Approximate Derivative Pricing for Large Class of Homogeneous Assets

Patrick Gagliardini  Christian Gouriéroux

First version: March 2008
Current version: June 2008

This research has been carried out within the NCCR FINRISK project on “New Methods in Theoretical and Empirical Asset Pricing”
Approximate Derivative Pricing for Large Class of Homogeneous Assets*

Patrick Gagliardini†
University of Lugano and Swiss Finance Institute

Christian Gouriéroux
CREST (Paris), CEPREMAP and University of Toronto

First Version: March 2008
This Version: June 2008

*Acknowledgements: We thank F. Galli, M. Gordy, A. Monfort and the participants at the International Financial Risk Forum 2008 in Paris and at the FINRISK Research Day 2008 in Gerzensee for very useful comments. The first author gratefully acknowledges financial support of the Swiss National Science Foundation through the NCCR FINRISK network. The second author gratefully acknowledges financial support of NSERC Canada and the chair AXA/Risk Foundation: "Large Risks in Insurance".

†Corresponding author: Patrick Gagliardini, Faculty of Economics, University of Lugano, Via Buffi 13, CH-6900 Lugano, Switzerland. Phone: +41 58 666 4660. Email: gagliarp@lu.unisi.ch.
Approximate Derivative Pricing for Large Class of Homogeneous Assets

Abstract

We consider an homogeneous class of assets, whose returns are driven by an unobservable factor. We derive approximated prediction and pricing formulas for the future factor values and their proxies, when the size $n$ of the class is large. Up to order $1/n$, these approximations involve solely well-chosen summary statistics of the basic asset returns, but not lagged factor values. The potential of the approximation formulas seems quite large, especially for credit risk analysis, which has to consider large portfolios of individual loans or corporate bonds, and for longevity risk analysis, which involves large portfolios of life insurance contracts.

Keywords: Derivative Pricing, Large Portfolios, Credit Risk, Longevity Risk, Credit Default Swap, Mortality Linked Securities, Default Correlation, Granularity Adjustments.
1 Introduction

The stochastic volatility model is a well-known example of factor model introduced for pricing derivatives [see e.g. Hull, White (1987), Heston (1993), Ball, Roma (1994)]. Under the historical probability, a stochastic volatility model can be defined as:

\[
\begin{align*}
    dS_t &= \mu S_t dt + \sigma_t S_t dW^S_t, \\
    d\sigma^2_t &= \alpha (\sigma^2_t - \beta) dt + \gamma \sigma_t dW^\sigma_t,
\end{align*}
\]

where \((W^S_t)\) and \((W^\sigma_t)\) are two independent (or correlated) Brownian motions. Compared to the standard Black-Scholes model [Black, Scholes (1973)], the introduction of the stochastic volatility as an additional factor induces a one-dimensional incompleteness, when only stock prices \(^1\) are known. More precisely, there exists a multiplicity of admissible prices for the derivatives written on stocks, which depends on the selected price of volatility risk.

To describe this multiplicity, it is generally assumed that the investors’ information includes the current and past values of both \(S_t, \sigma_t\), and that the joint process is Markov of order 1 under the admissible risk-neutral probabilities. Then, at time \(t\), the derivative prices depend on both \(S_t\) and \(\sigma_t\). The presentation of the stochastic volatility model is usually completed by a discussion of the investors’ information. In particular, the inclusion of \(\sigma_t\) in the information set is typically justified by the possibility to use the prices of highly traded derivatives to reconstitute the stochastic volatility value. Even if the volatility is not a traded asset, we only need some volatility dependent assets, such as derivatives [see the discussions in Garman (1976), Ross (1978), or Hull, White (1987), p. 281-283]. However, if the stock \(^2\) is the only traded asset, the value of the stochastic volatility cannot be deduced from observed asset prices, and the pricing formula cannot be applied. This drawback arises typically at the beginning of derivative markets, when a coherent quotation for derivative prices not yet highly traded has to be proposed. In fact, the use of the above stochastic volatility model requires at least one highly traded derivative\(^3\).

\(^1\)And the risk-free rate, assumed to be constant.
\(^2\)And a zero-coupon bond.
\(^3\)And even more to calibrate the price of volatility risk.
The aim of this paper is to explain how this drawback for the first quotation of derivatives can be circumvented if we observe the prices (or returns) of a large homogeneous class of assets driven by a same factor. Loosely speaking, the unobserved factor value and the derivative prices can be approximated by means of solely the observed asset prices (returns) at order $1/n$, where $n$ is the size of the class. Thus, it is not necessary to observe at least one derivative price to be able to price coherently the other ones.

The interest in homogeneous class of assets is growing rapidly with the new Basel 2 and Solvency 2 regulation and the introduction of basket or factor derivatives. The portfolios involved in the balance sheet of a bank or an insurance company can involve several millions of individual assets, and it is suggested to cluster them into homogeneous classes to facilitate the computation of the total VaR by simulation, for instance. Examples of large homogeneous classes of assets are:

(i) a set of corporate bonds for firms with given industrial sector, rating and country;

(ii) a cohort of mortgages with identical contractual interest rate, maturity, pattern of monthly payment and identical borrower’s rating;

(iii) a set of life insurance contracts with similar design and contractors with identical age.

A stationary common factor is usually introduced in examples (i) and (ii) to capture the so-called default correlation [see e.g. Vasicek (1987), Gupton et al. (1997), Schoenbucher (2001), Gordy, Heitfield (2002)]. A non-stationary factor is usually introduced in example (iii) to capture the so-called longevity risk, associated with the increase in life expectancy [see e.g. Lee, Carter (1992), Dahl, Moller (2004), Cairns, Blake, Dowd (2006), Schrager (2006), Gouriéroux, Monfort (2007)].

The notion of homogeneous class of risks is defined in Section 2 by means of the conditional historical and risk-neutral densities of risk variables $y_{1,t}, ..., y_{n,t}$, say, given the underlying factor path, and the historical and risk-neutral factor dynamics. We particularize this notion to different types of risks to account for quantitative as well as qualitative risk variables. Section 3 provides a convenient approximation of the historical (resp.
risk-neutral) density of a future factor value \( f_{t+h} \) given the information available at time \( t \), which includes the current and past values of \( y_{1,t}, \ldots, y_{n,t} \), for large \( n \). We discuss the consequences of this approximation formula in terms of first quotation of derivative prices, and the associated term structure of derivative prices. In Section 4 the methodology is illustrated by discussing carefully the approximate pricing of derivatives written on a common factor proxy, such as for instance basket default swaps and longevity bonds. A numerical illustration to the pricing of basket default swaps is given in Section 5. Section 6 concludes. The proofs are gathered in the Appendices.

2 Homogeneous class of assets

Let us consider \( n \) assets with values observed at \( T \) different dates. The values are denoted \( y_{it}, i = 1, \ldots, n, t = 1, \ldots, T \). These values have different interpretations and ranges according to the type of assets. For a class of stocks, \( y_{it} \) can be the return between \( t - 1 \) and \( t \) and is real valued. For a class of corporate bonds, \( y_{it} \) can be the spread, that is, the difference between the corporate interest rate and the risk-free interest rate. In this case \( y_{it} \) is non-negative. Finally, if we consider a class of digital Credit Default Swaps (CDS) with given time-to-maturity, \( y_{it} \) can be the ratio of the CDS price to the price of the risk-free zero-coupon bond with the same maturity. Then, \( y_{it} \) takes values between 0 and 1.

2.1 Definition and assumptions

The notion of homogeneous class is defined for instance in Gouriéroux, Tiomo (2007), Chapter 7.

**Definition 1:** The class of assets is homogeneous under the historical probability (resp. the risk-neutral probability), if and only if the joint historical (resp. risk-neutral) distribution of processes \((y_{1,t}), \ldots, (y_{n,t})\) is invariant by permutation.

In other words, the assets are exchangeable, when we focus on their distributional properties. The exchangeability property can be equivalently written in terms of underlying factors by applying de Finetti’s Theorem [de Finetti (1931)] and its generalization by Hewitt
and Savage (1955). In the rest of the paper, we consider a one-factor model, such that:

**Assumption A.1:** *Under the historical probability, the processes \((y_{1,t}), \ldots, (y_{n,t})\) are independent given the factor path \((f_t)\).*

**Assumption A.2:** *The conditional historical density of \(y_{1,t}, \ldots, y_{n,t}\) given the past values of the \(y_{i,t}\)’s, and the current and past values of the factors, is of the type \(\prod_{i=1}^{n} h(y_{it}|f_t)\).*

Thus, the conditional density is driven by the current factor value only, is independent of the date, and the variables \(y_{1,t}, \ldots, y_{n,t}\) are i.i.d. conditional on the current factor value. This explains the terminology ”conditionally independent risk model” introduced in the literature [see e.g. Schoenbucher (2001)]. Factor \(f_t\) represents systemic risk. Further, the dynamics of variables \(y_{i,t}, i = 1, \ldots, n\), is through the dynamics of factor \(f_t\) only, which is specified next.

**Assumption A.3:** *Under the historical probability, the factor process is Markovian with transition density \(g(f_{t}|f_{t-1})\).*

Under Assumptions A.1, A.2 and A.3, the joint conditional distribution of \(y_{1,t+1}, \ldots, y_{n,t+1}, f_{t+1}\) given the current and past values of the \(y_{i,t}\)’s and \(f_t\) is:

\[
L(y_{1,t+1}, \ldots, y_{n,t+1}, f_{t+1}|y_{1,t}, \ldots, y_{n,t}, f_t) = g(f_{t+1}|f_t) \prod_{i=1}^{n} h(y_{i,t+1}|f_{t+1}), \quad (2.1)
\]

where \(y_{i,t}\) denotes \(y_{i,t}, y_{i,t-1}, \ldots\), and similarly for \(f_t\). It is easily checked directly from equation (2.1) that the assets are exchangeable under the historical probability. The decomposition (2.1) has its analogue for any maturity \(h\). The historical joint conditional density at horizon \(h\) is:

\[
L(y_{1,t+h}, \ldots, y_{n,t+h}, f_{t+h}|y_{1,t}, \ldots, y_{n,t}, f_t) = g_h(f_{t+h}|f_t) \prod_{i=1}^{n} h(y_{i,t+h}|f_{t+h}), \quad (2.2)
\]

where:

\[
g_h(f_{t+h}|f_t) = \int \cdots \int g(f_{t+h}|f_{t+h-1}) \cdots g(f_{t+1}|f_t) df_{t+h-1} \cdots df_{t+1}.
\]

The set of assumptions above is now completed by assumptions on the investors’ information set and on the stochastic discount factor (sdf) \(m_{t,t+1}\), which characterizes the
change of measure to pass from the historical probability to the pricing operator [see e.g. Harrison, Kreps (1979), Hansen, Richard (1987)].

