Building a Portfolio to Have Limited Maximum Loss with SAS® Model Implementation Platform

João Pires da Cruz, Closer Consultoria, Lda and Center for Theoretical and Computational Physics; Carla Tempera, Closer Consultoria Lda; Mariana Eiras Soares, Closer Consultoria Lda; Manuel Fortes, SAS® Institute

ABSTRACT
Building a solid financial entity is the entire purpose of having a portfolio that is able to hold environmental impacts because we expect that as some assets depreciate, others simultaneously appreciate due to diversification. Here we present a model that uses time as diversification factor to build a solid portfolio that holds its value under severe macro economical changes. A theoretical model was built under two very generic assumptions and SAS® Model Implementation Platform was used to evaluate the model using empirical data taken from a real portfolio and to stress the assumptions to extreme environmental changes.

INTRODUCTION
Surprisingly, credit risk is far from being a fully solved problem. Put simply, credit risk relates the possibility of financial losses due to changes in the credit quality of the debtor, with the most severe change being the default event, i.e., the event in which the debtor stops fulfilling credit obligations. Conceptually, the problem may be reduced to measuring the probability of the event’s occurrence—a simple task that in fact lies within a mathematical conundrum. The difficulty is that the assumptions that underlie naive risk measures do not hold in practice. It seems remarkable that our current state of technology can discover planets light-years away from Earth using machine-learning techniques, and yet the solution to the everyday problem of credit decision making continues to be addressed using methods that are only proxies for a proper solution. Here, we address the problem by eliminating the sources of uncertainty that an incorrect probability measure brings using other portfolio parameters to establish a proper risk measure.

We can distinguish credit models into two separate classes (Lando (2009), Gordy (2000)): the structural models and the intensity models. Structural models are inspired by the Merton model (Merton (1974)), assuming that the total value of debtor assets follow a Brownian Motion. With that in mind, if the debtor cannot live up to debt payment requirements, the creditor can take possession of company assets. Meaning, in option jargon, that the amount of debt at the default instant can be looked upon as the strike for an option valuing the company that compares with the firm’s equity. This gives access to a complete set of financial mathematical tools, based on martingale theory, applied to the valuation of the firm’s assets. The expression ‘structural’ comes from the fact that risk is directly related with variations of the financial structure of the firm. Several similar approaches can be adopted based on the same fundamental Brownian Motion assumption, like taking a variable time to default (Black and Cox (1976)) or grouping loans in a portfolio to correlate latent structural changes with systemic risk factors (Vasicek (1987)), the latter being the model that supports the IRB approaches to the Basel Accords (on Banking Supervision (2006, 2011)).

Independently from the indisputable success of this type of approach, there is one issue to resolve. From observations of stock market price time series, equity values do not follow Brownian Motions (da Cruz and Lind (2010)) and, consequently, neither do their assets.
respective liabilities given their mathematical relationship with the residual or equity. Empirical evidences of asset and liability distributions (da Cruz et al. (2015)) show resilient power-law distributions in time which are not consistent with Markovian processes. Also, grouping loans in a portfolio with the goal of making empirical measures will carry the assumption that both the systemic risk and the asset value are assumed to be Gaussian distributions at the instant of the measure which is not and that is also not true (da Cruz and Lind (2010)). Practical approaches, like Basel (on Banking Supervision (2006, 2011)), use the concept of unconditional probability of default, which in a non-stationary system is even worse than improperly estimating the correlation between two non-Gaussian distributions. In summary, despite the interest in structural models stemming from their ability to ascribe causality to credit risk, they are poorly supported in mathematical terms and usually provide erroneous results.

Intensity models, on the other hand, are less dependent on the causality of default events and more on the intensity of events in time. How we model that intensity is the issue. If we consider the economic system as having the same possible states through time, we could model the intensity of defaults based on historical data and consider it constant through time. Again, we cannot assume that because every empirical and theoretical evidence shows us otherwise. For the same reason, we also cannot model it as a Brownian Motion based stochastic process (Schönbucher (2005)), because the parameters must come from similar measures. So, apparently, we cannot use an intensity model unless we could capture the future behavior of the intensity based on economic states that do not exist yet - seemingly an impossible task.

