DERIVATIVES & STRUCTURED PRODUCTS: PROJECT

THE "POLLUPTION": A WAY TO INSURE THE INSURER

Yoan Rebetez

Professor Tony Berrada
Assistant Maxime Auberson

University of Geneva
Spring 2020

TABLE OF CONTENTS

Introduction

Climate-related risks have always been an important factor to consider for firms. They can lead to substantial losses and it turns out that these kinds of risks can be very hard to hedge. To overcome this issue weather derivatives have been created and the first product of this kind was issued in 1996 (Aquilla Energy). Since then, they have been widely developed on OTC markets and the CME (Chicago Mercantile Exchange) became the unique organised market to propose these products. It remains today the leader in terms of transactions in weather derivative contracts.

As the market of climate-related derivatives is well expanded in the USA but not very much in Switzerland, I wanted to create such a product for the swiss market. Moreover, if one looks for weather derivatives contracts it is very likely to find products which rely on temperature, tornados, rainfalls, or earthquakes but not on pollution. This factor now represents a big threat for the public health and therefore, for the insurance sector.
Thus, the purpose of the product is to generate a certain payoff when the pollution increases too much (see section payoff structure), which could lead to financial losses for health insurances (I reasonably assume that as the air pollution becomes more present in the air, people will go more often to the doctor or to the hospital due to increases of diseases). To have an order of magnitude, the total costs of air pollution in the EU are in the range of 330-940 billion Euros per year[1].

[1] According to the European Parliamentary Research Service

Part 1: Marketing

The underlying

The idea is to create an index that can be a good proxy for the level of pollution in the air in Switzerland. To face this problem, I use data from January 1, 2010 to May 12, 2020 indicating the level of Ozone (O_3) measured in 6 main swiss cities: Basel, Bern, Lausanne, Lugano, Sion, and Zurich. The choice of this pollutant was driven by the availability of data[2] and because it reflects well the quality of air and is directly and highly related to emission degrees of firms, that are called "precursor of Ozone".

Table I : Pollution Index standards

O_3	Pollution
0-90	Low
91-120	Acceptable
121-150	Significant
151-180	Impactful
181-240	High
>240	Very high

The index is composed as follows for the pollutant Ozone. If it lies from 0 to 90, the little pollution has a low impact on health ; it goes crescendo to a level of 240 where it is very likely to suffer from diseases caused by air pollution (trouble breathing, cancer for smokers, pneumonia etc). In fact, the European health commission has as an objective to lower the concentration of polluted air to reach the threshold of 80 (instead of 120 before 2012) by 2030.

In terms of relevance of my product, it would be interesting for health insurance companies to get payoffs when the index lies from 91 to 240 over a certain period, with a certain coefficient factor for each tranche (that is not proportionally increasing).
You can find below a "summary statistics" of my raw data over the period divided into seasons.

Table III: Descriptive statistics of the raw data, divided into seasons

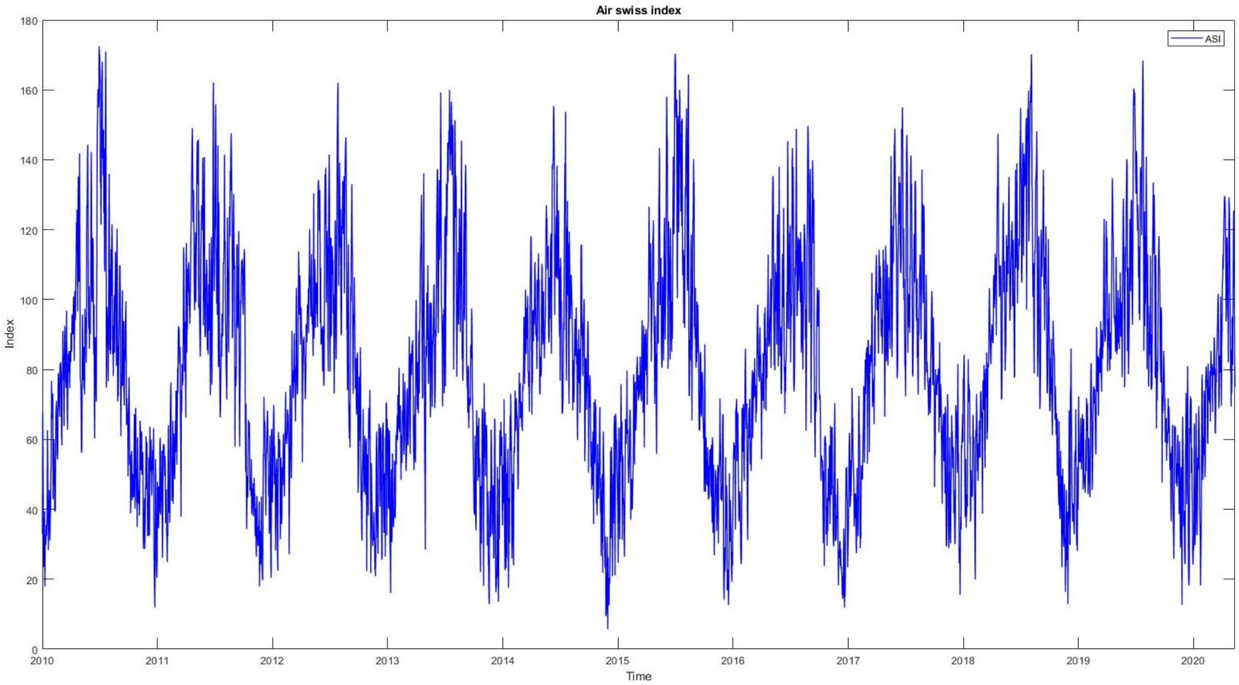
	Winter						Spring						Summer						Fall					
	Bern	Lausanne	Lugano	Zurich	Basel	Sion	Bern	Lausanne	Lugano	Zurich	Basel	Sion	Bern	Lausanne	Lugano	Zurich	Basel	Sion	Bern	Lausanne	Lugano	Zurich	Basel	Sion
Min	6.1	8.2	1.8	5.2	4.0	1.9	14.9	20.7	18.5	19.7	24.3	19.5	18.2	35.0	34.5	20.0	32.1	32.6	1.2	2.1	2.9	1.7	1.9	2.0
Max	121.9	114.1	137.0	130.3	138.3	126.8	168.3	165.1	199.1	178.4	196.6	151.0	171.8	161.7	199.4	198.7	195.2	183.0	92.0	87.6	163.1	110.6	126.3	101.3
Mean	51.9	55.4	62.0	60.4	64.5	65.0	89.2	89.8	118.1	103.2	103.7	102.1	86.9	89.4	128.8	102.7	103.5	97.3	35.7	41.4	48.6	43.0	47.2	44.1
SD	20.8	18.5	27.2	23.0	24.1	24.5	19.4	18.0	31.0	23.4	24.1	17.2	23.3	21.7	35.0	29.0	29.3	22.5	19.8	18.2	23.7	22.1	24.1	18.5