**Assumption A.4:** The investors’ information set at date $t$ is $\Omega_t = (y_{1,t}, \ldots, y_{n,t}, f_{t-1})$.

The information set of the investors includes the current and past values of the $y_{i,t}$’s, and the past values of the factor.  The current factor value $f_t$ is unobservable for the investor at date $t$. It means that the variables $y_{i,t}$, $i = 1, \ldots, n$, are observed at the beginning of period $(t, t+1)$, whereas $f_t$ is observed by the investor at the end of that period only. Note that in continuous time, it is not possible to distinguish between the current factor value $f_t$ and the most recent lagged value $f_{t-1}$, say, for a factor with continuous path. In discrete time, the most recent factor value $f_{t-1}$ differs from $f_t$, and Assumption A.4 on information becomes relevant.

**Assumption A.5:** The sdf is of the type $m_{t,t+1} = m(f_t)$.

The current factor value unobservable at time $t$ is included in the sdf $m_{t,t+1}$ and is used for risk correction and time discount between $t$ and $t+1$. The existence of a sdf is a consequence of the absence of arbitrage opportunities assumption. In general, the sdf depends on the investors information, that is $y_{1,t+1}, \ldots, y_{n,t+1}, f_t$, and on the market, in particular on the number $n$ of assets. Under Assumption A.5, the sdf is supposed to depend only on the current value of the common risk factor, and neither on the idiosyncratic risks specific of the individual assets, nor on the size $n$ of the homogeneous class. This latter assumption is standard for credit risk [see e.g. Gouriéroux, Monfort, Polimenis (2006)], or for mortality linked securities (MLS) [see Schrager (2006)].  

---

\(4\) Lamb et al (2008) consider a dynamic model where the factor follows a transformed autoregressive process. In their approach, the investors’ information set consists of the current and past factor values.

\(5\) This standard assumption merits to be discussed in some detail. Indeed, let us assume that $m_{t,t+1} = m(f_t)$ and suppose an infinite population, $n = \infty$. Then, $f_t$ becomes known and the short term interest rate is $r_t = -\log E|m_{t,t+1}|_{\Omega_t} = -\log m(f_t)$. In some sense, either function $m$ and the interest rate are constant, or factor $f_t$ can be identified with the interest rate up to a given transformation. If we want to get a pricing model with constant interest rate $r$ and non constant factor, it is necessary to suppose $m_{t,t+1} = m(n, f_t)$, with $m(\infty, f_t) = \exp(-r)$, and if we want to get a stochastic interest rate $r_t$ and another stochastic factor $f_t$ to suppose $m_{t,t+1} = m(r_t, n, f_t)$, with $m(r_t, \infty, f_t) = \exp(-r_t)$. 

7
A.5, the short term pricing operator is:

\[ p(y_{1,t+1}, \ldots, y_{n,t+1}, f_t|\Omega_t) = m(f_t)g(f_t|\Omega_t) \int \prod_{i=1}^{n} h(y_{i,t+1}|f_{t+1}) g(f_{t+1}|f_t) df_{t+1}, \quad (2.3) \]

where:

\[ g(f_t|\Omega_t) = \frac{g(f_t|f_{t-1}) \prod_{i=1}^{n} h(y_{i,t}|f_t)}{\int m(f_t)g(f_t|\Omega_t) df_t}, \quad (2.4) \]

denotes the density of \( f_t \) given the investor information \( \Omega_t \) at \( t \). The pricing operator depends on \( \Omega_t \) through the current observations \( y_{1,t}, \ldots, y_{n,t} \) and the past factor value \( f_{t-1} \) only. The price at \( t \) of the short term zero-coupon bond is:

\[ B(t, t+1) = \int p(y_{1,t+1}, \ldots, y_{n,t+1}, f_t|\Omega_t) dy_{1,t+1} \ldots dy_{n,t+1} df_t = \int m(f_t)g(f_t|\Omega_t) df_t. \]

From the short term pricing operator (2.3), we deduce the short term risk-neutral distribution by standardizing with the zero-coupon bond price. The risk-neutral conditional distribution is:

\[ l^*(y_{1,t+1}, \ldots, y_{n,t+1}, f_t|\Omega_t) = \frac{p(y_{1,t+1}, \ldots, y_{n,t+1}, f_t|\Omega_t)}{B(t, t+1)} = g^*(f_t|\Omega_t) \int \prod_{i=1}^{n} h(y_{i,t+1}|f_{t+1}) g(f_{t+1}|f_t) df_{t+1}, \]

where:

\[ g^*(f_t|\Omega_t) = \frac{m(f_t)g(f_t|\Omega_t)}{\int m(f_t)g(f_t|\Omega_t) df_t}. \]

Thus, the historical and risk neutral conditional distributions of \( y_{1,t+1}, \ldots, y_{n,t+1} \) given \( f_t \) are the same, whereas the historical and risk-neutral distributions of \( f_t \) given \( \Omega_t \) are related by the sdf \( m(f_t) \).

2.2 Examples

Let us now derive the expression of the so-called micro-density \( \prod_{i=1}^{n} h(y_{i,t}|f_t) \) [see Gagliardini, Gouriéroux, Monfort (2007)] for standard models encountered in the literature for homogeneous class of stocks, corporate bonds, or digital CDS. For expository purpose,
we consider one-factor models. In general, the factor has a nonlinear effect. It can be a stochastic mean, a stochastic variance, a stochastic overdispersion, ...

i) Linear factor model

Let us consider a linear one-factor model:

\[ y_{it} = a + bF_t + \sigma u_{it}, \quad i = 1, \ldots, n, \]

where the errors \( u_{it} \) are independent standard Gaussian variables. The common factor \( F_t \) impacts the conditional mean of the variables \( y_{it} \). When the variables \( y_{it}, i = 1, \ldots, n, \) are stock returns, and the factor \( F_t \) is the market portfolio return, we get the standard market model written for an homogeneous class of stocks, since the alphas, betas and idiosyncratic volatilities are stock independent. The conditional correlation between any two assets is zero, but the (unconditional) correlation is non-zero when the factor is integrated out. By introducing the transformed factor \( f_t = a + bF_t \), this model satisfies Assumptions A.1 and A.2 with micro-density:

\[
\prod_{i=1}^{n} h(y_{i,t}|f_t) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_{it} - f_t)^2 \right\}.
\]

ii) One-factor stochastic volatility model

Let us consider the model:

\[ y_{it} = \mu + f_t^{1/2} u_{it}, \quad i = 1, \ldots, n, \]

where the factor \( (f_t) \) is a positive Markov process, and the errors \( u_{it} \) are independent standard Gaussian variables. In this model, the joint stochastic volatility-covolatility matrix is \( \Sigma_t = f_t Id_n \), where \( Id_n \) denotes the identity matrix of order \( n \). The associated micro-density is:

\[
\prod_{i=1}^{n} h(y_{i,t}|f_t) = \frac{1}{(2\pi)^{n/2}} \exp \left\{ -\frac{n}{2} \log f_t - \frac{1}{2f_t} \sum_{i=1}^{n} (y_{it} - \mu)^2 \right\}.
\]

iii) One-factor corporate spread model
Corporate spreads are positive and their dynamics is generally specified to be compatible with affine term structures [see e.g. Duffie, Filipovic, Schachermayer (2003) in continuous time, Darolles, Gouriéroux, Jasiak (2006) and Gouriéroux (2007) in discrete time]. For instance, let us assume that the conditional distribution of \( y_{it} \) given \( f_t \) is a gamma distribution \( \gamma (f_t, \lambda) \) with stochastic degree of freedom \( f_t \) and scale parameter \( \lambda \). The conditional density is:

\[
h (y_{i,t}|f_t) = \frac{1}{\Gamma(f_t)} \exp (-\lambda y_{i,t}) y_{i,t}^{f_t-1} \lambda^{f_t} 1_{y_{i,t}>0},
\]

and the micro-density becomes:

\[
\prod_{i=1}^{n} h (y_{i,t}|f_t) = \frac{1}{\Gamma(f_t)^n} \exp \left( -\lambda \sum_{i=1}^{n} y_{i,t} \right) \left( \prod_{i=1}^{n} y_{i,t} \right)^{f_t-1} \lambda^{nf_t} 1_{\min_{y_{i,t}}>0},
\]

where \( \Gamma \) denotes the gamma function. The associated conditional Laplace transform is given by:

\[
E \left[ \exp (-uy_{i,t}) | f_t \right] = \exp \left[ f_t \log(1 + u/\lambda) \right].
\]

This Laplace transform is an exponential affine function of the factor, which simplifies the derivation of nonlinear predictions.

Other specifications compatible with affine term structures can be written as:

\[
y_{it} = f_t + u_{i,t}, \quad i = 1, ..., n,
\]

where \( f_t \) and \( u_{i,t} \) are positive variables. However, the conditional distribution of \( y_{it} \) given \( f_t \) will admit a support \((f_t, \infty)\) depending on the factor value. As seen below, the approximation theorem is valid for a micro-density, which is twice-differentiable with respect to \( f_t \). Thus, it cannot be applied to this type of affine model.

iv) Homogeneous class of CDS

As mentioned above, the variables \( y_{it} \) can represent the values of digital CDS with identical time-to-maturity divided by the associated zero-coupon price. Such variables take continuous values between 0 and 1, and their distribution can be chosen in the class of beta distributions. For instance, we can choose:

\[
h (y_{i,t}|f_t) = \frac{\Gamma(f_t)}{\Gamma(\alpha f_t) \Gamma \left[ (1 - \alpha) f_t \right]} y_{i,t}^{\alpha f_t-1} (1 - y_{i,t})^{(1-\alpha) f_t-1} 1_{0<y_{i,t}<1},
\]

(2.5)
where $\alpha$ is a scalar between 0 and 1, and $f_t$ is a positive factor. The conditional mean $E(y_{it}|f_t) = \alpha$ is constant. Further, it is known that the conditional variance of a variable on $(0, 1)$ is upper bounded:

$$V(y_{it}|f_t) \leq E(y_{it}|f_t) \left[1 - E(y_{it}|f_t)\right],$$

and that the upper bound is reached when the total mass is on the two-points set $\{0, 1\}$. It is easily checked that:

$$f_t + 1 = \frac{E(y_{it}|f_t) \left[1 - E(y_{it}|f_t)\right]}{V(y_{it}|f_t)}.$$  

Thus, factor $f_t$ measures the concentration of the distribution, taking into account the existence of the upper bound. We get a model with constant conditional mean and stochastic concentration parameter. The micro-density is:

$$\prod_{i=1}^{n} h(y_{i,t}|f_t) = \left(\frac{\Gamma(f_t)}{\Gamma(\alpha f_t)\Gamma(1 - \alpha f_t)}\right)^n \left(\prod_{i=1}^{n} y_{i,t}\right)^{\alpha f_t - 1} \left(\prod_{i=1}^{n} (1 - y_{i,t})\right)^{(1-\alpha) f_t - 1} \prod_{i=1}^{n} 1_{0 < y_{i,t} < 1}.$$  

By interchanging the role of $\alpha$ and $f_t$ in density (2.5), we get an alternative model with stochastic conditional mean $f_t$ and constant concentration parameter $\alpha$.