Given this negative background for both model types and the fact that credit has been a profitable business for centuries, long before mathematical modeling, option theory or sophisticated probability measures, we were motivated to the modeling work presented here. Most of credit portfolios in the past were created without any probability-based selection process and were profitable and accessible to a wide band of customers. Our goal is to provide quantitative support for this manner of building a credit portfolio and provide the tools for its fine-tuning. Here we will present a model that minimizes the lack of mathematical support for probability measures by keeping these measures as free variables and using the remaining portfolio parameters as the adjustable ones.

This work is a version of a previous theoretical work (da Cruz et al. (2017)) that was not possible to develop in several scenarios, something that become possible with SAS® Model Implementation Platform. Here we will see that with a real portfolio that theoretical results do not differ from empirical results, even when we apply a relatively intense hit on the environment effect.

**MODEL**

As we have noted, there is reason to be somewhat distrustful of credit selection mechanisms. But it would be absurd to take such a position to the point where we trust no selection mechanisms and thereby give up on credit risk decision making, since we know from practice it is possible to build profitable credit portfolios. Consider, for argument, the alternative of trusting every selection process. The reasoning for not attempting to distinguish the quality of selections mechanisms it that is not possible to know these differences *a priori*. What we want to achieve is a form of temporal independence, and it turns out that this can be done as long as there is a selection mechanism, regardless of which mechanism it is. Let us imagine that we decide to give credit to anyone and not to use any selection process whatsoever. Obviously, we would get a huge rate of default, but that effect could be resolved by setting a proper spread. The problem would come from the effect known as the critical behavior of
economic agents (da Cruz and Lind (2010)), due to the inflationary nature of the economic system. By nature, economic agents are correlated forming a complex network. This means that a default event is not, in principle, independent from others and that a portfolio can be washed out by avalanche effects in the economy, since one bankruptcy can lead to another, that can lead to another, and so on. In other words, default in the portfolio, in this case, depends on history. Thus, when we decide to give credit without selection criteria, what we capture is also the dependency between the default events. Then, statistics measures will fail if we take debtors as independent. To achieve some independence between debtors in our portfolio we need to have selection, any selection, as long as it is made with the aim of avoiding defaults.

Since we got rid of the selection problem and we gain independence between the default events by assuming that there is one (or more) selection process, we still have the stability problem, i.e., how to predict the future if the number of states in the system is not stable? Again, the obvious solution is to not predict. And that is exactly what we do and why we need a complete risk platform as SAS® Model Implementation Platform. But first we need to put some math in the explanation. Let us begin with assumptions:

a) (Certainty) Everybody defaults on a perpetual loan;

b) (Independence) The possibility of an event of default at one instant is independent from the possibility of the event of default in the previous instant.

The first assumption is the basis for the usage of a hazard and survival rate, there is a ‘stream’ of defaults that will fall in time bins in the future. How they will fall, we do not know, but we will take as certain that they will fall independently from each other because there is a selection process.

Let \( N(t) \) be the number of defaults in the time interval \([0, t]\). \( N(t) \) is a counting process sustained by the assumption a). It is a process that increases in unit steps at isolated times and is constant between these times. It can be shown (Van Kampen (1992)) that, with probability one, \( N(t) \) is a (homogenous) Poisson process (Kingman (2005)) with parameter \( \lambda \). Hence

\[
p(N(t) = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}
\]