[2] Taken from « Office fédérale de la statistique »

The first thing to notice is that the air in Switzerland is quite healthy. Except for some rare cases, the index does not find itself in the “high” or “very high” zone. On average, only Lugano in Summer suffers from a bad air quality but apart from that, the country breathes quite well. However, the table aggregates the values into seasons, which can omit some intra-season trends. Even though scientists sound the alarm when only daily values are above a certain threshold, they also stipulate that a month of exposition at an acceptable level can cause some real troubles and lead to cardiovascular diseases[3]. Smokers are an easy prey and they are likely to develop cancers if they are too much exposed to high levels of pollution.

The air swiss index is then composed of the 6 cities, weighted by their level of pollution. You can find below the behaviour of the index that I named “ASI”.

Figure 1 : Air swiss Index



At first sight, we can see the cyclicity (or seasonality) of the series as well as the mean-reversion. The upward trend is hardly distinguishable, but we will see later that it exists. The dispersion is also bigger in summer, as we could think when we looked at the descriptive statistics (Table II). You can find its analogue in table III.

Table III : Summary statistics of the ASI

	Winter	Spring	Summer	Fall
Min	16.1	28.4	43.4	5.7
Max	123.1	172.5	170.9	114.5
Mean	63.2	103.3	104.7	47.9
SD	18.6	19.8	24.8	16.3

[3] https://www.ersnet.org/pdf/publications/air-quality-FRE.pdf

The Payoff Structure

You can find hereafter the design of my payoff structure

$$CF_{T\_month} = a \sum_{t}^{T\_month} Max\,(ASI_t - (\overline{ASI_\tau} + B * t), 0) \qquad a = \begin{cases} 0 & 0 < ASIt < 90 \\ 1 & 90 \leq ASIt < 120 \\ 3 & 120 \leq ASIt < 150 \\ 9 & 150 \leq ASIt < 180 \\ 18 & ASIt \geq 180 \end{cases}$$

Where

$\overline{ASI_\tau}$ *is the sample average for each month for* $\tau = 1,2...,12$
B represents a coefficient of trend that will be estimated later on
ASI_t *is the average value of the index at time t for the month* $t = 1,2...,T\_month$

Variable II: $\overline{ASI_\tau}$

Jan	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
49.1	59.7	80.4	97.8	101.5	110.5	118.8	107.5	88.5	57.6	43.6	43.1

The idea of my payoff structure is to compensate the potential losses in case of an increase of the Index compared to a previous benchmark (I consider a sample average over the observation period for each month) plus a small trend as the ASI has a slightly upward trend. As the underlying shows a high cyclicity and is mean-reverting, it makes sense to design a payoff structure related to a deviation from a past benchmark. For each month, I define ASI_t as the average over that month until the maturity month. "T_month" is therefore equal to the number of months until the maturity of the product. $\overline{ASI_\tau}$ as shown in "variable I" is the empirical average for each month of the index level (the benchmark). The coefficient a is a multiplicator that depends on where ASI_t lies on average each month. Regarding table I, the more you scale down in the chart, the better the payoff should get. According to the research of the European respiratory society "Air quality and Health", it would be 3 times more harmful to be exposed to bad air than to acceptable quality of air. That is why I choose to multiply the payoff by 3 if the index lies between 120 and 150. Even though the financial losses of air pollution for an insurance firm is nearly impossible to quantify this is the best approximation I could make. In the same research, scientists also say that breathing polluted air whose level falls down the scale gets worse and worse and that is the reason why I do not choose a linear increase in the coefficient if the Index falls into bad level of air quality. I have good reasons to believe that the more we get down the scale, the bigger the financial loss gets (more than proportionally!).

Maturity

I choose a maturity of 2 years for one main reason. As the costs (or consequences) do not arise directly after the observation of the increase in air pollution, it seems logical to choose such a period. Patients take some time before seeing a doctor and the disease is potentially contracted after a certain time that I estimate as being 2 years (the authors talk about medium term in the research paper).