## 3 Large portfolio approximation

This section provides the approximation theorem of the predictive factor density valid for large $n$. Then, we explain how this theorem can be used for either pricing, or prediction purposes.

### 3.1 Approximation theorem

Let us consider an homogeneous class satisfying Assumptions A.1 to A.3. The approximation theorem is derived along the lines of the Laplace method, which is well-known in statistics [e.g., Jensen (1995)]. Define:

$$\hat{f}_{nt} = \arg \max_{f_t} \sum_{i=1}^{n} \log h(y_{i,t}|f_t).$$  

(3.1)
Proposition 1. The conditional Laplace transform of $f_t$ given $y_{1,t}, \ldots, y_{n,t}, f_{t-1}$ is such that:

$$E \left[ \exp (u f_t) | y_{1,t}, \ldots, y_{n,t} \right] = \exp \left\{ u \left( \hat{f}_{nt} + \frac{1}{n} I_{nt} \frac{\partial \log g}{\partial f_t} \left( \hat{f}_{nt} | \hat{f}_{n,t-1} \right) + \frac{1}{2} \frac{1}{n I_{nt}} S_{nt} \right) + \frac{1}{2} \frac{I_{nt}^{-1}}{n} u^2 + o(1/n) \right\},$$

where:

$$I_{nt} := -\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 \log h}{\partial f_t^2} (y_{it} | \hat{f}_{nt}) \quad \text{and} \quad S_{nt} := \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^3 \log h}{\partial f_t^3} (y_{it} | \hat{f}_{nt}).$$

Proof. See Appendix 1. 

Up to order $1/n$, the lagged factor values are non-informative, and the path dependence is captured by means of the four summary statistics $\hat{f}_{nt}$, $\hat{f}_{n,t-1}$, $I_{nt}$ and $S_{nt}$. The statistic $\hat{f}_{nt}$ is the cross-sectional maximum likelihood estimator of the current factor value $f_t$. Statistic $I_{nt}$ is the estimated Fisher information corresponding to this estimation problem. Finally, $S_{nt}$ is the additional statistic involved in asymptotic bias correction. This set of statistics is independent of the factor dynamics, that is, of function $g$. The conditional density of $f_t$ given $y_{1,t}, \ldots, y_{n,t}$ is approximately normal at order $1/n$:

$$N \left( \hat{f}_{nt} + \frac{1}{n} I_{nt}^{-1} \frac{\partial \log g}{\partial f_t} \left( \hat{f}_{nt} | \hat{f}_{n,t-1} \right) + \frac{1}{2} \frac{1}{n I_{nt}} S_{nt} \right), \frac{1}{n I_{nt}} \right).$$

(3.2)

Proposition 1 can be seen as an approximation of a posterior distribution in Bayesian statistics [see e.g. Lindley (1980)]. Indeed, let us assume that $f_t$ is an unknown parameter. Then, distribution $g(\cdot | f_{t-1})$ can be interpreted as the prior density, and $E \left[ \exp (u f_t) | y_{1,t}, \ldots, y_{n,t}, f_{t-1} \right]$ characterizes the posterior distribution. Proposition 1 explains how the posterior moments of $f_t$ can be approximated by expanding around their maximum likelihood estimate. 

6These statistics can satisfy deterministic relationships in some applications. The expressions of statistics $\hat{f}_{nt}$, $I_{nt}$ and $S_{nt}$, and the associated approximations of the predictive density, for the examples of Sections 2.2 are provided in Appendix 2.

7This Bayesian interpretation considers $f_t$ as the parameter, and is valid for any parametric or nonparametric specification of the factor distribution. It has to be distinguished from Bayesian approaches concerning the mean reversion and volatility parameters of the factor distribution [see e.g. Duffie et al. (2006), Section 7].
The posterior density is asymptotically approximated by a well-defined density, i.e. the normal approximation given in (3.2). This is especially important when the approximation concerns a risk-neutral distribution; indeed, the possibility to interpret the approximation as a density is equivalent to the absence of arbitrage opportunities in the approximate pricing. It is known in the statistical literature that more accurate approximations at order $1/n^2$ can be derived [see e.g. Tierney, Kadane (1996)]. However, when these more accurate methods are used to approximate a moment $E[\varphi(f_t)|\Omega_t]$ by $E_n[\varphi(f_t)|\Omega_t]$, say, the mapping $\varphi \rightarrow E_n[\varphi(f_t)|\Omega_t]$ is not a linear operator [see e.g. formula (A.2) in Tierney and Kadane (1986)]. Thus, these approximations are not appropriate for pricing purposes.

We immediately deduce from Proposition 1 the approximated conditional expectation of any smooth function of the factor.

**Corollary 2.** For any twice continuously differentiable function $\varphi(f_t)$ of $f_t$, we have:

$$E\left[\varphi(f_t)\mid y_{1,t}, \ldots, y_{n,t}\right] = \varphi(\hat{f}_{nt}) + \frac{1}{n} \frac{d\varphi}{df_t}(\hat{f}_{nt}) \left[I_{nt}^{-1} \frac{\partial \log g}{\partial f_t}(\hat{f}_{nt} \mid \hat{f}_{nt-1}) + \frac{1}{2} I_{nt}^{-2} S_{nt}\right]$$

$$+ \frac{1}{2n} \frac{d^2\varphi}{df_t^2}(\hat{f}_{nt}) I_{nt}^{-1} + o(1/n).$$

**Proof.** The approximation is derived by expanding function $\varphi$ at second-order around $\hat{f}_{nt}$:

$$\varphi(f_t) = \varphi(\hat{f}_{nt}) + \frac{d\varphi}{df_t}(\hat{f}_{nt}) (f_t - \hat{f}_{nt}) + \frac{1}{2} \frac{d^2\varphi}{df_t^2}(\hat{f}_{nt}) (f_t - \hat{f}_{nt})^2 + o\left((f_t - \hat{f}_{nt})^2\right),$$

and computing the conditional expectation w.r.t. the density in (3.2).

Corollary 2 describes how the first- and second-order derivatives of function $\varphi$ are involved in the expansion of the conditional expectation. This is the analogue of the Ito’s formula for large homogeneous class. When function $\varphi$ is not differentiable, for instance if $\varphi(f_t) = (f_t - K)^+$, the approximation formula is no longer valid. However, we can apply Corollary 2 to the imaginary exponential transformation $\varphi(f_t) = \exp(iuf_t)$ and apply the Fourier Transform Inversion formula to get the appropriate approximation [see Duffie, Pan, Singleton (2000)]. We will carefully discuss this point in Section 4 in the application to the pricing of derivatives written on a factor proxy.

---

For $\varphi(f_t) = (e^{iu f_t} - K)^+$, the approximation of the conditional expectation can be directly computed using distribution (3.2) and the Black-Scholes formula.
Finally, Proposition 1 suggests an alternative approximation at order $1/n$ for the conditional moment of any function $\varphi(f_t)$ of the factor. Since from (2.4) the distribution of $f_t$ given the investors’ information depends on the past factor values through the most recent lag $f_{t-1}$ only, Proposition 1 implies that we can replace $f_{t-1}$ by $\hat{f}_{n,t-1}$ at order $1/n$. Then:

$$E \left[ \varphi(f_t) \middle| y_{1,t}, \ldots, y_{n,t}, f_{t-1} \right] = \int \varphi(f_t) g(f_t \mid \hat{f}_{n,t-1}) \prod_{i=1}^{n} h(y_{i,t} \mid f_t) df_t$$

$$+ o(1/n). \quad (3.3)$$

Approximation (3.3) differs from Corollary 2 since it relies only on the possibility to replace at order $1/n$ the lagged factor value by its ML estimate, while Corollary 2 also relies on the Gaussianity at order $1/n$ of the predictive density of the current factor value.

### 3.2 Approximate pricing formula

Let us consider a European derivative with time-to-maturity 1 and a payoff $a(y_{1,t+1})$, say. Its price at date $t$ is:

$$\pi_t(a, 1) = \int \cdots \int a(y_{1,t+1}) p(y_{1,t+1}, \ldots, y_{n,t+1}, f_t \mid \Omega_t) dy_{1,t+1} \cdots dy_{n,t+1} df_t$$

$$= \int m(f_t) \left( \int a(y_{1,t+1}) h(y_{1,t+1} \mid f_{t+1}) g(f_{t+1} \mid f_t) dy_{1,t+h} df_{t+1} \right) g(f_t \mid \Omega_t) df_t$$

$$= E \left[ m(f_t) \alpha(f_t) \mid \Omega_t \right], \text{ say.}$$

Thus, it is equivalent to price a derivative written on the traded assets with payoff $a(y_{1,t+1})$, or to price a derivative written on the unobserved factor value with virtual payoff $\alpha(f_t)$. This argument extends to a general payoff $a(y_{1,t+h}, \ldots, y_{n,t+h})$ and any horizon $h$.

The approximation theorem provides pricing formulas at order $1/n$ for European payoffs written on the factor. Let us first consider the short horizon and denote $\pi_t(\alpha, 0)$ the price at date $t$ of the payoff $\alpha(f_t)$. We have:

$$\pi_t(\alpha, 0) = E \left[ m(f_t) \alpha(f_t) \mid y_{1,t}, \ldots, y_{n,t}, f_{t-1} \right].$$

The approximation given in Corollary 2 with $\varphi(f) = m(f) \alpha(f)$ shows that the derivative price can be computed at order $1/n$ from the observed values $y_{1,t}, \ldots, y_{n,t}$ even if $f_t$ is

---

9This payoff is observed at the end of period $(t, t+1)$, not at the beginning of this period.
not observed at any date, whenever the sdf is given. Loosely speaking, the approximate pricing formula can be used to define a coherent system of derivative quotations, even if no derivative is highly traded. This system of quotations is fixed by solely the choice of the sdf $m$, that is, by the selected levels of risk premia.