To show a duality between the waiting time until a fixed number of defaults and the number of defaults that happen in a fixed length of time (which we represent by \( N \)), let \( W(k) \) be the distribution of the waiting time until the \( k \)-th event occur. If \( W(k) \leq t \) then we observed the \( k \)-th default before instant \( t \) occur. Then, at least \( k \) events occur during the time interval \([0, t]\), i.e. \( N(t) \geq k \). Using the duality, we can say that

\[
P(W(k) \leq t) = P(N(t) \geq k)
\]

\[
P(W(k) > t) = 1 - P(W(k) \leq t) = P(N(t) < k)
\]

which leads to

\[
P(W(k) \leq t) = 1 - \frac{\Gamma(k + 1, \lambda t)}{\Gamma(k + 1)}
\]

where \( \Gamma(a, b) \) is the upper incomplete gamma function and \( \Gamma(a) \) is the gamma function. Thus, the probability density function of \( W(k) \) in time will be given by deriving \( P \) in order to time and
\[ p(k; \lambda, t) = \frac{\lambda^k t^{k-1}}{\Gamma(k)} e^{-\lambda t} \]

So, we stopped looking at the default as a timeless default rate measured over the portfolio and now we are looking for the default of one credit in a time horizon. What did we gain from that? Statistical consistency. When we divide the number of defaults by the number of total credits, a change in the latter means the same as a change in the surrounding conditions of the portfolio. In other words, we are comparing things that are not comparable, like throwing a dice with 6 sides and, afterwards, a dice with 12 sides. They are both dice, and any ratio would give a result between 0 and 1, but they are not the same system and we should not merge results into the same measuring apparatus. By ignoring the total credits in the portfolio and measure through time, we ‘close’ the space of events. We will have defaults that distribute themselves in time because we cannot predict the future. Moreover, we assume that everybody defaults sometime in the future, like an irresistible attraction. With that, all underlying assumptions that govern statistics and probability theory are present and verifiable.

We should note that \( k \) is not necessarily an event of default. And, in fact, we will not take it as such. \( k \) can be taken as a ‘distance in wealth to default’. As if it was steps in a ladder that we should step down to reach the floor, which in this case represents default. This is, in practice, the best way to see \( k \) because default events should be rare events and caused by those steps in the ladder. The greater the number of steps in the ladder to default, the more effective is the selection process.

Now, we have two parameters governing our probability density function: \( k \) and \( \lambda \). We will only calculate the former from the data. Why? \( 1/\lambda \) is equal to characteristic time to default (Kingman (2005)). \( k \) is a short-term parameter in the sense that it can be measured from the very short-term events \( t = 1 \). The greater \( k \), more steps to default and less probable to default on the first instants of the credit. In terms of credit selection, this means we say that the debtor will not default on the next few months. Obviously, greater \( k \) means that the probability of default will be stretch in time and it will be less probable to have a default in the earlier instants of the credit. Since \( k \) can be measured in the first interval of time (in \( 1/\lambda \) time units) and \( k \) depends on the selection, it can be easily measured numerically taking empirical data without major concerns about the stability of environment. It is all about internal processes.

Having measured the parameter that reflects selection, \( k \), we are left with the parameter that reflects the market, \( \lambda \). The goal of risk management is to be able to quantify independently from market conditions and that is what we will do in the next steps. First, let us fix a value for \( \lambda \), for now, and start to look at the credit contracts. Let us take a non-perpetual credit contract that can be looked as time sequence of interest at maturity deals, with fixed rate, \( r \), a start date, \( t_s \), and a maturity \( M \) (see Figure 1. ).

![Figure 1. Schematics of contract’s cashflow.](image-url)
With this in mind we know that the value of the contract at an instant \( t = 0 \) as

\[
v(r, t) = (1 + r_m)^{-t_0} \sum_{t = t_0}^{M} C(1 + r)^{t - t_0} \approx e^{-r_m t_0} \sum_{t = t_0}^{M} C e^{r(t - t_0)}
\]

where \( r_m \) is the market interest rate that we consider as constant for simplicity sake and, also, and \( C \) represents the nominal amount of the loan that we also consider fixed and credit risk free (this is simply discounting cashflows). Even if we consider it as variable, the amount would not change due to customer risk changes so, for simplicity and without losing generality, we will take is constant.