Target clients

As parsimoniously mentioned, the payoff would be of high interest for a firm whose costs raise when the level of pollution ekes out. A typical example is the health insurances or even the reinsurances. As we know from the cited research that air pollution is the 3rd factor to cause deaths (after alcohol and car accidents), health insurance cannot be "hedged" financially against this bad feature. Indeed, everyone undergoes the same pollution and a premium of a health insurance contract is not driven by the area which the patient breathes in. Furthermore, the air pollution is not ready to be cut down and seems to slightly increase over time. It would give an opportunity not to pass the costs on the taxpayer. We can imagine that in the future, air pollution could be a factor to increase even more the health insurance premium so it could give a competitive advantage for a health insurance company to deal with the product which prevents it from increasing the premium too much. Knowing how much swiss people suffer from high premia, it would be clever to avoid increasing them too much…

Product advantages and defaults

The product has the feature that it is a unique one in Switzerland. There exists no other way to hedge pollution so it is novel and in case of sudden need of such a product, it would have the monopoly on the market. From the derivative point of view, having the monopoly also means to have the possibility to charge juicy fees. Also, the product can be easily understood by any investor and through its simplicity, the pricing is quite easy to make. To conclude, we will see in the pricing section that it is very likely that the client gets a payoff according to the model. It is always good news before investing to know that it is very probable to get something from an investment.

However, it turns out to have some defaults and the biggest one is that it is nearly impossible to hedge it perfectly except if we find an investor ready to take the opposite bet (I'll go further in details in the hedging section). Another relevant drawback is the fact that we do not know exactly the financial losses compared to air pollution so the coefficient "a" is very problematic to guess. In the same way, as the pollution is non traded, I need to make an assumption about the market price of risk that can lead to substantial different in price if I choose one that is far from the true one.

Part 2: Pricing and hedging

Let us switch to the technical part of the project. After defining the design of the product, it is essential to formulate a coherent model to obtain a fair price.

The model

As we can see from Figure 1, the underlying shows some cyclicity and some mean-reversion. It is the reason why I choose to implement an Ornstein-Uhlenbeck process with time-varying variables. The model is also very popular to model temperature as it cannot rise day after day for a long time. The same reasoning is applied to air pollution. The model under P is as follows:

$$dASI_t = \Theta(ASI_t^P - ASI_t)dt + \sigma_t dW_t$$

Where Θ represents the speed of mean-reversion, ASI_t^P the mean of air pollution at time t that will also be modelled and σ_t being the quadratic variation of pollution.

Under Q, the dynamic changes and after application of Girsanov theorem I obtain:

$$dASI_t = \Theta\left(ASI_t^Q - ASI_t\right)dt + \sigma_t dW_t^Q$$

Where ASI_t^Q = ASI_t^P - $\frac{\sigma_t \gamma}{\Theta}$ and γ = market price of risk of air pollution. The latter will need to be assumed later.

However, the model might not be optimal since it can lead to negative values whilst air pollution cannot be. We will see in the paths simulation section that it might happen but not significantly. If I would really like to be accurate, I would have chosen a Cox-Ingersoll-Ross model, which exhibits a square root and therefore avoid negative values. The choice of the Ornstein-Uhlenbeck was driven by its simplicity since an exact discretization of the process exists.

As the mean of air pollution is time-varying, I will need to model it. Looking at figure 1, we can see that the curve behaves like a sine function. As it is also widely used in the literature (Alaton, Djehiche & Stillberger 2002), I will use such a model:

$$ASI_t^P = A + Bt + C * \sin(\omega t + \varphi)$$

Where A is a constant, B represents the trend, C is the amplitude of the function and φ is the width between each cycle. $\omega = \frac{2\pi}{365}$ as the period of oscillation is one year. The estimation of the parameters will be the topic of the next section.

Parameter estimation

I first start with the presentation parameters estimation of the average temperature over time. I choose the method of minimization of least squares as it is simple and efficient. After estimation, I obtain the following fitted model:

$$ASI_t^P = 77.16 + 0.0016t - 37.64 * \sin\left(\frac{2\pi}{365}t + 114.85\right)$$

We can see that there exists a small trend that over 10 years, leads to an increase of 5.84 units on the index level. The amplitude is 37.64 in absolute value and shows that between a summer day (where the level is at its peak) and a normal winter day, the difference is about 75 units.

Once I plot the model to the data, I obtain this figure:

Figure 2 : Model fit to the data

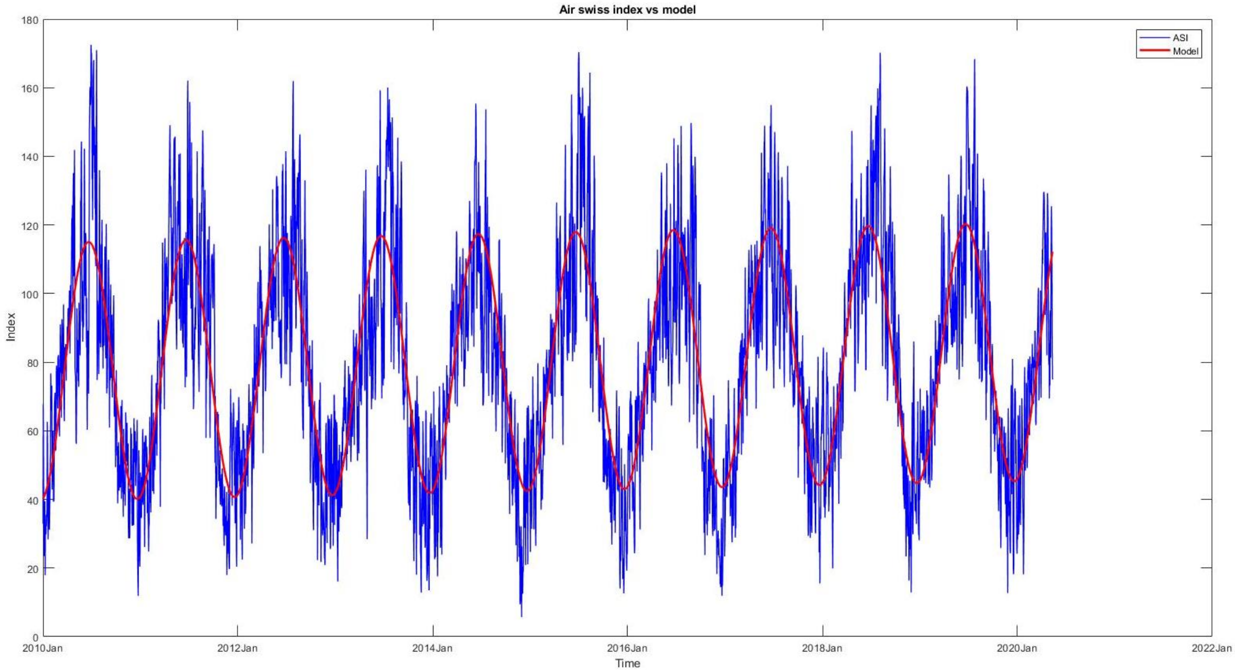
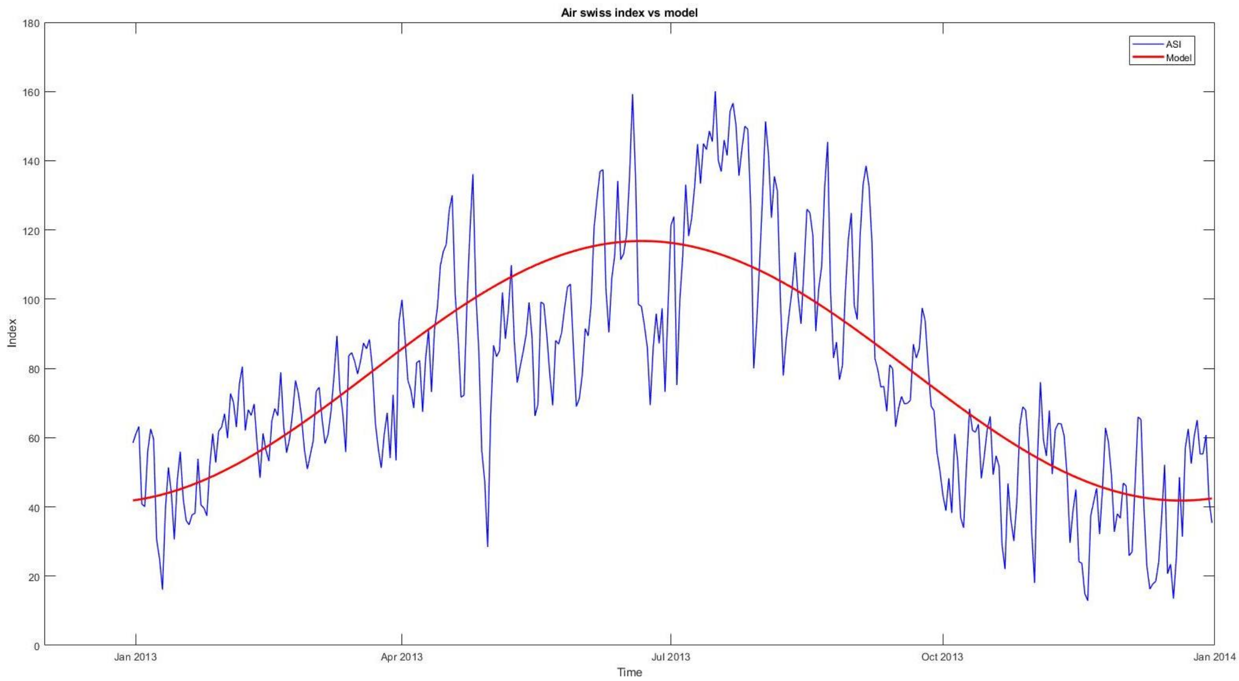


Figure 3 : Model fit to the data over 1 year (2013-2014)



As we can see, the model seems to be a good fit.

Estimation of σ_t and Θ

After trying to estimate constant paramteres by maximum likelihood that lead to bad results in terms of paths simulation, I decided to stick to a time-varying σ_t. Accuracy is a clear argument towards time-varying coefficients. To estimate the coefficients, I refer to the work of Alaton, Djehiche & Stillberger (2002) to find an estimate based on the quadratic variation of $\mathrm{ASI_t}$ (Basawa and Prasaka Rao, 1980). The estimator is constructed as follows:

$$\hat{\sigma}_\tau^2 = \frac{1}{N_\tau}\sum_{t=1}^{N_\tau}(ASI_{t+1} - ASI_t)^2$$

Where N_τ is the number of days occurring during a month.