Let us now consider horizon $h$ and denote $\pi_t(\alpha, h)$ the price at date $t$ of the payoff $\alpha(f_{t+h})$. We have:

$$\pi_t(\alpha, h) = E \left[ m(f_t)m(f_{t+1}) \cdots m(f_{t+h}) \alpha(f_{t+h}) | \Omega_t \right].$$

By an iterated expectation argument, we can first condition on $\Omega_t, f_t$ to get:

$$\pi_t(\alpha, h) = E \left[ E \left[ m(f_t)m(f_{t+1}) \cdots m(f_{t+h}) \alpha(f_{t+h}) | \Omega_t, f_t \right] | \Omega_t \right].$$

where $\Psi(f_t, h, \alpha) = E \left[ m(f_{t+1}) \cdots m(f_{t+h}) \alpha(f_{t+h}) | \Omega_t, f_t \right]$ is a function of $f_t$ only by the Markov property. Function $\Psi(f_t, h, \alpha)$ corresponds to the price at $t$ of payoff $\alpha(f_{t+h})$ with time-to-maturity $h$ computed by an informed investor who has the larger information set $\Omega_t, f_t$ available at date $t$, and uses the sdf $m(f_{t+1})$ to discount risk between $t$ and $t+1$.

From Corollary 2 with $\varphi(f; h, \alpha) = m(f) \Psi(f, h, \alpha)$, price $\pi_t(\alpha, h)$ can be approximated at order $1/n$ using the observed values $y_{1,t}, \ldots, y_{n,t}$. In particular, the term structure of derivative prices, that is the function $h \rightarrow \pi_t(\alpha, h)$, is a combination of the three baseline term structures $h \rightarrow \varphi(\hat{f}_n; h, \alpha)$, $h \rightarrow \frac{d\varphi}{df} \left( \hat{f}_n; h, \alpha \right)$, and $h \rightarrow \frac{d^2\varphi}{df^2} \left( \hat{f}_n; h, \alpha \right)$.

When some derivatives are actively traded on the market, the corresponding time-$t$ prices can be used to correct the predictive density of $f_t$ derived from Proposition 1. We follow the canonical valuation approach introduced in Stutzer (1996) for pricing stock derivatives. This approach suggests to consider the distribution of $f_t$, that is the closest to the Gaussian approximation in (3.2) in terms of the Kullback-Leibler proximity measure and satisfies the pricing restrictions implied by the observed derivative prices. Specifically, suppose that at time $t$ the option with virtual payoff $\beta(f_t)$ and time-to-maturity 1 is actively traded, with time-$t$ price $q_t$. Let us denote by $\tilde{g}_n(f_t | \Omega_t)$ the Gaussian approximation of

---

10 The extension to several observed derivative prices with generic time-to-maturity is straightforward.
the predictive density in (3.2). Then, the corrected time-
time predictive density \( \hat{g}_t \) is given by:
\[
\hat{g}_t = \arg \min_g \int \log \left[ g(f) / \tilde{g}_n(f | \Omega_t) \right] g(f) df
\]
\[
\text{s.t.} \quad \int g(f) df = 1
\]
\[
\int m(f) \beta(f) g(f) df = q_t
\]

The solution of this constrained minimization problem is:
\[
\hat{g}_t(f) = \frac{\tilde{g}_n(f | \Omega_t)e^{\lambda \beta(f)}}{\int \tilde{g}_n(f | \Omega_t)e^{\lambda \beta(f)} df},
\]
where \( \lambda \) is the associated Lagrange multiplier, that solves \( \int m(f) \beta(f) \tilde{g}_n(f | \Omega_t) e^{\lambda \beta(f)} df = q_t \). Then, the approximate time-
time price of a derivative with virtual payoff \( \alpha(f_t) \) and time-
to-maturity 1 is given by \( \int m(f) \alpha(f) \hat{g}_t(f) df \).

### 3.3 Default and migration correlation

The large portfolio approximation can also be used under the historical distribution to de-
ervive (nonlinear) predictions of the unobservable factor. This application is especially im-
portant for analyzing default or migration dependence [e.g., Vasicek (1987), Gupton et al.
Gouriéroux (2005), Feng, Gouriéroux, Jasiak (2008)]. We recall below the standard spec-
fications used for a dynamic analysis of defaults and rating migration in homogeneous
populations. They involve a common default (or migration) risk factor, that changes ran-
domly over time and is sometimes called frailty [see e.g. Duffie et al. (2006)].

i) Generalized one-factor firm value model

The one-factor firm value model is written as [see Vasicek (1987) and the Credit Metrics
framework in Gupton, Finger, Bhatia (1997)]:
\[
\log \left( \frac{A_{it}}{L_{it}} \right) = m + \sigma \sqrt{\rho} F_i + \sigma \sqrt{1 - \rho} u_{it},
\]
where the error terms are independent standard Gaussian variables, \( A_i \) and \( L_i \) denote the
asset value and liability of firm \( i \), respectively, and \( \rho \in (0, 1) \). When the variance of
common factor $F_t$ is normalized to 1, the variance of the log asset-to-liability ratio of any firm is $\sigma^2$, and the correlation between the log asset-to-liability ratios of any two firms is $\rho$. This structural model is used to characterize the joint distribution of default occurrence: $y_{it} = 1$, if $A_{it} < L_{it}$, and $= 0$, otherwise. Thus, conditional on factor value $F_t$, the dichotomous variables $y_{i,t}$, $i = 1, \ldots, n$, are i.i.d. with Bernoulli distribution $\mathcal{B}(1, f_t)$, where the transformed factor is defined by $f_t = \Phi \left( -\frac{m + \sigma \sqrt{\rho F_t}}{\sigma \sqrt{1 - \rho}} \right)$. This transformed factor $f_t$ corresponds to the conditional default probability at time $t$. The joint conditional distribution of default indicators is:

$$\prod_{i=1}^{n} h(y_{i,t}|f_t) = (f_t)^{\tilde{y}_{nt}} (1 - f_t)^{n(1-\tilde{y}_{nt})},$$ (3.4)

where $\tilde{y}_{nt}$ is the proportion of obligors that actually defaulted in period $t$. The approximation theorem can be used to derive the predictive distribution of $f_t$ at order $1/n$, that is [see Appendix 2 v]):

$$N \left( \tilde{y}_{nt} + \frac{1}{n} \left[ \tilde{y}_{nt} (1 - \tilde{y}_{nt}) \frac{\partial \log g}{\partial f_t} (\tilde{y}_{nt}|\tilde{y}_{n,t-1}) + 1 - 2\tilde{y}_{nt} \right], \frac{1}{n} \tilde{y}_{nt} (1 - \tilde{y}_{nt}) \right),$$ (3.5)

$^{11}$A better notation is $\prod_{i=1}^{n} h(y_{i,t}|f_t, y_{i,t-1} = 0)$, since the model is valid for firms which have not yet defaulted and since default is an absorbing state. More precisely, conditional on a factor path, the individual histories are i.i.d. Markov chains with time-varying default probability:

$$P[y_{i,t} = 1|y_{i,t-1} = 0, (f_t)] = f_t,$$

and default as an absorbing state:

$$P[y_{i,t} = 1|y_{i,t-1} = 1, (f_t)] = 1.$$

Let $n_{t-1} = \sum_{i=1}^{n} (1 - y_{i,t-1})$ denote the number of firms which are still operating at the beginning of year $t$. For the purpose of pricing credit derivatives with payoffs that are independent of the obligors’ names (such as $\alpha$-to-default swaps and longevity bonds), we can assume that the operating firms at time $t$ are those with indices $i = 1, \ldots, n_{t-1}$. The micro-density of these firms is:

$$\prod_{i=1}^{n} h(y_{i,t}|f_t, y_{i,t-1} = 0) = \prod_{i=1}^{n_{t-1}} h(y_{i,t}|f_t) = (f_t)^{\tilde{f}_{n,t}} (1 - f_t)^{n_{t-1}(1-\tilde{f}_{n,t})},$$

where $\tilde{f}_{n,t} = \frac{1}{n_{t-1}} \sum_{i=1}^{n_{t-1}} y_{i,t}$ is the cross-sectional ML estimator of default probability $f_t$ for year $t$. The past information on default events is summarized in the size $n_{t-1}$ of the operating pool in year $t$. To simplify the exposition, the impact of the past information has not be explicitly considered in formula (3.4). For factor approximation and derivative pricing, it is sufficient to update the pool size $n_{t-1}$ w.r.t. time $t$. 17
where $g(.|.)$ denotes the transition density of the transformed autoregressive Gaussian process $(f_t)$.

**ii) Ratings**

The one-factor firm value model can also be used to analyze jointly the ratings of several firms. Let us introduce thresholds $a_1 \leq \ldots \leq a_{K-1}$, say, and define the rating as:

$$ y_{it} = k, \text{ if } a_{k-1} < \log A_{it} - \log L_{it} < a_k, $$

with $k = 1, \ldots, K$, and $a_0 = -\infty$, $a_K = +\infty$ by convention. We have:

$$ P[y_{it} = k | f_t] = \Phi \left( \frac{a_k - m - \sigma \rho f_t}{\sigma \sqrt{1 - \rho^2}} \right) - \Phi \left( \frac{a_{k-1} - m - \sigma \rho f_t}{\sigma \sqrt{1 - \rho^2}} \right), $$

where the factor is $f_t = F_t$. The micro-density is:

$$ \prod_{i=1}^n h(y_{i,t} | f_t) = \prod_{k=1}^K \left[ \Phi \left( \frac{a_k - m - \sigma \rho f_t}{\sigma \sqrt{1 - \rho^2}} \right) - \Phi \left( \frac{a_{k-1} - m - \sigma \rho f_t}{\sigma \sqrt{1 - \rho^2}} \right) \right] \bar{y}_{k,n,t}, $$

where $\bar{y}_{k,n,t} = \frac{1}{n} \sum_{i=1}^n 1 \{ y_{it} = k \}$ is the observed frequency of rating $k$. This approach extends the Generalized one-factor firm value model to any number of alternatives. When the factor is integrated out, the model captures the migration correlations due to the common factor (for instance, to the business cycle). Proposition 1 can be used to derive an approximation for the predictive distribution of $f_t$ at order $1/n$.

### 4 Derivatives written on a factor proxy

As noted earlier, the common factor is often introduced to capture the nondiversifiable component of individual risks, also called systematic risk. It is not surprising to see market solutions to reduce the nondiversifiable risk exposure. Since the underlying factor is not directly observable, it is not possible to propose derivatives written on the factor itself. However, such derivatives can be replaced by derivatives written on an observable factor proxy. A typical example is provided by basket default swaps. These are digital derivatives that pay off at maturity, if the frequency of defaults in a pool of obligors is larger than a given threshold. The frequency of default is a proxy for a systematic credit risk factor.
The goal of this section is to introduce an approximation formula for the price of derivatives written on a Maximum Likelihood (ML) factor proxy. We first consider in Section 4.1 the case where the factor proxy corresponds to the default frequency. We discuss in details several examples of such derivative products proposed on the market. Then, in Section 4.2 we compare the prices of derivatives written on the factor proxy (default frequency), with the prices of the corresponding derivatives written on the unobservable factor value (default probability). Finally, in Section 4.3 the results are extended to the general framework of a homogeneous class of assets and derivatives written on a factor proxy that corresponds to the ML estimate of the factor value.

