Politecnico di Milano

Financial Engineering

Group 9
Assignment 5

Martino Bernasconi
Eleonora Capelli
Alessandro Lazzaretti
Heykel Rajhi

1 Historical (HS & WHS) Simulation, Bootstrap and PCA for VaR ES

1.1 Portfolio A. Historical Simulation & Bootstrap

The Historical simulation method is done by ordering the losses obtained with today weights and past observed results:

$$L_s = \{X_s\}_{s=t-n+1,...,t}$$

So that we have

$$L^{i,n} : L^{n,n} \leq ... \leq L^{1,n}$$

Then we measure the Value at Risk and our Expected Shortfall as the $1 - alpha$ percentile, hence:

$$Var_\alpha = L^{([n(1-\alpha)],n)}$$

$$ES_\alpha = mean\{L^{(i,n)}, i = [n(1-\alpha)], ..., 1\}$$

where $[n(1-\alpha)]$ is the largest integer not exceeding $n(1-\alpha)$.
The Bootstrap is the same technique with a random sampling with replacement from the original dataset, this might help if the original sample size is too big. Clearly we have a random component in this results. Since we are looking at tail events, the variance can be high since even few data points, missed or included in the new sample, can shift the result by a great deal. For the first portfolio we obtain the following results.

	VaR_α	ES_α
HS	315K	234K
Bootstrap	297K	198K

Table 1: Historical simulation Vs Bootstrap. Notional: 3.2Mio.

1.2 Weighted Historical Simulation approach

The Historical simulation technique can be modified to give more importance to most recent losses by weighting each loss in a exponentially decreasing way. Then we order again the losses by magnitude and we consider the i^*-th one as the VaR measure, were the index i^* is such that $\sum_{i=1}^{i^*} w_i \leq 1 - \alpha$ where the sum is done by considering the ordering imposed by the losses.
More precisely the weights are defined as:

$$w_s = C\lambda^{t-s}$$

with $\lambda = 0.9$ with $C = \dfrac{1-\lambda}{1-\lambda^n}$ as normalization constant.
For the second portfolio the results obtained are the following:

	VaR_α	ES_α
WHS	37K	34K

Table 2: Weighted Historical simulation. Notional: 1Mio.

1.3 Standard Gaussian parametric approach and Gaussian parametric PCA approach

In short the PCA methodology is a way to describe a dataset in $\mathbb{R}^n$ as a dataset of a smaller space $\mathbb{R}^m$, with $m < n$. This is done by considering a base for $\mathbb{R}^m$ in $\mathbb{R}^n$ such that we are minimizing the variance loss in the projection, hence by choosing its first components.
It turns out that the optimal base of $\mathbb{R}^m$ is found by considering the greatest eigenvector of the covariance matrix of the dataset.
The results obtained are:

	VaR_α	ES_α
PCA	138K	120K
Analytical	138K	120K

Table 3: 10-day VaR and ES. PCA vs Analytical Gaussian approach. Notional 1Mio. PCA components: 4.

This is the error of the PCA varying the number of components used:

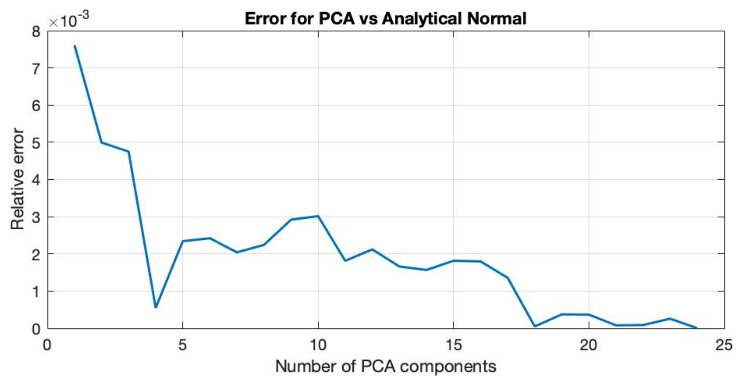

Figure 1: Error in PCA vs Analytic.

1.4 Plausibility Check

To compute an estimation of the order of magnitude of portfolio VaR we use a Plausibility check, to summarize we have:

	VaR_α	ES_α	VaR_{Check}
HS	315K	234K	180K
Bootstrap	315K	227K	180K
WHS	37K	34K	52K
PCA	138K	120K	118K
Analytic Normal	138K	120K	118K

Table 4: PCA vs Standard Gaussian approach .