I repeat the procedure every year, which will give 10 estimates of $\hat{\sigma}_\tau^2$ for each month. I then take the average of those to get a single estimate for each month. The results show up in the next table:

Variable II: $\hat{\sigma}_t$

Jan	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
12.00	10.30	11.17	12.26	13.91	16.03	18.66	16.33	13.25	11.75	12.02	13.63

Also derived from the paper of Alaton & al. the estimation of Θ relies on a martingale approach first suggested by Bibby & Sorensen (1995). The paper derives plenty of estimates for different processes and for the Ornstein-Uhlenbeck, an efficient estimator $\hat{\Theta}$ is obtained by setting $G_T(\hat{\Theta}_n)$ =0

$$\text{Where } G_T(\Theta) = \sum_{t=1}^{T}\frac{(ASI_{t-1}^P - ASI_{t-1})}{\hat{\sigma}_{t-1}^2}(ASI_t - E[ASI_t|ASI_{t-1}])$$

To determine $E[ASI_t|ASI_{t-1}]$ I start from the solution of the differential equation which is (for $t \geq s$):

$$ASI_t = (ASI_s - ASI_s^P)e^{-\Theta(t-s)} + ASI_t^P + \int_s^t e^{-\Theta(t-\tau)}\hat{\sigma}_\tau dW_\tau$$

If we take the conditional expectation of the process, we get:

$$E[ASI_t|ASI_{t-1}] = (ASI_{t-1} - ASI_{t-1}^P)e^{-\Theta} + ASI_t^P$$

Since the conditional expectation of the stochastic integral = 0. It gives us:

$$G_n(\Theta) = \sum_{t=1}^{T}\frac{(ASI_t^P - ASI_t)}{\hat{\sigma}_{t-1}^2}(ASI_t - (ASI_{t-1} - ASI_{t-1}^P)e^{-\Theta} + ASI_t^P)$$

After some steps, we can check that $G_T(\hat{\Theta}_n)$ = 0 for:

$$\hat{\Theta}_T = -\log\left(\frac{\sum_{t=1}^{T}\frac{(ASI_{t-1}^P - ASI_{t-1})}{\hat{\sigma}_{t-1}^2}(ASI_t - ASI_t^P)}{\sum_{t=1}^{T}\frac{(ASI_{t-1}^P - ASI_{t-1})}{\hat{\sigma}_{t-1}^2}(ASI_{t-1} - ASI_{t-1}^P)}\right)$$

In my case, $\hat{\Theta}_T$ = 0.318

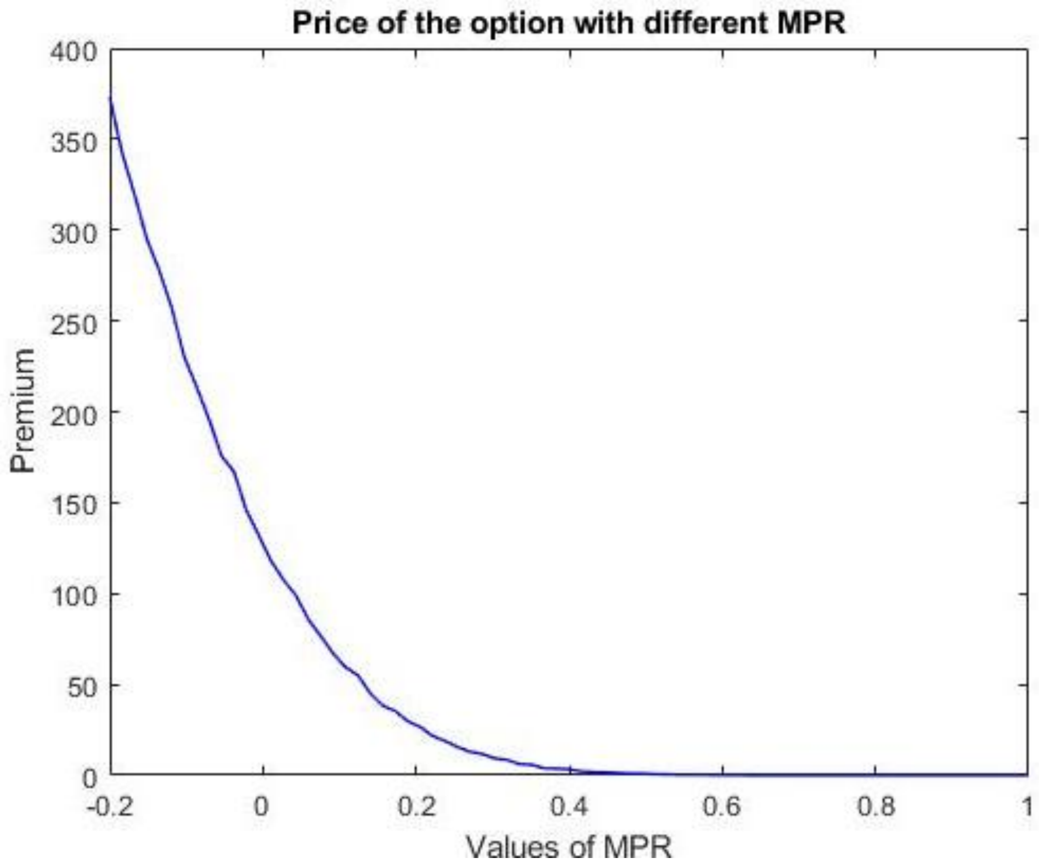
Monte Carlo method for paths simulation

After estimating all the parameters, I need to simulate my process through a Monte Carlo simulation. As no closed form exists, this method is the easiest and most practical one. In order to simulate my paths, I need to define my market price of risk for my non-traded asset.

The assumption on γ

To simulate my paths under the probability measure Q, I need to assume the coefficient γ. As no derivative market on air pollution exists, I cannot find out an implied market price of risk as it could be the case for a product on temperature. To come up with a consistent but arbitrary value, I focus on three factors. The first one is the stability of the premium after simulation for different values of γ.