Now, considering the expression for the probability of default the time variable used here is not the same. The time in the expression of the probability of default is the time from the selection, i.e., from the origination of the contract. To keep both times compatible, we write the probability as

\[
p(k; \lambda, t, t_0) = \frac{\lambda^k (t - t_0)^{k-1}}{\Gamma(k)} e^{-\lambda(t-t_0)}
\]

and the expected risk adjusted value of the contract in time is written as

\[
v(r, t, t_0) \approx e^{-r_m t_0} \sum_{t = t_0}^{M} \left[ 1 - C \frac{\lambda^k (t - t_0)^{k-1}}{\Gamma(k)} e^{-\lambda(t-t_0)} \right] e^{r(t - t_0)}
\]

Now, the way we are going to build the portfolio is to put one contract after the other, i.e., one with \( t_0 = 0 \), another with \( t_0 = 1 \), and so on; and we assume that the contracts are exactly equal (we can do it on average). So, we integrate the last expression in order to \( t_0 \) and \( t \) which leads to

\[
v(r, \lambda) \approx g \left[ 1 + \frac{r}{e^r M - 1} \left( \frac{\lambda}{\lambda - r} \right)^{k} \left( \frac{\Gamma(k, M(\lambda - r))}{\Gamma(k)} - 1 \right) \right]
\]

where \( \Gamma(\cdot, \cdot) \) is the upper incomplete gamma function, \( \Gamma(\cdot) \) is the normal Gamma function that converts to factorial when the argument is an integer and \( g \equiv g(C, r_m) \) is a function of the nominal value of the contracts and the market interest rate which here we will ignore since it does not depend of credit risk (neither \( \lambda \) or \( r \)) and we will just bear in mind that there is a present value of the portfolio at instant \( t = 0 \), that depends on several factors, but not on the ability of the customers to repay the loan.

To understand the behavior of the \( v(r, \lambda) \), we can make \( r \ll \lambda \) without loss of generality, and that reduces to

\[
v(r, \lambda) \approx g \left[ 1 + \frac{r}{e^r M - 1} \left( \frac{\Gamma(k, \lambda M)}{\Gamma(k)} - 1 \right) \right]
\]

Since \( 1/\lambda \) is the characteristic time to default, the r.h.s of the equation depends mainly on how high is \( \lambda M \). A high \( \lambda M \) means a very low time to default with a very high maturity. Which seems very unlikely to happen in a consistent way. On the other conditions, \( v(r, \lambda) \) is very close to be constant.

Now, we are interested on the behavior of the portfolio, the value of \( v(r, \lambda) \), in terms of credit risk. Since we can assume that our selection process is what it is, people or machines selecting on ratings or scorings, we are especially interested on how it behaves with variations of \( \lambda \). For
that we need to apply to our portfolio several values of $\lambda$ and see how it behaves, historically and presently in order that we can manage the portfolio without the need for mathematical misleading assumptions.

IMPLEMENTATION

The first step was the importation of all the data with SAS Studio. The data consists of a portfolio with three different types of contracts Consumer Credit (CRED), Equipment Leasing (MOB) and Real Estate Leasing (IMOB).

The SAS® solution used to implement the model proposed in the paper was the SAS® Model Implementation Platform, in which all the following steps took place:

1. We started implementing the model in Model Groups, where we constructed the MDL_Credit_Multihorizon model group.
2. Next, we moved to Model Implementation: We constructed a model Run_Multihorizon where we considered 456-time horizons (in months) which correspond to the month of the start of the oldest contract, until the end of the last one. For every contract in study, the respective cash flows were simulated from the start date to the final one of the contracts. (see SAS® Model Implementation Platform 3.1 User’s Guide )

   a. **Source Risk Cube:** We chose the data set to use. There are around 70,239 contracts and the ones that were considered in the model, are those with the starting date 1st of January 2002

   b. **Scenario Data:** We selected the simulated macroeconomic scenarios to be considered, which we previously defined as: 6 scenarios for each type of product. Those were: the base one, two optimistic scenarios with shocks of +20% and +70%, and the exact opposite: -20% and -70%. An additional scenario was set which was the one with the real defaults occurred. We also defined the interval as month, the number of horizons as 456 and the currency as euro

   c. **Maps and Methods:** In this part we mapped the model group to consider (which was previously created in the model groups tab) and the defined output variables for the evaluation of the models.