The formula with which we compute the plausibility check follows from the arguments below:

$$VaR(UL) = N^{-1}(\alpha)\sqrt{\underline{w}\cdot\Sigma\underline{w}} = \sqrt{\sum_{ij}^{d} N^{-1}(\alpha)w_i\sigma_i C_{ij} N^{-1}(\alpha)w_j\sigma_j}$$

But we know that $VaR_{x_i}(\alpha) = u_i$ and $VaR_{x_i}(1-\alpha) = l_i$, by symmetry:

$$sVaR_i = sens\Big(\frac{|l_i| + |u_i|}{2}\Big)$$

So we can identify $sVaR_i = N^{-1}(\alpha)w_i\sigma_i$ and this leads to

$$VaR^{pft} = \sqrt{sVaR \cdot CsVaR}$$

The plausibility check implemented is for a 1-day VaR, while the risk measures for the last portfolio are for 10 days. In the above table we wrote the 10 days plausibility check, while the function is only able to compute the 1 day measure (we would need an extra input argument). The 1 day Plausibility check for the last portfolio is: 37K.

2 Exercise: Full Monte-Carlo and Delta Normal VaR

At the end of the 14th of Sep 2009 we have considered a portfolio formed by stocks of Vivendi for 1,913,220 Euro and the same number of put options with expiry on the 16th of Nov 2009, with strike 23 Euro and volatility equal to 21.4% (r = 3.8% and dividend yield of 5.1%). We computed a 10days 99% VaR via a Full Monte-Carlo and a Delta normal approaches using a 2y Historical Simulation for the underlying.
The value of the losses in the full MC methodology appears to be given by the difference between the initial portfolio and the simulated one:

$$L = \Pi(S_t) - \Pi(\tilde{S}_t)$$

Where $\tilde{S}_t$ is a simulation of the new underlying. In our case the simulation will be historical, hence an historical sampling of past data, where data are taken avoid the overlapping phenomenon. The name full Monte-Carlo refers to the fact that the new portfolio $\Pi(\tilde{S}_t$ is recomputed completely considering also the time difference in the maturity of the derivative, this is not always feasible as we will see.
After ordering the losses in decreasing order we took the $n(1-\alpha)$-th as

$$VaR_\alpha = L^{([n(1-\alpha)],n)}$$

For the calculation of the VaR with the Delta Normal method the derivative price is approximated with a first order expansion:

$$L = \Pi(S_t) - \Pi(\tilde{S}_t)$$

$$\tilde{S}_t = S_t \exp\sum_{s=1}^{H} X_s$$

$$\Delta S = \tilde{S}_t - S_t = S_t\left(\exp\left(\sum_{s=1}^{H} X_s\right) - 1\right)$$

This implies that:

$$\tilde{P}_t = P_t + \Delta S \Delta^P$$
$$L = -N^s \Delta S - N^P \Delta S \Delta^P \tag{1}$$

where H is the time lag between considered in the Var computation (10 days in our case), N^S, N^P are respectively the notional in the Stock and in the Put, Δ^P is the first order sensitivity of the derivative and X_s is a past one day log return.
The results obtained with this methods are summarized in table 5:

MC	Δ
15	34

Table 5: 10-day VaR in absolute value (bps).

We can observe how the first order linearization gives a VaR which is more that double the VaR computed with the MC approach.
A few observations are to be made about this. First of all our portfolio is almost a Δ-hedged portfolio (the Δ of the Put is 97% while we own 100% of the underlying) and hence, as we expect the, VaR of the portfolio is close to zero.
Moreover, and most importantly, we are considering polynomial (linear up to now) expansions of the value of the derivative and then looking at extreme values (a 10 day log return can have a large absolute value), so the linear approximation will not hold far from the spot. In fact the approximation should be better if we were considering a 1-day VaR. If fact for the 1-day VaR we are getting much better results:

MC	Δ
9	14

Table 6: 1-day VaR in absolute value (bps).

The Losses of the Δ methods are linear in the returns, as shown in eq. (1), this in fact is not the same relationship that holds in the real portfolio since the Put option introduces convexity in the payoff.
To improve this methodology we considered, firstly, the Γ method in which we consider the second order derivative wrt to the underlying, now the Loss profile is approximated by:

$$L = -N^S \Delta S - N^P \left[\Delta S \Delta^P - \frac{1}{2} \Delta S^2 \Gamma^P \right]$$

The results obtained with the second order approximation are reported in the following table:

MC	Δ	$\Delta\Gamma$
15	34	6

Table 7: 10-day VaR in absolute value (bps).