Figure 4: Value of the option for different market prices of risk



The decreasing pattern shows a sort of stability zone between 0.2 and 0.4. Before 0.2, the curve grows exponentially, and it might not be a good choice to take a value below 0.2 as it might not reflect a fair price. Above 0.4, the premium is too flat and too low in my opinion.

The second factor is the slight comparison of such an MPR with a product that might be close in terms of features. I decided to compare my product with an HDD (call option on temperature) in Sweden because this is the closest to mine in terms of payoff and the Nordic country has similarities with Switzerland in terms of climate. From Alaton et al., the market price of risk for such a contract amounts to 0.08.

The third factor is simply the fact that the MPR should be the same order of magnitude as the Sharpe Ratio of the market on which the asset could be traded. It would make no sense to choose an MPR = 6 if the Sharpe ratio of the market is 0.3. As a proxy, the Sharpe ratio of the SMI was 0.3 over the two last years of data collection (05/12/2018-05/12/2020).

Of course, I give more credit to the first factor and I decide to choose arbitrarily an MPR of 0.25.

Paths simulation
To simulate the paths and approximate the continuous variation of my index, I use the exact discretization of the Ornstein-Uhlenbeck process:

$$ASI_{t+\Delta t} = ASI_t e^{-\hat{\Theta}\Delta t} + ASI_t^Q\left(1 - e^{-\hat{\Theta}\Delta t}\right) + \hat{\sigma}_t\sqrt{\frac{1 - e^{-2\hat{\Theta}\Delta t}}{2\hat{\Theta}}}Z$$

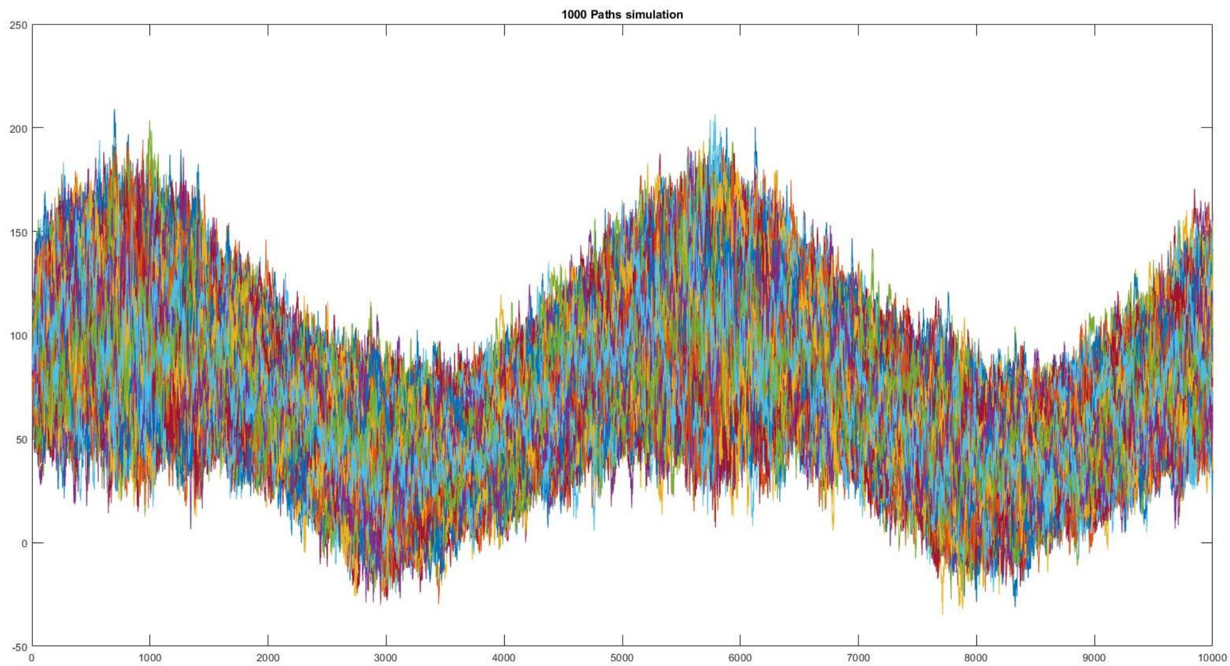
Where Z ~N(0,1) and $\Delta t = \frac{maturity}{T\_steps}$

In my case, the following setup needs to be taken into consideration:

MPR=0.25
Maturity=730 (2 years)
Number of simulations: 1000
Number of time steps: 10000
Δt = 0.073 which means that each step represents 0.073 day which roughly corresponds to 1h45.

As matlab is a slow language in terms of loop computations, I could not really choose a greater number of time steps. Find below the graphical result of the simulation:

Figure 5: Paths simulation of the O-U process



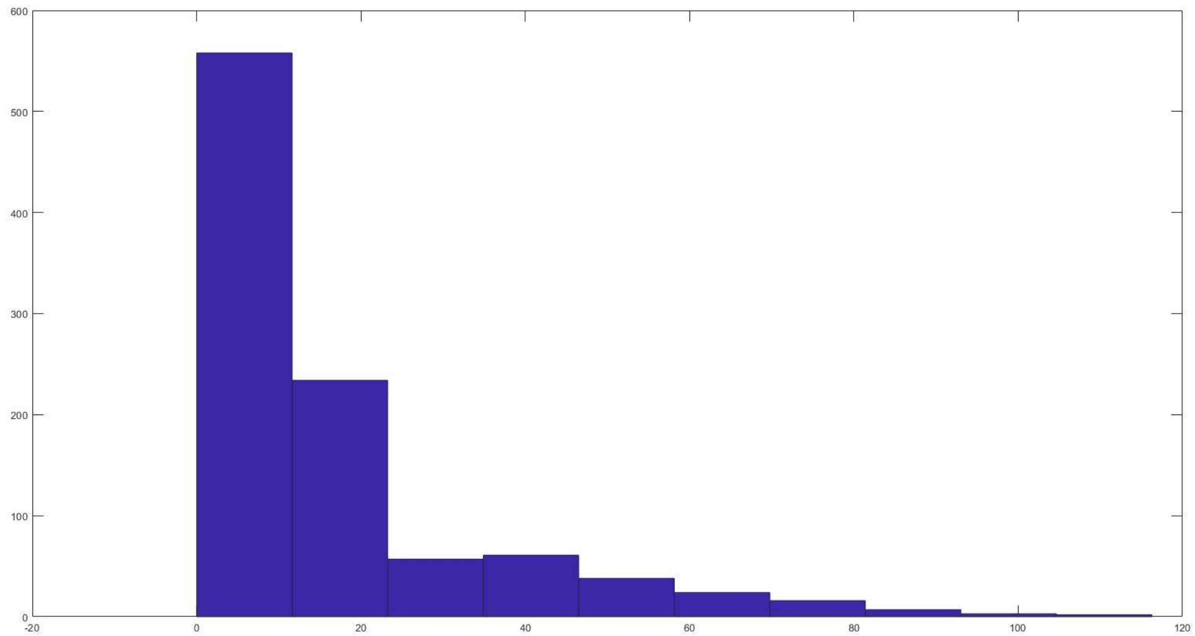
We see that the simulations are aligned with what one could expect. The cyclicity remains and the higher volatility appears in Summer. Although the model is easy to use thanks to its discretization, we see that it allows some negative values which can be problematic if they appear too often. Only 0,1% of the values simulated are below 0 which does not seem to be much of a problem. As previously mentioned, even though this point can be avoided using a Cox-Ingersoll-Ross process, my choice was driven by its easiness of use.

As the paths are random, there exists some techniques to lower the variance of the Monte Carlo method and one of them is called antithetic variates. Its use will be displayed graphically in the next section.

The pricing

Once we simulated the paths under Q, the price of the option will only be the discounted expectation of those. By construction, I obtain a matrix "CF" (Number of simulations x 24) in which I compare the average of the simulated values and the empirical average over a month. I then sum all those cash flows to get a vector "CF_T" (Number of simulations x 1). The important assumption of my structure is that the risk-free rate is 0% due to the swiss economic conditions, and it is supposed to remain the same during the next 2 years. By taking the mean of the vector "CF_T", I get a price of 16.35 CHF. The maximum payoff that can be obtained is 116.25CHF. Here displayed the distribution of the payoff through a histogram:

Figure 6: Histogram of the payoff at maturity



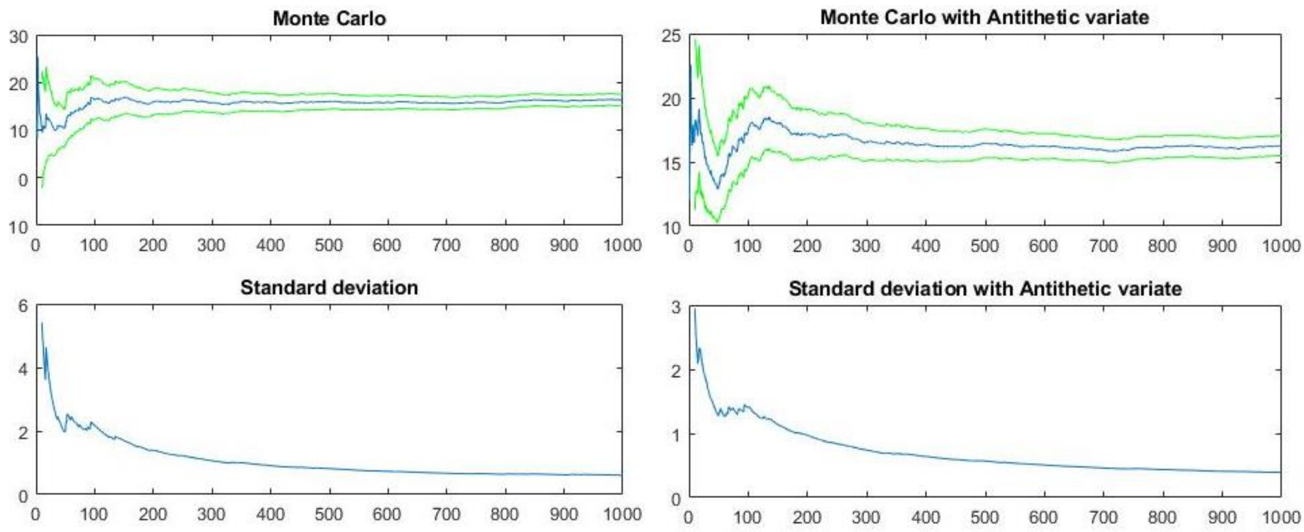
In 10% of cases, the payoff will be null and in 69% of cases it will be below the premium. The price of the option is explained by the large payoffs one could get which is translated into an average of the index above 120 or even 150. An interesting thing as an advantage of the product, is that the client will have 80% chance to get a payoff between 10 CHF and 47 CHF so it is always good news to know that it is likely to get something.