4.1 Approximate pricing of derivatives written on a default frequency

i) Short-term $\alpha$-to-default swap

Let us consider a large pool of firms of similar size in a given industrial sector. Let $y_{it}$ denote the indicator for default occurrence, that is $y_{it} = 1$, if firm $i$ defaults at date $t$, and $= 0$, otherwise. Suppose that the joint distribution of the default indicators is given by the generalized one-factor firm value model of Section 3.3, in which $y_{1,t}, \ldots, y_{n,t}$ are i.i.d. with the same Bernoulli distribution $\mathcal{B}(1, f_t)$, conditional on factor $f_t$. A $\alpha$-to-default swap with maturity $t+1$ pays one Euro, if the fraction of firms in the pool which are in default at $t+1$ is above $100\alpha$ percent, $\alpha \in [0, 1]$, and 0, otherwise. The payoff of this derivative is given by:

$$a(y_{1,t+1}, \ldots, y_{n,t+1}) = I_{\bar{y}_{n,t+1} \geq \alpha},$$

where $\bar{y}_{n,t+1} = \frac{1}{n} \sum_{i=1}^{n} y_{i,t+1}$.

The basket default swap can be seen as a derivative written on a factor proxy. Indeed, let $n_t = \sum_{i=1}^{n} (1 - y_{i,t})$ denote the number of firms which are still operating at the end of year $t$. Without loss of generality, we can assume that these firms correspond to indices $i = 1, \ldots, n_t$. Then, we have $\bar{y}_{n,t+1} = \frac{n_t}{n} \hat{f}_{n,t+1}$, where $\hat{f}_{n,t+1} = \frac{1}{n_t} \sum_{i=1}^{n_t} y_{i,t+1}$ is the ML estimator of default probability $f_{t+1}$ at date $t + 1$ in the pool of $n_t$ firms. Estimator $\hat{f}_{n,t+1}$ corresponds to the future default frequency. Thus, the payoff of the $\alpha$-to-default swap can
be written as:
\[
a (y_{1,t+1}, \ldots, y_{n,t+1}) = \mathbb{I}_{f_{n,t+1} \geq \alpha_t},
\]
where \( \alpha_t = \frac{n}{n_t} \alpha \) is known at date \( t \).

For expository purpose, let us assume that the sdf is constant, with zero risk-free rate. The price at time \( t \) of the \( \alpha \)-to-default swap for maturity \( t+1 \) is given by:
\[
p_{n,t}(\alpha, 1) = E \left[ \mathbb{I}_{f_{n,t+1} \geq \alpha_t} \bigg| \Omega_t \right]. \tag{4.1}
\]
This price can be computed from the prices of exponential derivatives written on \( \hat{f}_{n,t+1} \) by using the Fourier Transform Inversion formula. More precisely, the price at date \( t \) of the derivative with exponential payoff \( \exp(\hat{f}_{n,t+1}) \) at \( t+1 \) is given by:
\[
\pi_{n,t}(u, 1) = E \left[ \exp \left( u \hat{f}_{n,t+1} \right) \bigg| \Omega_t \right] = E_t \left[ E_t \left[ \exp \left( u \hat{f}_{n,t+1} \right) \bigg| f_{t+1} \right] \right]
= E_t \left( \left\{ 1 + \left( \exp \left( \frac{u}{n_t} \right) - 1 \right) f_{t+1} \right\}^{n_t} \right), \tag{4.2}
\]
where \( E_t \) denotes the conditional expectation given the investors’ information \( \Omega_t = (y_{i,t}, i = 1, \ldots, n, f_{t-1}) \). From the Law of Iterated Expectation, we get:
\[
\pi_{n,t}(u, 1) = E \left[ \varphi(f_t; u) \bigg| \Omega_t \right], \tag{4.3}
\]
where \( \varphi(f_t; u) = E \left[ \left\{ 1 + \left( \exp \left( \frac{u}{n_t} \right) - 1 \right) f_{t+1} \right\}^{n_t} \bigg| f_t \right] \). The density of \( f_t \) given \( \Omega_t \) is computed by formula (2.4) applied to the pool of \( n_{t-1} \) firms operating in period \( (t-1, t) \):
\[
g \left( f_t \bigg| \Omega_t \right) = \frac{g(f_t|f_{t-1}) \left( f_t \right)^{n_{t-1}-n_t} (1-f_t)^{n_t}}{\int g(f_t|f_{t-1}) \left( f_t \right)^{n_{t-1}-n_t} (1-f_t)^{n_t} df_t}. \tag{4.4}
\]
Thus, the derivative price \( \pi_{n,t}(u, 1) \) depends on the investor information \( \Omega_t \) through \( f_{t-1}, n_t \) and \( n_{t-1} \). Then, from the Fourier Transform Inversion formula in Proposition 2 of Duffie, Pan, Singleton (2000), which applies to purely imaginary argument \( u = iv \), say, we get: \(^{12}\)
\[
p_{n,t}(\alpha, 1) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{Im \left[ \pi_{n,t}(iv, 1) \exp(-iv\alpha_t) \right]}{v} dv, \tag{4.5}
\]
\(^{12}\)The Fourier Transform Inversion formula is presented in Duffie, Pan, Singleton (2000) to compute expectations of indicator and truncated functions of continuous-time affine processes, but is valid for the indicator function of a general random variable.
where $Im$ denotes the imaginary component of a complex number.

The approximation of the $\alpha$-to-default swap price is derived by the approximation of the exponential derivative prices. By equation (4.3) and Corollary 2 with $n_{t-1} \to \infty$, we get [see Appendix 4 ii)]:

$$
\pi_{n,t}(u,1) = \varphi(\hat{f}_{nt}; u) + \frac{1}{n_{t-1}} \frac{d \varphi}{df_t}(\hat{f}_{nt}; u) \left[ \hat{f}_{nt} \left( 1 - \hat{f}_{nt} \right) \frac{\partial \log g}{\partial \hat{f}_{nt}} \left( \hat{f}_{nt} \hat{f}_{n,t-1} \right) - 2 \hat{f}_{nt} + 1 \right] 
+ \frac{1}{2n_{t-1}} \hat{f}_{nt} \left( 1 - \hat{f}_{nt} \right) \frac{d^2 \varphi}{df_t^2}(\hat{f}_{nt}; u) + o(1/n_{t-1}). (4.6)
$$

Since $\hat{f}_{nt} = (n_{t-1} - n_t)/n_{t-1}$, the approximate price depends on the past default history through the counts $n_{t-2}, n_{t-1}$ and $n_t$ of operating firms at the end of year $t-2$, $t-1$ and $t$, respectively. An equivalent summary of past default history is $n_{t-1}, n_t, \hat{f}_{n,t-1}$. Then, the approximation of the $\alpha$-to-default swap price is obtained by the Fourier Transform Inversion formula (4.5) applied with approximation (4.6) evaluated at purely imaginary argument.

ii) $\alpha$-to-default swap at longer horizon

Let us now consider a $\alpha$-to-default swap with a general time-to-maturity $h$. The payoff at $t + h$ is given by:

$$
a(y_{i,t+h}, \ldots, y_{n,t+h}) = 1 \{ \bar{y}_{n,t+h} \geq \alpha \}.
$$

The approximation of the derivative price can be obtained along the same lines as for the short horizon $h = 1$. Indeed, conditional on a factor path, the default probability at horizon $h$ for a firm which is operating at $t$ is:

$$
P[y_{i,t+h} = 1|y_{i,t} = 0, (f_t)] = 1 - (1 - f_{t+1}) \cdots (1 - f_{t+h}).
$$

Thus, conditional on a factor path, the default indicators $y_{i,t+h}, i = 1, \ldots, n_t$, of the firms operating at $t$ are i.i.d. with Bernoulli distribution $B(1, \lambda_{t+h})$, where $\lambda_{t+h} = 1 - (1 - f_{t+1}) \cdots (1 - f_{t+h})$. The same approach as above can be applied, with $\varphi(f_t; u) = E \left[ \left\{ 1 + \left( \exp \left( \frac{u}{nt} \right) - 1 \right) \lambda_{t+h} \right\}^{nt} \bigg| f_t \right]$. 

iii) Other examples of derivatives written on a default frequency

13 Another one is $n_{t}, \hat{f}_{nt}, \hat{f}_{n,t-1}$. 

21
The approximate pricing approach introduced for $\alpha$-to-default swaps can be extended to other derivatives written on a default frequency. For instance, a synthetic CDO tranche is a derivative that offers protection against the losses in an underlying portfolio that are in a specific range. Assuming for expository purpose a portfolio with homogeneous nominals and zero recovery rates, the percentage portfolio loss at horizon $h$ is $\hat{f}_{n,t+h}$, and the payoff is $\left(\hat{f}_{n,t+h} - \alpha_1\right)^+ - \left(\hat{f}_{n,t+h} - \alpha_2\right)^+$, where $\alpha_1 < \alpha_2$ are called attachment, resp. detachment, points. An actively traded contract of this type is written for instance on the iTraxx Europe index, which is an index based on the CDS spreads of 125 companies. Finally, derivatives written on an observable factor proxy have been proposed also in the insurance industry. A typical example is provided by the so-called longevity bond. Such a bond has a contractual maturity (e.g. 25-year) and pays regularly a coupon proportional to the current frequency of surviving people in a contractual cohort, for instance a national population with given age at the date of bond issuing. These observed frequencies are proxies of the successive values of a longevity factor and correspond to the cross-sectional maximum likelihood estimates of these factor values. Approximate pricing formulas for longevity bonds can be derived by considering for instance a one-factor model similar to the one in Sections 3.3 and 4.1 i), ii), where the common factor is assumed nonstationary to capture a trend in the mortality risk.