MC	Δ	$\Delta\Gamma$
9	15	7

Table 8: 1-day VaR in absolute value (bps).

This results are much more satisfactory even if not as much as expected. Hence we will take a further analysis on the problem.
As already said the polynomial expansion of the Put works in a neighborhood of the Spot price and not further away from the spot, as the following figures exemplify. We also included the sensitivity wrt to passage of time (Θ), which is a shift in value, independently of the return observed.

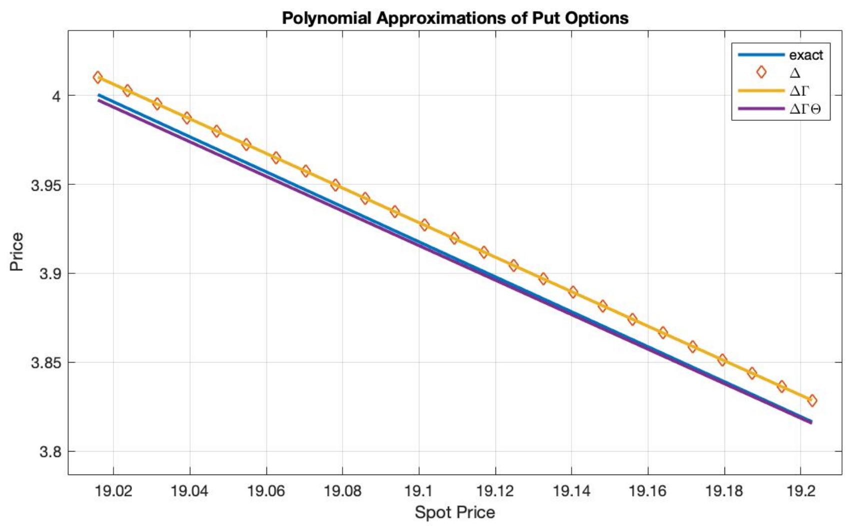

Figure 2: Polynomial expansion of put price.

Moreover we explored the relationship between the loss profile and the returns in the different approaches:

As we can see the expansion works reasonably good in a neighborhood of 0 returns wile for further returns the approximation works really poorly. Let's examine closely the loss profile in the $\Delta\Gamma\Theta$ case. If we assume the position is fully Δ hedged then we can perform the following considerations, where P is the price of the put:

$$L_P = P - \tilde{P} = -\frac{1}{2}\Gamma\ \Delta S - \Theta\ \Delta T$$

and by the well known relation $\Theta = -\frac{1}{2}\Gamma S^2\sigma^2$ we obtain:

$$L_P \sim \frac{1}{2}\Gamma S^2 \left[\sigma^2 \Delta T - \left(\frac{\Delta S}{S}\right)^2\right]$$

which states that the losses are proportional to the difference of the hedging variance σ^2 and the realized variance in the time window ΔT, which is $(\Delta S/S)^2$. Hence for big returns we will incur

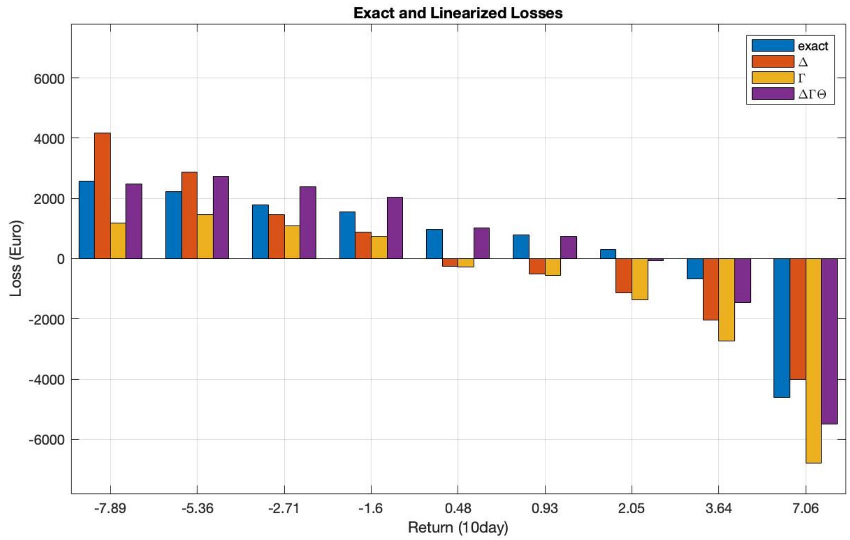

Figure 3: Loss profile.