   d. **Advanced options:** No changes were made in this part, we used the default options.

   e. **Execution Options:** We finally introduced the cash flow analysis. Clicking in Analyze cashflow and selecting the time bucket scheme as monthly and duration modified.

The code for MDL_Credit_Multihorizon is the following:

```plaintext
BEGINBLOCK MAIN;
put simulationhorizon=;
put instid=;
currentdate=_date_

dtend=intnx('MONTH',data_realizacao,nrperiod,'s');
dtbegin=data_realizacao;
```
if year(currentdate)=year(dtbegin) and month(currentdate)=month(dtbegin) then do;
    outstanding=MONTANTE_INICIAL;
    _cashflow_.balance.matamt[simulationhorizon+1]=outstanding;
end;

if data_realizacao > currentdate then return;
if dtend < currentdate then return;

_idx simulationhorizon+1;

_outstanding = _cashflow_.balance.matamt[idx-1];
_rate = tx_ini_ef/100;
_interest = outstanding*(rate/12);
_capital = sum(prestacao,-interest);
_outstanding = max(0,sum(outstanding,-capital));
_outstanding = _cashflow_.balance.matamt[idx]=outstanding;
_outstanding = _cashflow_.interest.matamt[idx]=interest;
_outstanding = _cashflow_.capital.matamt[idx]=capital;
modelname=compress("Lambda_"!!insttype);

actual_mf=MF_LAMBDA;
call run_model(modelname);
model_ret=_model_.result;

actual_lambda=model_ret*actual_mf;
actual K=k_FACTOR;
num=(actual_lambda**actual_k)*((idx-1)**(actual_k-1));
den=fact(round(actual_k-1));
p=(num/den)*exp(-actual_lambda*(idx-1));

_b = _cashflow_.balance.matamt[idx]* incarcerationrate;

_d = d-p;
if rf_default=1 then do;
    d=1;
    if default_date not in (0,.) then do;
        if _date_ >= default_date then do;
            d=0;
        end;
    end;
end;
The reason why we used SAS® Model Implementation Platform was to take advantage of the platform to run the same code, on the same portfolio using several scenarios, both theoretical and empirical.

RESULTS
To understand the results a proper description of the portfolio we are using is needed. The portfolio is composed by 70456 contracts of consumer credit, equipment leasing (including commercial and non-commercial cars) and real estate leasing that started from Jan 2002 to Jan 2015. In terms of type of client, 75% are companies and 25% are individuals, all Portuguese or completely exposed to the Portuguese market. This is an important characteristic of the portfolio since the Portuguese economy suffered major crisis between 2008 and 2012. The first was due to the global impact of the US credit crisis in 2008 and the second due to the Euro Sovereign Debt Crisis in which the Portuguese Republic was at the ground zero that led to the bail out of the state (see Figure 2).

![Figure 2 Portuguese economy (unemployment and GDP)](chart)

So, considering the theoretical model described above, we made the portfolio suffer an impact of 20% and 70% on the inverse of the characteristic time to default ($\lambda$) from the average measured from real defaults and called the scenarios SCN_BASE for the measured average
(see Figure 3), SCN_M20 and SCN_P20 to an impact of minus 20% and plus 20%, respectively; and SCN_M70 and SCN_P70 have the same meaning for a 70% impact.