4.2 Derivatives written on default frequencies vs derivatives written on default probabilities

In this section, we focus on the comparison between the price of a derivative written on a frequency that proxies the latent factor, and the price of the corresponding derivative written on the value of the factor itself. As in the previous section, let $y_{i,t}$ denote the default indicator for firm $i$ and date $t$ and assume that $y_{1,t}, \ldots, y_{n,t}$ are i.i.d. with the same Bernoulli distribution $B(1, f_t)$, conditional on factor $f_t$. For expository purpose, let us again assume a constant stochastic discount factor and a zero risk-free rate. Furthermore, let us focus on a short-term derivative with exponential payoff written on the future default frequency
\[ \hat{f}_{n,t+1} = \frac{1}{n} \sum_{i=1}^{n} y_{i,t+1}. \] From (4.2), the price at date \( t \) of the derivative with payoff \( \exp\left(u\hat{f}_{n,t+1}\right) \) is equal to:

\[
\pi_{n,t}(u, 1) = E_t \left[ \left\{ 1 + \left( \exp \left( \frac{u}{n} \right) - 1 \right) f_{t+1} \right\}^n \right] = E_t \left[ \exp \left\{ n \log \left[ 1 + \left( \exp \left( \frac{u}{n} \right) - 1 \right) f_{t+1} \right] \right\} \right] + o(1/n) = E_t \left[ \exp \left\{ n \left[ \left( \frac{u}{n} + \frac{u^2}{2n^2} \right) f_{t+1} - \frac{u^2}{2n^2} f_{t+1}^2 \right] \right\} \right] + o(1/n) = E_t \left[ \exp \left( uf_{t+1} + \frac{u^2}{2n} f_{t+1} (1 - f_{t+1}) \right) \right] + o(1/n). \tag{4.7} \]

This price differs from the price of the similar derivative written on the factor itself, which is equal to:

\[ \pi_{\infty,t}(u, 1) = E_t \left[ \exp \left( uf_{t+1} \right) \right]. \]

Up to order \( 1/n \), the ratio between the two derivative prices is equal to:

\[
\frac{\pi_{n,t}(u, 1)}{\pi_{\infty,t}(u, 1)} = \frac{E_t \left[ \exp \left( uf_{t+1} + \frac{u^2}{2n} f_{t+1} (1 - f_{t+1}) \right) \right]}{E_t \left[ \exp \left( uf_{t+1} \right) \right]} = E_t^P \left[ \exp \left( \frac{u^2}{2n} f_{t+1} (1 - f_{t+1}) \right) \right],
\]

where \( P_u \) is a modified probability with density \( \exp \left( uf_{t+1} \right) / E_t \left[ \exp \left( uf_{t+1} \right) \right] \). In particular, the price \( \pi_{n,t}(u, 1) \) is always larger than \( \pi_{\infty,t}(u, 1) \) for large \( n \), and decreases with the size \( n \) of the underlying cohort. In fact, this ratio represents the price of the aggregate idiosyncratic risk, which does not vanish, when the cohort has a finite size. This is seen by noting that \( V_t \left[ \hat{f}_{n,t+1} | f_{t+1} \right] = \frac{1}{n} f_{t+1} (1 - f_{t+1}) \). The correction term \( \exp \left( \frac{u^2}{2n} f_{t+1} (1 - f_{t+1}) \right) \) is the granularity adjustment for exponential derivative prices [see e.g. Gordy (2003), (2004) and references therein for granularity adjustements of the Value-at-Risque and expected shortfall of the portfolio loss distribution in a static framework].

\textsuperscript{14} For expository purpose, we note \( n_t = n \).
4.3 Approximate pricing of derivatives written on a factor proxy

Let us now consider the general framework of an homogenous class of assets satisfying Assumptions A.1-A.5 and derivatives written on the factor proxy
\[ \hat{f}_{n,t+h} = \arg \max_{f_{t+h}} \sum_{i=1}^{n} \log h \left( y_{i,t+h} | f_{t+h} \right), \]
that is, the ML estimator of the factor value \( f_{t+h} \). We focus on approximate pricing of derivatives with exponential payoff \( \exp\left( u \hat{f}_{n,t+h} \right) \). These are the basis for approximate pricing of more general payoffs by means of the Fourier Transform Inversion formula (see Section 4.1).

**Proposition 3.** The price \( \pi_{n,t}(u, h) \) at time \( t \) of the derivative with payoff \( \exp\left( u \hat{f}_{n,t+h} \right) \) at \( t + h \) is such that:
\[
\pi_{n,t}(u, h) = E \left[ \varphi \left( f_t \right) | \Omega_t \right] + o(1/n), \tag{4.8}
\]
where:
\[
\varphi \left( f_t \right) = E \left[ m(f_t)m(f_{t+1})...m(f_{t+h-1}) \exp \left( uf_{t+h} + \frac{u^2}{2n} I_{t+h} \right) | f_t \right], \tag{4.9}
\]
and \( I_{t+h} = E \left[ -\frac{\partial^2 \log h \left( y_{i,t+h} | f_{t+h} \right) }{\partial f^2} | f_{t+h} \right] \).

**Proof.** See Appendix 3. \( \square \)

The quantity \( \frac{1}{n} I_{t+h}^{-1} \) is the asymptotic variance of the ML estimator \( \hat{f}_{n,t+h} \) of \( f_{t+h} \) at date \( t + h \), that is, the aggregate idiosyncratic risk for finite class size \( n \). This explains the difference at order \( 1/n \) between \( \pi_{n,t}(u, h) \) and the price:
\[
\pi_{\infty,t}(u, h) = E \left[ m(f_t)m(f_{t+1})...m(f_{t+h-1}) \exp \left( uf_{t+h} \right) | \Omega_t \right],
\]
of the derivative written on the factor value \( f_{t+h} \). The result in Proposition 3 generalizes the formula (4.7) valid for derivatives written on frequencies, for which \( I_{t+h}^{-1} = f_{t+h} (1 - f_{t+h}) \).

The approximation at order \( o(1/n) \) of derivative price \( \pi_{n,t}(u, h) \) is deduced by Corollary 2:
\[
\pi_{n,t}(u, h) = \varphi \left( \hat{f}_n \right) + \frac{1}{n} \frac{d \varphi}{d f} \left( \hat{f}_n \right) \left[ I_n^{-1} \frac{\partial \log g}{\partial f_t} \left( \hat{f}_n | \hat{f}_n,t-1 \right) + \frac{1}{2} I_n^{-2} S_n \right] + \frac{1}{2n} \frac{d^2 \varphi}{d f^2} \left( \hat{f}_n \right) I_n^{-1} + o(1/n),
\]
with \( \varphi \left( f_t \right) \) given in (4.9), and \( I_n \) and \( S_n \) as in Proposition 1.
4.4 Comparison with the literature on large portfolio approximations

There exist two streams of literature on large portfolio approximations developed for static factor models. The first stream is interested in the prediction of future aggregate risk, assumes unobservable the underlying factor, and thus proposes large $n$ approximations of the marginal historical distribution of $\hat{f}_{n,t+h}$ [see e.g. Vasicek (1991), Lucas et al. (2001), Schoenbucher (2002), Schloegl, O’Kane (2005), Dembo et al. (2004)]. These approximations consist in replacing the distribution of $\hat{f}_{n,t+h}$ by that of $f_{t+h}$, on the basis of the Law of Large Numbers applied conditional on the factor value. The second stream of literature is interested in pricing, assumes observable the underlying factor and proposes large $n$ approximations of the risk-neutral conditional distribution of $\hat{f}_{n,t+h}$ given $f_{t+h}$. These approximations are generally based on Stein’s method [see e.g. El Karoui, Jiao, Kurtz (2007), Bastide, Benhamou, Ciua (2007)]. More precisely, the factor proxy $\hat{f}_{n,t+h}$ is approximately equal to:

$$\hat{f}_{n,t+h} \simeq f_{t+h} + \frac{1}{\sqrt{n}}E\left[ - \frac{\partial^2 \log h (y_{i,t+h}|f_{t+h})}{\partial f^2} \right]^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial \log h}{\partial f} (y_{i,t+h}|f_{t+h}) ,$$

(see Appendix 3). The second component of the right hand side involves a sum of terms, which are i.i.d. conditional on the factor value $f_{t+h}$. The standard Stein’s method could be used to get a Gaussian approximation with bias correction term of the distribution of $\hat{f}_{n,t+h}$ given $f_{t+h}$. Thus, it is appropriate to price derivatives written on $\hat{f}_{n,t+h}$ when $f_{t+h}$ is observable.

In our paper, we consider a dynamic framework, assume unobservable the underlying factor and propose large $n$ approximations for both the conditional historical and risk-neutral distributions of $\hat{f}_{n,t+h}$ given $\Omega_t$. As seen in Sections 4.1 and 4.3, this requires an appropriate replacement of the distributional effects of the unobserved lagged factor values $f_{t-1}$ by means of the observed basic risks $y_{1,t}, \ldots, y_{nt}$.

5 Numerical illustration to basket default swap

In this section we provide a numerical illustration to approximate pricing of $\alpha$-to-default swap derivatives with a one-factor generalized firm value model (see Sections 3.3 and 4.1).
5.1 Parameter values of the one-factor firm value model

The time period corresponds to one year. The asset volatility is $\sigma = 0.20$, and parameter $m$ is set equal to 0.3501 in order to match an historical (unconditional) default probability of 4%. The asset correlation parameter $\rho$ can be calibrated in order to match relevant values of default correlation. In the empirical literature, different orders of magnitude for estimated default correlations have been proposed, according to the country, firm size and characteristics used to group the firms in homogeneous classes. For instance, values of about 1% have been found when large US firms are grouped into industrial sectors [De Servigny, Renault (2002)], while the estimated default correlations are about 0.1% for small and medium size French firms classified according to both industrial sector and rating [Gagliardini, Gouriéroux (2005)]. These values of default correlation imply an asset correlation $\rho$ below 0.10, much smaller than the value of about 0.30 given by the formula proposed by the Basle Committee for a default probability of 4%. To cover the range obtained with these different approaches, we consider three values of asset correlation, that are $\rho = 0.01$, $\rho = 0.10$ and $\rho = 0.30$. Finally, the common factor $F_t$ is a Gaussian autoregressive process:

$$F_t = \beta F_{t-1} + \sqrt{1-\beta^2} \eta_t,$$

where the autocorrelation coefficient is $\beta = 0.5$ and the innovations $\eta_t$ are i.i.d. standard Gaussian variables. For expository purpose, the sdf is assumed constant with zero risk-free rate.

---

15The unconditional default probability is equal to $\Phi (-m/\sigma)$.  
16The unconditional default correlation between any two firms $i$ and $j$ is given by [see e.g. Gouriéroux, Tiomo (2007), Chapter 7]:

$$\text{corr} (y_{i,t}, y_{j,t}) = \frac{\Phi (-m/\sigma, -m/\sigma; \rho) - \Phi (-m/\sigma)^2}{\Phi (-m/\sigma) [1 - \Phi (-m/\sigma)]},$$

where $\Phi (., .; \rho)$ denotes the joint cdf of the bivariate standard Gaussian distribution with correlation coefficient $\rho$. 


5.2 Patterns of factor distributions and derivative prices

Let us first illustrate the patterns of the factor distribution and \( \alpha \)-to-default swap price for a specific past default history. The numbers of operating firms at the end of years \( t - 1 \) and \( t \) are \( n_{t-1} = 1000 \) and \( n_t = 960 \), respectively, which imply \( \hat{f}_{nt} = 0.04 \). Further, we consider three different values of \( n_{t-2} \), which correspond to \( \hat{f}_{n,t-1} = 0.04 \), \( \hat{f}_{n,t-1} = 0.0025 \), and \( \hat{f}_{n,t-1} = 0.125 \), respectively.