in profits and losses only when the realized variance is very small, hence for very small returns, as show in the detailed loss profile of the above figure. In the real case (the loss profile of the MC) we can show that the loss profile is equivalent of the one of a portfolio of a bond position (of Notional as the Strike) and a Call with same strike as the put. Hence for negative return the loss would be positive close to zero, while for positive return the profit would be (almost) linear in Spot, as shown in blue in the right picture of figure 4.

To complete the investigation we report the results obtained adding the time sensitivity.
At the end we obtain the following results:

MC	Δ	$\Delta\Gamma$	$\Delta\Gamma\Theta$
15	34	6	12

Table 9: 10-day VaR in absolute value (bps).

In this case study we are in a particular situation in which our portfolio is almost Δ-hedged (it would be perfectly delta hedged if we had 97.13K or 102.9K stocks instead of 100K) and so that approximate measures are not capable to manage such a little sensitivity.
The argument above follows, and is justified in a certain way, from the fact that my expected loss is almost negligible respect to the total value of the portfolio because we are (almost) immunized respect to the change of the stock price.

2.1 Question

If we would like to use the MC method for exotic derivatives with no closed formula we would encounter some computational difficulties. Indeed the method would need a number of valuations that is KN where K is the number of instruments and N is number of MC simulations needed in each valuation of the K instruments. While in the approximate methodologies we will need only K and the values of the sensitives, for each risk factor.

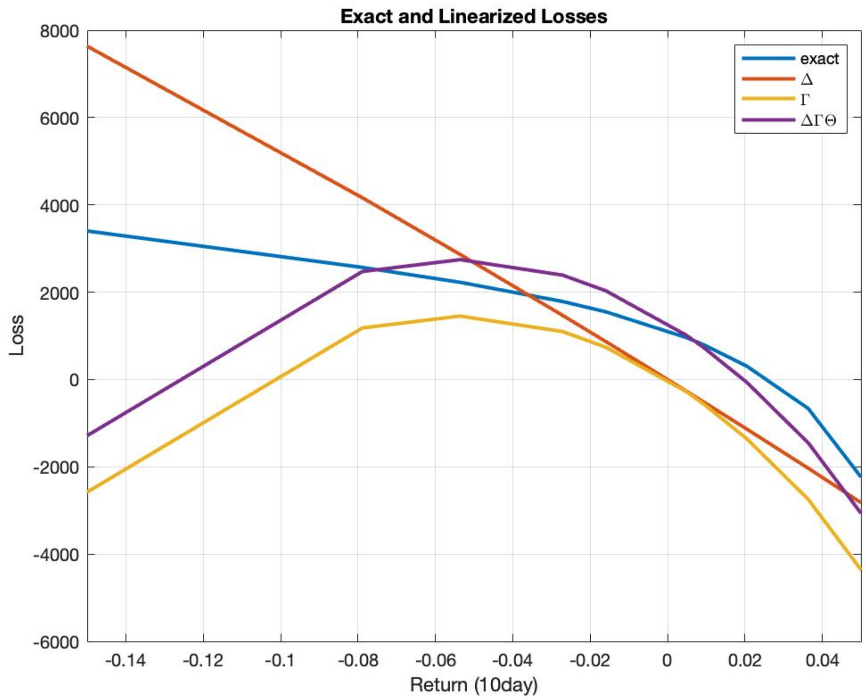

Figure 4: Loss profile.

3 Credit Simulation

Given a constant intensity model, we have that the survival probabilities are distributed as an exponential:

$$P(0,t) = e^{-\lambda t} \implies \mathbb{P}(\tau < t) = 1 - e^{-\lambda t} \implies \tau \sim \mathcal{E}(\lambda)$$

