqrt{T}}$ is the moneyness of the options measured in standard deviations. In this definition σ is calculated as the implied volatility of an at-the-money option, i.e. $\sigma = BS^{-1}(C_{MRT}, T, S_t, S_t, r)$, where C_{MRT} is the Merton price of a call option with strike equal to S_t. The equispaced grid is constructed as follows. First, I recover $K_{min} = S_t e^{-4\sigma\sqrt{T}}$ and $K_{max} = S_t e^{4\sigma\sqrt{T}}$, then I divide the range $[K_{min}, S_t]$ and $[S_t, K_{max}]$ in 20 intervals each, and finally I compute the Merton option prices of these out-of-the-money puts and calls. In this way I keep constant the number of options and the moneyness range, because I always have 40 option prices, 20 calls and 20 puts, which span the range $[-4SD, +4SD]$.

The fixed leg of the swap (Equation 1 in the main text) is computed using these model-based option prices, where the discreteness of the options is addressed with the same quadrature-based approximation of the integral explained in Equation 7 in the main text. The floating leg of the swap (Equation 2 in the main text) is given by the sum of the payoff of the same option portfolio and a continuous delta-hedge in the forward market. The return of the swap is computed with Equation 11 in the main text.

Table 4 in the main text reports the average return of the model-based skewness swap for the S&P500 index and for the cross-section of stocks in the pre-crisis and post-crisis subsample. While the risk premium for the S&P500 increases by 20%, the model-based skewness risk premium for single stocks goes from 32.45% to 43.56%, which corresponds to a 35% increase. Moreover, while before the crisis the model-based skewness risk premium is statistically significant for 298 stocks, after the financial crisis around 600 stocks have a statistically significant model-based skewness risk premium.

References

G. Bakshi, C. Cao, and Z. Chen. Empirical performance of alternative option pricing models. *The Journal of Finance*, 52(5):2003–2049, 1997.

G. Bakshi, N. Kapadia, and D. Madan. Stock return charachteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies*, 16 (1):101–143, 2003.

I. Ben-David, F. Franzoni, and R. Moussawi. Do etfs increase volatility? *The Journal of Finance*, 73(6):2471–2535, 2018.

P. Christoffersen and K. Jacobs. The importance of the loss function in option valuation. *Journal of Financial Economics*, 72(2):291–318, 2004.

P. S. Hagan, D. Kumar, A. S. Lesniewski, and Woodward D. E. Managing smile risk. *The Best of Wilmott*, page 249, 2002.

R. C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2):125–144, 1976.

P. Schneider and F. Trojani. Fear trading. *Swiss Finance Institute Research Paper*, 15 (3), 2015.