Figure 1 displays the conditional distribution of factor \( f_t \) given \( f_{t-1} \), for different values of \( f_{t-1} \) and asset correlation \( \rho = 0.10 \). The conditioning values of \( f_{t-1} \) are given in terms of their corresponding Gaussian factor \( F_{t-1} \); they are \( F_{t-1} = 0 \), \( F_{t-1} = 2 \), and \( F_{t-1} = -2 \), respectively. Figure 2 displays the conditional distribution of factor \( f_t \) given the investors’ information \( f_{t-1}, n_{t-1}, \) and \( n_t \), for different values of \( f_{t-1} \) and asset correlation \( \rho = 0.10 \). The conditioning values of the lagged factor are \( F_{t-1} = 0 \), \( F_{t-1} = 2 \) and \( F_{t-1} = -2 \), respectively. The density is obtained by formula (4.4), where the integral in the denominator is computed numerically. \(^{17}\) Figure 3 displays the approximate distribution of factor \( f_t \) given the past default history \( n_t, n_{t-1}, \) and \( \hat{f}_{n,t-1} \), for different values of \( \hat{f}_{n,t-1} \). Asset correlation is \( \rho = 0.10 \). The approximation is obtained by formula (3.5). Whereas the sole knowledge of the lagged factor value results in skewed distributions with very different patterns (Figure 1), we see that the observed default frequencies are very informative. When this additional information is introduced, the predictive distributions given the investors’ information are close to Gaussian distributions, and much less sensitive to the lagged factor value (see Figure 2). Similarly, the approximate predictive distributions given the default history are not very sensitive to the estimated lagged factor value (see Figure 3), and close to the predictive distributions given the investors’ information displayed in Figure 2. These findings are consistent with Proposition 1.

The price of \( \alpha \)-to-default swap at time-to-maturity 1 year is displayed in Figure 4 as a function of \( \alpha \), for three different values of asset correlation \( \rho = 0.01, \rho = 0.10 \) and \( \rho = 0.30 \). The initial size of the pool is \( n = 1000 \), and the current size is \( n_t = 960 \). The

\(^{17}\) Another possibility is to simulate a long sample from the distribution of \( f_t \) given \( f_{t-1}, n_{t-1}, \) and \( n_t \) by the acceptance-rejection algorithm [see Appendix 4 i)], and apply a kernel method to estimate the density. We have checked that this approach provides similar results.
lagged Gaussian factor value is $F_{t-1} = 0$. The price is computed with the Fourier Transform Inversion formula (4.5), where the exponential derivative prices are obtained from (4.2) by Monte-Carlo simulation with the acceptance-rejection algorithm [see e.g. Robert, Casella (2004) and Appendix 4 i) for a discussion of this algorithm]. The pattern of the price as a function of $\alpha$ corresponds to the (risk-neutral) survivor function of the future factor proxy $\hat{f}_{n,t+1}$ [see equation (4.1)]. For small values of $\rho$, this pattern closely corresponds to that of a Gaussian survivor function because of the Central Limit Theorem. As expected, default correlation is a key parameter to measure the quality of a basket derivative. The price of the $\alpha$-to-default swap decreases in $\rho$ (resp. increases in $\rho$) for small (resp. large) values of $\alpha$. This is due to the positive effect of $\rho$ on the variance of the (risk-neutral) distribution of $\hat{f}_{n,t+1}$. Figure 5 displays the price of the $\alpha$-to-default swap at time-to-maturity 1 year as a function of $\alpha$, for three different values of the lagged factor $F_{t-1} = 0$, $F_{t-1} = -2$, and $F_{t-1} = 2$. Asset correlation is $\rho = 0.10$. Since the observed current and past default frequencies are included in the investor’s information set, the derivative price is not very sensitive to the lagged factor value.

Figure 6 displays the approximate price of the $\alpha$-to-default swap at time-to-maturity 1-year as a function of $\alpha$, for three different values of asset correlation $\rho = 0.01$, $\rho = 0.10$ and $\rho = 0.30$. The past default history is such that $\hat{f}_{n,t-1} = 0.04$. The approximation is obtained by the Fourier Transform Inversion formula (4.5) and the approximation (4.6) for exponential derivatives [see also Appendix 4 ii) for the implementation]. The approximate prices with $\hat{f}_{n,t-1} = 0.04$ are close to the theoretical prices with $F_{t-1} = 0$ (see Figure 4). Finally, Figure 7 displays the approximate price of the $\alpha$-to-default swap at time-to-maturity 1 year as a function of $\alpha$, for asset correlation $\rho = 0.10$ and three different values of $\hat{f}_{n,t-1}$. The approximate price is not very sensitive to the estimate of the lagged factor value.

18Strictly speaking, the CLT cannot be applied to $\hat{f}_{n,t+1} = \frac{1}{m} \sum_{i=1}^{n} y_{i,t+1}$, since the common factor introduces an equicorrelation structure across the individuals. However, for $\rho = 0.01$ the equicorrelation is weak and the Gaussian approximation implied by the CLT is rather accurate.
5.3 Monte-Carlo

Let us now investigate the size of the pricing errors implied by the approximation formula. Let us consider a $\alpha$-to-default contract issued at date $t = 0$ on a pool of $n = n_0$ obligors and maturing at date $t = 3$. We consider two different values for the initial size of the pool, which are $n = 100$ and $n = 1000$. The price is computed at date $t = 2$, for default frequencies $\alpha = 2.5\%$, $\alpha = 5\%$, $\alpha = 10\%$ and $\alpha = 12.5\%$. To compare the theoretical and approximate prices, we perform a Monte-Carlo experiment to simulate the factor path $f_t$, $t = 1, 2$, and the default history summarized by $n_t$, $t = 1, 2$. The price approximation is not valid when $\hat{f}_{n,t}$ is equal to either 0 or 1, since then the approximate factor density is degenerate with zero variance. Thus, we disregard the realizations with either $\hat{f}_{n,t} = 0$ or $\hat{f}_{n,t} = 1$, which amounts to simulate the process conditional on the event $\hat{f}_{n,t} \in (0, 1)$. For each Monte-Carlo replication, we compute i) the theoretical prices at $t = 2$ using $n_t, n_{t-1}, f_{t-1}$, and ii) the approximate prices using $n_t, n_{t-1}, \hat{f}_{n,t-1}$. We perform 2500 replications.

Table 1 reports the mean, the median, the upper and lower 5% quantiles of the distribution of the percentage pricing errors, as well as the mean and median of the absolute value of the percentage pricing errors. The three panels refer to asset correlation $\rho = 0.01$, $\rho = 0.10$ and $\rho = 0.30$, respectively. The percentage pricing error is defined as the difference between approximated price and theoretical price, in percentage of the theoretical price. The approximate prices are obtained by the method discussed in Section 5.2. The bias is overall rather small, and increases for small size of the class and large asset correlation (see the Panel with $\rho = 0.30$ and $n = 100$). For default frequencies $\alpha = 2.5\%$, $\alpha = 5\%$, asset correlations $\rho = 0.10$, $\rho = 0.30$ and class size $n = 1000$, the mean absolute value of the percentage pricing errors is below 5%. Percentage pricing errors are less than 10% in absolute value with probability at least 0.90. At the contrary, the mean absolute value of the percentage pricing errors is larger than 10% for either frequencies $\alpha = 10\%$, $\alpha = 12.5\%$, or for class size $n = 100$, or for asset correlation $\rho = 0.01$. To explain these findings, note that frequencies $\alpha = 10\%$ and $\alpha = 12.5\%$ correspond to rather extreme joint

\footnote{Here the bias reflects the strong negative skewness of the distribution of the percentage pricing errors.}
default events compared to the historical default probability of 4%. The associated theoretical derivative prices are often very close to zero. In practice, since prices are displayed in discrete ticks, these small values correspond to zero prices, yielding infinite percentage approximation errors. To avoid this problem, the statistics in Table 1 are computed using only the Monte-Carlo values with theoretical prices larger than .005. Still, for theoretical prices close to this lower bound, the percentage approximation errors are quite large and explain the large values of the quantiles displayed in Table 1 for frequencies \(\alpha = 10\%, 12.5\%\). For asset correlation \(\rho = 0.01\), only the statistics for frequencies \(\alpha = 2.5\%, \alpha = 5\%\) and pool size \(n = 1000\) are displayed. Indeed, for \(\alpha = 10\%\) and \(\alpha = 12.5\%\) the derivative prices are always below .005 (see Figure 4). Moreover for \(\rho = 0.01\) the percentage approximation errors are rather large, and often above 100\% for pool size \(n = 100\) (not displayed). This is because for \(\rho = 0\) the factor \(f_t\) does not impact the individual default indicators, and then the approximation theorem breaks down.

Figure 8 investigates the relationship between percentage pricing errors and theoretical prices, and between percentage pricing errors and estimated factor \(\hat{f}_{n,t}\), by displaying scatter plots (upper Panels) and estimated conditional quantile curves at 5\%, 50\% and 95\% level (lower Panels), for asset correlation \(\rho = 0.10\), initial size of the pool \(n = 1000\) and frequency \(\alpha = 2.5\%\). In the scatter plots, each circle corresponds to a pair of theoretical and approximate prices for a Monte-Carlo replication. The conditional quantile curves are computed by a kernel method, and are displayed for the central part of the distribution. The same statistics are displayed in Figure 9 for \(\rho = 0.10\), \(n = 1000\) and \(\alpha = 12.5\%\). Percentage pricing errors are large for Monte-Carlo repetitions with small theoretical prices, and with small estimated factor value \(\hat{f}_{n,t}\). Indeed, when \(\hat{f}_{n,t}\) is small we expect that the predictive density of \(f_t\) (see Figure 2) is squeezed at 0 and rather different from a Gaussian distribution.

Table 2 reports the statistics of the percentage pricing errors for the approximation method based on formula (3.3). Specifically, exponential derivative prices are approximated using (3.3) by computing the integral with the acceptance-rejection algorithm [see Appendix 4 i)]. The Fourier Transform Inversion formula is used to get the approximated \(\alpha\)-to-default derivative prices. Comparing Tables 1 and 2 we get a decomposition of the
pricing error into i) the component due the approximation of the lagged factor value by its ML estimate \( \hat{f}_{n,t-1} \) (Table 2), and ii) the component due to the Gaussian approximation of the predictive density (the difference between Table 1 and Table 2). The latter component is rather small for default frequencies \( \alpha = 2.5\%, \alpha = 5\% \), asset correlation \( \rho = 0.10, \rho = 0.30 \) and class size \( n = 1000 \), and the performance of the two approximation methods are close in this setting. The pricing errors induced by the Gaussian approximation get rather large for either frequencies \( \alpha = 2.5\%, \alpha = 5\% \), or class size \( n = 100 \), or asset correlation \( \rho = 0.01 \), which explains the better performance of the approximation method in Table 2 in these cases.