To simulate the default times, we have to use the inverse sampling theorem that in short states that the if random variable X has cdf $F_X(x)$ then $F_X^{-1}(U) \overset{\mathcal{L}}{=} X$. If we consider a maximum likelihood estimator for given an independent sample of $\{\tau_i\}_{i=1,\dots,N}$ we will obtain

$$\hat{\lambda}_{MLE} = \frac{N}{\sum_{i=1}^{N} \tau_i}$$

This estimator however is biased. Since $\sum_{i=0}^{N} \tau_i \sim \Gamma(N, \lambda^{-1})$ we have that $\hat{\lambda}_{MLE} \sim N\Gamma^{-1}(N, \lambda)$ which is the inverse gamma distribution. (abuse of notation: $X \sim N\Gamma$ would mean that the variable X is distributed as a Γ variable times N).
Hence:

$$\mathbb{E}[\hat{\lambda}_{MLE}] = \frac{N}{N-1}\lambda$$

To use an unbiased estimator we will use

$$\hat{\lambda} = \frac{N-1}{\sum_{i=1}^{N} \tau_i} \sim (N-1)\Gamma^{-1}(N, \lambda)$$

To find a confidence interval of level $1-\alpha$ we have to find a pivotal quantity, this would be $T = \lambda \frac{\sum_{i=1}^{N} \tau}{N} \sim \Gamma N, N$ Then we can write that, defining $CI = [LB(\hat{\lambda}), UB(\hat{\lambda})]$

$$1-\alpha = \mathbb{P}\left(\lambda \in CI\right) = \mathbb{P}\left(LB(\hat{\lambda})\frac{N-1}{N} < T < UB(\hat{\lambda})\frac{N-1}{N}\right)$$

$$= F_T\left(LB(\hat{\lambda})\frac{N-1}{N}\right) - F_T\left(UB(\hat{\lambda})\frac{N-1}{N}\right)$$

we can arbitrarily take the first cdf as $1-\alpha/2$ and the second as $\alpha/2$. Then we will take

$$LB(\hat{\lambda}) = F_T^{-1}(1-\alpha/2)\frac{N}{N-1}$$

and

$$UB(\hat{\lambda}) = F_T^{-1}(\alpha/2)\frac{N}{N-1}$$

Here we report the obtained results:

$\hat{\lambda}$	LB	UB
49	46	54

Table 10: Estimation of the default intensity from the simulation (bps).

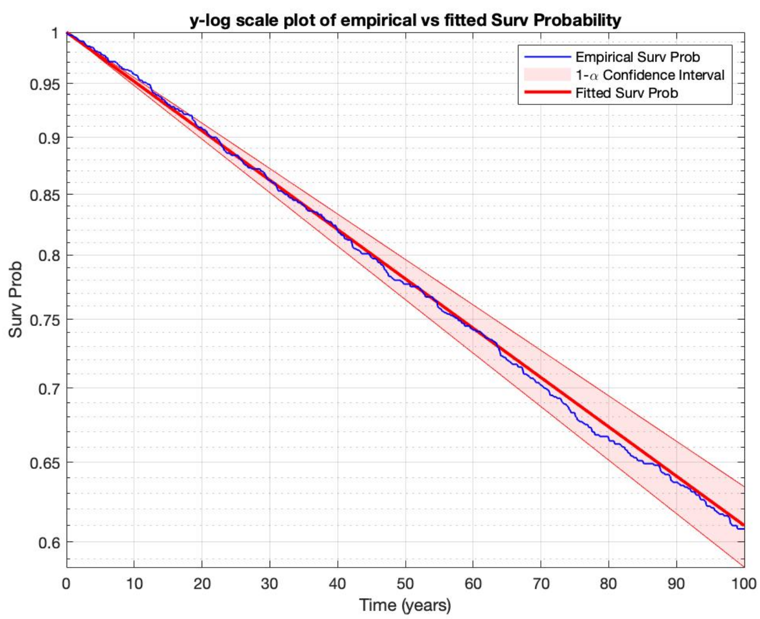

Figure 5: Calibrated and empirical default probabilities

4 Theoretical Exercise

We start from the following (stochastic) differential equation:

$$\begin{cases} dS_t = [r(t) - d(t)]dt + \sigma(t)dW_t \\ S(0) = S_0 \end{cases}$$

Applying Ito's formula to $f(S_t, t) = \log(S_t)$ we obtain:

$$d[\log(S_t)] = \frac{dS_t}{S_t} - \frac{1}{2}\sigma^2(t)dt = \left(r(t) - d(t) - \frac{\sigma^2(t)}{2}\right)dt + \sigma(t)W_t$$

This lead, with a trivial integration, to:

$$S_t = \exp\left[\int_0^T \left(r(s) - d(s) - \frac{\sigma^2(s)}{2}\right)ds + \int_0^T \sigma(s)dW_s\right]$$