The price with antithetic variates

A nice technique to reduce the error in the Monte Carlo simulation is called antithetic variates. Basically, the idea is to create a copy of the variable of interest with the same mean and variance but negatively correlated. If we add the two variables and divide by 2 (taking the average), the mean of the variable of interest will remain the same but the variance will decrease. In the path simulation, it is translated into adding a minus sign before the volatility of the process.

Here I show the difference of precision:

Figure 7: Comparison of the price with Antithetic variates method



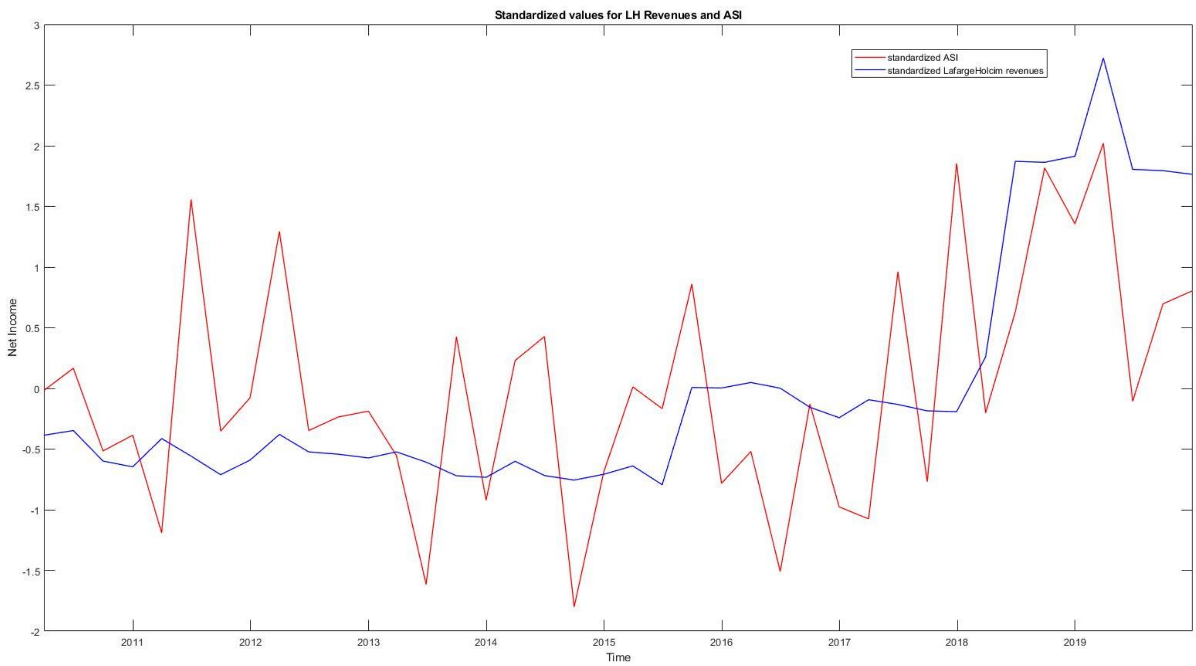
We can witness the reduction of variance with the antithetic variates method to obtain a more accurate price. The price found with this method is 16.25 CHF.

The hedging

The hedging part is probably the weakness of the product. I cannot guarantee a perfect hedge as it could be the case if the underlying were traded. Instead, I assume that some firms might be interested in the opposite payoff.

After selling the product to the health insurance firm what will give rise to a premium, I will use this premium to buy the same option to a firm that needs the opposite payoff (i.e. that will pay me every time the index is above than its historical average against a premium). This might seem hard to find but I thought of selling such a product (buying them the call) to firms that are prone to pollute. To back up my idea, I analyse the link between the revenues of Lafarge Holcim (quarterly), which is the most polluting firm in Switzerland, and the air pollution index. As Ozone is mainly generated by firms (they are called precursor of Ozone), there should be a positive link between how much firms can pollute and their revenues.

Figure 8: Air pollution and LH revenues



After standardizing the two variables for purpose of comparison, we can notice a certain trend. Indeed, the values for the Pearson's and Spearman's correlation are respectively 0.52 and 0.36. Even though there might be an endogenous effect like there is less pollution in Winter and less sales as well, I still believe that a link between the two exists.

As regulations in terms of pollution are becoming stricter and stricter (the objective of Switzerland is to have an upper threshold of 90 in terms of index level), polluting enterprises could suffer from a reduction of pollution. They could be interested to compensate their loss in selling such a product. If the pollution is great (the business of such firms goes well) I would get a payoff that I will transfer to the health insurance company. This point is assuming that a firm like Lafarge Holcim would be willing to pay such a payoff as they would not mind since the revenues are juicy. On the opposite side, if the pollution is low and their revenues too, they would be compensated through the premium.

Conclusion

The price of the option is somehow aligned with what I expected: A lot of low payoffs and some big payoffs but not frequent. The modelling part displays coherent results even though it could have been improved using a process that guarantees positive values (like a CIR process) but the easiness of the implementation got me going this way. From the designing part of the payoff, I am aware that the main issue is to allocate coefficients to how much a health insurance company would lose if air pollution raises dramatically but I believe that such a product is feasible.

Nevertheless, I think that this novel product has interesting features that might be exploited in the near feature. As pollution level is becoming a hot topic aligned with global warming, some firms could be interested in subscribing such a product. Despite that, it would be hard to implement it here in Switzerland as the level of pollution is not extreme and is controllable even though the government has a strict program to lower emissions. I chose this country as data was available, but it would be interesting to see if a demand could be tangible in a country like China where air pollution is a burden.