![Figure 3 Normalized portfolio value](image)

We called the normalized portfolio value to the reason between the value of the portfolio with risk and the value of the portfolio without risk in the same date. Notice the scale for which the normalized portfolio value is plotted, between .9875 and 1.0, that gives a maximum of 1.25% loss considering the theoretical probabilities. This value is so low because in the measure we are considering that an equal portfolio is being rolled periodically, meaning that the commercial result of placing credit would be constant in time.

The real portfolio is not like that because in the crisis of 2010 the banks were forced to cut on credit due the bail out agreement (Memorandum of Understanding On Specific Economic Policy Conditionality), so commercial production drop considerably (see Figure 4).

![Figure 4 Evolution of the portfolio value with production](image)
In Figure 4 all series are normalized to the correspondent maximum. We can see that as credit production is more or less stable, the value of the portfolio is growing to its maximum, despite of the 2008 crisis. The value starts to drop as production also drops, both in simulated scenarios (SCN_BASE, SCN_M20, SCN_P20, SCN_M70, SCN_P70) and in the real scenario (SCN_REAL).

To see how the simulated scenario performed when compared with the theoretical one, we normalized the real scenario (SCN_REAL) with the simulated one (SCN_BASE). The result is presented in Figure 5.

![Figure 5 Real/Theoretical performance](image)

Again, until the production of credit is decreased, the real scenario performs as good as the theoretical and we can say that credit risk is diversified in time. As credit production is cut, the diversification effect disappears, and losses begin to be relevant.

**CONCLUSION**

Why are these results expectable? Mathematically, there is only a probability measure if we can have a stable number of possibilities in time and that cannot happen unless we consider a time frame for which we can consider that everyone will eventually ‘die’. This is the closing condition necessary to assume that there is a natural tendency for a credit to default, and that it will happen in some bin of time in the future, even if it is in infinite. From this assumption, is straightforward that we will have a Gamma distribution for this behavior by aggregating the Bernoulli trials in each time bin.

As we can tell from analytical, simulated and empirical results, the risk of a credit portfolio can be efficiently erased by time diversification, as long as the commercial production remains constant. Or, in other words, this technique of time diversification can be used to build solid portfolios from existing ones by selecting credits to a ‘riskless’ pool. From a practical point of view, the possibility of using a platform that can run several risk scenarios at the same time with several combinations of portfolios, like SAS® Model Implementation Platform make this technique an interesting one.

This result is interesting when compared with the generally accepted Basel (capital rules) and impairment rules. From the analytical, simulated and empirical results, we should reach the
conclusion that the value of a portfolio is considerably stable in terms of average contract value, as long as business keeps rolling at the same rate. But the generally accepted procedure, and dominant in global regulation, is to measure probabilities by considering the default rate at one instant in time by averaging it through the past instants. The problem is that this is not a probability measure since it does not fill the Kolmogorov axioms that any textbook on probability theory explains.

Also, from this way of dealing with a credit portfolio, as we can tell, there is no room for concepts like unexpected loss since every loss can be associated to a decrease in selling and not to an abstract, and badly sustainable, notion of systemic catastrophe.

There are some limitations that we need to take into account. First, the assumption of the independence of the defaults from instant to instant. We stated that the selection of the credits would give the independence. Theoretically, there are conditions of such severity that the independence assumption drops, and the above results are no longer valid. However, we should take into account that the Portuguese conditions were quite severe, and that the empirical data corroborates the analytical result. Second, there is no guarantee that a commercial policy can be kept under any market conditions without significantly changing the selection process. Nevertheless, the suggested strategy is that a bank should work with parameters that are not probability measures, namely interest rate and maturity.

The reader should take in account that the value of the portfolio is not immune to interest rate conditions in the market. The goal in this work was to look exclusively to credit risk without considering interest rate risk, for which there already is a panoply of hedging instruments.

REFERENCES


Vasicek, O. 1987, ‘Probability of Loss on Loan Portfolio’, KMV.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

    Joao Pires da Cruz
    Closer Consultoria, LDA
    +351 913591042
    joao.cruz@closer.pt

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.