Now we want to show that Garman-Kohlhagen formula for an European Call holds for an underlying with interest rates, continuous dividends and volatility as a deterministic function of time:

$$C(0, t; K) = \mathbb{E}\left[e^{-\int_0^T r(s)ds}(S_T - K)^+\right]$$

For the linearity property of the expected value follows:

$$\mathbb{E}\left[e^{-\int_0^T r(s)ds}(S_T - K)^+\right] = e^{-\int_0^T r(s)ds}\left[\mathbb{E}\left[S_T\mathbb{I}_{S_T>K}\right] - \mathbb{E}\left[K\mathbb{I}_{S_T>K}\right]\right]$$

The quantities that we want to compute are the following:

$$\mathbb{E}\left[S_T\mathbb{I}_{S_T>K}\right] = \int_{S_T>K} S_T d\mu \quad \text{and} \quad \mathbb{E}\left[K\mathbb{I}_{S_T>K}\right] = \int_{S_T>K} K d\mu$$

In order to compute the two expectations we have to compute $\mathbb{P}(S_T > K)$:

$$\mathbb{P}(S_T > K) = \mathbb{P}\left(S_0 \exp\left[\int_0^T \left(r(s) - d(s) - \frac{\sigma^2(s)}{2}\right)ds + \int_0^T \sigma(s)dW_s\right] > K\right)$$

We can write the probability above as:

$$\mathbb{P}(S_T > K) = \mathbb{P}\left(-\int_0^T \sigma(s)dW_s < ln(\frac{S_0}{K}) + \int_0^T (r(s) - d(s) - \frac{\sigma^2(s)}{2})ds\right)$$

Using the Ito-isometry and assuming $\sigma \in Ł^2([0, T])$ we can show that:

$$\int_0^T \sigma(s)dW_s \sim N\left(0, \int_0^T \sigma^2(s)ds\right)$$

Thanks to this result we can write the probability above in terms of a normal variable $g \sim N(0,1)$:

$$\mathbb{P}(S_T > K) = \mathbb{P}\left(g < \frac{ln(\frac{S_0}{K}) + \int_0^T (r(s) - d(s))ds}{\sqrt{\int_0^T \sigma^2(s)ds}} = d_2(t)\right)$$

Using the fact that $\mathbb{P}(S_T > K) = \mathbb{P}(g < d_2(t))$ we can write the following equality changing our integration domain:

$$\int_{S_T>K} S_T d\mu = \int_{-\infty}^{d_2(t)} \frac{S_0}{\sqrt{2\pi}} \exp\left(\frac{g^2}{2}\right) \exp\left(\int_0^T \left(r(s) - d(s) - \frac{\sigma^2(s)}{2}\right) ds + g\sqrt{\int_0^T \sigma^2(s)ds}\right) dg$$

With algebraic manipulation this is equal to:

$$\int_{S_T>K} S_T d\mu = S_0 \exp\left(\int_0^T \left(r(s) - d(s) - \frac{\sigma^2(s)}{2}\right) ds\right) \int_{-\infty}^{d_2(t)} \exp\left[-\left(\frac{g^2}{2} - g\sqrt{\int_0^T \sigma^2(s)ds}\right)\right]$$

It remains to complete the square, we can do it with the change of variable $\bar{g} = g - \sqrt{\int_0^T \sigma^2(s)}$, so we have, where $d_1(t) = d_2(t) + \sqrt{\int_0^T \sigma^2(s)}$:

$$\int_{S_T>K} S_T d\mu = S_0 \exp\left[\int_0^T \left(r(s) - d(s) - \frac{\sigma^2(s)}{2}\right) ds\right] \int_{-\infty}^{d_1(t)} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{\bar{g}^2}{2}\right) \exp\left(\frac{1}{2}\sqrt{\int_0^T \sigma^2(s)ds}\right) d\bar{g}$$

This lead to:

$$\mathbb{E}\Big[S_T \mathbb{I}_{S_T>K}\Big] = S_0 \exp\left[\int_0^T (r(s) - d(s))\, ds\right] N(d_1(t))$$

And similarly:

$$\mathbb{E}\Big[K \mathbb{I}_{S_T>K}\Big] = K\mathbb{P}(S_T > K) = KN(d_2(t))$$

To conclude the proof we have that:

$$C(0, t; K) = e^{-\int_0^T d(s)ds} S_0 N(d_1(t)) - e^{-\int_0^T r(s)ds} KN(d_2(t))$$