6 Concluding remarks

A large variety of basket derivatives with nonlinear payoffs are currently introduced to capture the nonlinear dynamic feature of common latent risk factors influencing an homogeneous pool of individual contracts. The associated securitization usually concerns either individual loans, or insurance contracts. In our paper, we have derived approximated pricing formulas for derivatives written on an observable factor proxy for large underlying pools. These pricing formulas do not involve the unobservable factor values, which are replaced by well-chosen summaries constructed from the observable asset returns.

The methodology of this paper can be easily extended to more complicated dynamic models. First, to models featuring an idiosyncratic dynamics, that is, an additional effect of lagged \( y_{i,t-1} \) on \( y_{i,t} \), whenever these idiosyncratic effects appear in the joint likelihood by means of appropriate summary statistics [see Gagliardini, Gourieroux (2007) for the extension of the approximation formula]. Second, to models including both observable and unobservable common factors [see Duffie et al. (2006) for an example of such a model].
References


Appendix 1: Proof of Proposition 1

i) Let us first derive an approximation for the conditional Laplace transform of \( f_t \) given \( y_{1,t}, \ldots, y_{n,t} \) and \( f_{t-1} \):

\[
\Psi_{nt}(u) = E\left[ \exp \left( uf_t \right) \mid y_{1,t}, \ldots, y_{n,t}, f_{t-1} \right] = \frac{\int e^{uf_t} g(f_t \mid f_{t-1}) \prod_{i=1}^{n} h(y_{i,t} \mid f_t) \, df_t}{\int g(f_t \mid f_{t-1}) \prod_{i=1}^{n} h(y_{i,t} \mid f_t) \, df_t},
\]

(A.1)

which depends only on \( y_{1,t}, \ldots, y_{n,t} \) and \( f_{t-1} \).

Let us expand the micro-density around \( \hat{f}_{nt} \):

\[
\sum_{i=1}^{n} \log h(y_{i,t} \mid f_t) = \sum_{i=1}^{n} \log h(y_{i,t} \mid \hat{f}_{nt}) + \frac{1}{2n} \sum_{i=1}^{n} \frac{\partial^2 \log h}{\partial f_t^2} (y_{i,t} \mid \hat{f}_{nt}) \left[ \sqrt{n} \left( f_t - \hat{f}_{nt} \right) \right]^2
\]

\[+ \frac{1}{6\sqrt{n}} \sum_{i=1}^{n} \frac{\partial^3 \log h}{\partial f_t^3} (y_{i,t} \mid \hat{f}_{nt}) \left[ \sqrt{n} \left( f_t - \hat{f}_{nt} \right) \right]^3
\]

\[+ \frac{1}{24n} \sum_{i=1}^{n} \frac{\partial^4 \log h}{\partial f_t^4} (y_{i,t} \mid \hat{f}_{nt}) \left[ \sqrt{n} \left( f_t - \hat{f}_{nt} \right) \right]^4 + o(1/n).
\]

Let us introduce the change of variable:

\[
X = I_{nt}^{1/2} \sqrt{n} \left( f_t - \hat{f}_{nt} \right) \quad \iff \quad f_t = \hat{f}_{nt} + \frac{1}{\sqrt{n}} I_{nt}^{-1/2} X.
\]

Then, we have:

\[
\sum_{i=1}^{n} \log h(y_{i,t} \mid f_t) = \sum_{i=1}^{n} \log h(y_{i,t} \mid \hat{f}_{nt}) - \frac{1}{2} X^2 + \frac{1}{6\sqrt{n}} J_{nt} X^3 + \frac{1}{24n} K_{nt} X^4 + o(1/n),
\]

where:

\[
J_{nt} = I_{nt}^{-3/2} S_{nt} \quad \text{and} \quad K_{nt} = I_{nt}^{-2} \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^4 \log h}{\partial f_t^4} (y_{i,t} \mid \hat{f}_{nt}).
\]

Thus:

\[
\prod_{i=1}^{n} h(y_{i,t} \mid f_t) = \prod_{i=1}^{n} h(y_{i,t} \mid \hat{f}_{nt}) \exp \left( -\frac{1}{2} X^2 \right) \exp \left( \frac{1}{6\sqrt{n}} J_{nt} X^3 + \frac{1}{24n} K_{nt} X^4 + o(1/n) \right)
\]

\[= \prod_{i=1}^{n} h(y_{i,t} \mid \hat{f}_{nt}) \exp \left( -\frac{1}{2} X^2 \right)
\]

\[\quad \left[ 1 + \frac{1}{6\sqrt{n}} J_{nt} X^3 + \frac{1}{24n} K_{nt} X^4 + \frac{1}{72n} J_{nt}^2 X^6 + o(1/n) \right].
\]

(A.2)
Similarly, we have an expansion for \( \log g(f_t|f_{t-1}) \) as:

\[
\log g(f_t|f_{t-1}) = \log g \left( \hat{f}_{nt} + \frac{1}{\sqrt{n}} I_{nt}^{-1/2} X | f_{t-1} \right) = \log g \left( \hat{f}_{nt} | f_{t-1} \right) + \frac{1}{\sqrt{n}} I_{nt}^{-1/2} A_{nt} X + \frac{1}{2n} I_{nt}^{-1} B_{nt} X^2 + o(1/n),
\]

where:

\[
A_{nt} = \frac{\partial \log g}{\partial f_t} \left( \hat{f}_{nt} | f_{t-1} \right) \quad \text{and} \quad B_{nt} = \frac{\partial^2 \log g}{\partial f_t^2} \left( \hat{f}_{nt} | f_{t-1} \right) .
\]

Thus:

\[
g(f_t|f_{t-1}) = g \left( \hat{f}_{nt} | f_{t-1} \right) \exp \left( \frac{1}{\sqrt{n}} I_{nt}^{-1/2} A_{nt} X + \frac{1}{2n} I_{nt}^{-1} B_{nt} X^2 + o(1/n) \right) = g \left( \hat{f}_{nt} | f_{t-1} \right) \left[ 1 + \frac{1}{\sqrt{n}} I_{nt}^{-1/2} A_{nt} X + \frac{1}{2n} I_{nt}^{-1} B_{nt} X^2 + \frac{1}{2n} I_{nt}^{-1} A_{nt}^2 X^2 + o(1/n) \right].
\]

(A.3)

Finally, we have an expansion for \( \exp (uf_t) \):

\[
\exp (uf_t) = \exp (u \hat{f}_{nt}) \exp \left( \frac{u}{\sqrt{n}} I_{nt}^{-1/2} X \right) = \exp \left( u \hat{f}_{nt} \right) \left[ 1 + \frac{u}{\sqrt{n}} I_{nt}^{-1/2} X + \frac{u^2}{2n} I_{nt}^{-1} X^2 + o(1/n) \right]. \quad (A.4)
\]

Let us now substitute expansions (A.2)-(A.4) into the numerator in equation (A.1) (the denominator is obtained by setting \( u = 0 \)). We have:

\[
\int e^{uf_t} g(f_t|f_{t-1}) \prod_{i=1}^n h(y_{i,t}|f_t) df_t = e^{u \hat{f}_{nt}} \prod_{i=1}^n h \left( y_{i,t} | \hat{f}_{nt} \right) g \left( \hat{f}_{nt} | f_{t-1} \right)
\]

\[
E_X \left[ \left( 1 + \frac{u}{\sqrt{n}} I_{nt}^{-1/2} X + \frac{u^2}{2n} I_{nt}^{-1} X^2 + o(1/n) \right) \left( 1 + \frac{1}{\sqrt{n}} I_{nt}^{-1/2} A_{nt} X + \frac{1}{2n} I_{nt}^{-1} \left( B_{nt} + A_{nt}^2 \right) X^2 + o(1/n) \right) \left( 1 + \frac{1}{6\sqrt{n}} J_{nt} X^3 + \frac{1}{24n} K_{nt} X^4 + \frac{1}{72n} f_{nt}^2 X^6 + o(1/n) \right) \right],
\]

where the expectation \( E_X \) is w.r.t. the standard normal variable \( X \). Since odd power moments of \( X \) are equal to zero, the terms of order \( 1/\sqrt{n} \) (and similarly the terms of order
$1/(n\sqrt{n})$, if the expansion is considered up to order $1/n^2$ can be obtained by replacing $f_{t-1}$ by $\hat{f}_{n,t-1}$. We have:

$$
\Psi_{nt}(u) = \exp \left\{ u \hat{f}_{nt} + \frac{u}{n} \left( I_{nt}^{-1} \frac{\partial \log g}{\partial f_t} (\hat{f}_{nt}|f_{t-1}) + \frac{1}{2} I_{nt}^{-2} S_{nt} \right) + \frac{1}{2n} u^2 I_{nt}^{-1} + o(1/n) \right\}.
$$

Then, Proposition 1 follows.
Appendix 2: Examples

The aim of this Appendix is to derive the summary statistics \( \hat{f}_{nt}, I_{nt} \) and \( S_{nt} \), and the associated approximations of the predictive density, for the examples considered in Section 2.2 and 3.3.

i) Linear factor model

For the linear state space model we have:

\[
\hat{f}_{nt} = \bar{y}_{nt}, \quad I_{nt} = \frac{1}{\sigma^2}, \quad S_{nt} = 0.
\]

The predictive density of the factor at order \( 1/n \) becomes:

\[
N \left( \bar{y}_{nt} + \frac{\sigma^2}{n} \frac{\partial \log g}{\partial \hat{f}_t} (\bar{y}_{nt}|\bar{y}_{n,t-1}) ; \frac{\sigma^2}{n} \right).
\]

If the common factor is Gaussian autoregressive:

\[
f_t|f_{t-1} \sim N (\mu + \rho (f_{t-1} - \mu) ; \eta^2),
\]

with \(|\rho| < 1\), we get:

\[
\frac{\partial \log g}{\partial \hat{f}_t} (\bar{y}_{nt}|\bar{y}_{n,t-1}) = -\frac{1}{\eta^2} [\bar{y}_{nt} - \rho (\bar{y}_{n,t-1} - \mu)].
\]

The predictive density becomes:

\[
N \left( \mu + \left(1 - \frac{1}{n \eta^2} \right) (\bar{y}_{nt} - \mu) + \frac{1}{n \eta^2} \rho (\bar{y}_{n,t-1} - \mu) ; \frac{\sigma^2}{n} \right).
\]

The predictive mean corresponds to the unconditional factor mean \( \mu \) corrected by a convex linear combination of \( \bar{y}_{nt} - \mu \) and \( \rho (\bar{y}_{n,t-1} - \mu) \), with weights \( 1 - \frac{1}{n \eta^2} \) and \( \frac{1}{n \eta^2} \), respectively.

ii) One-factor stochastic volatility model

We have:

\[
\hat{f}_{nt} = \frac{1}{n} \sum_{i=1}^{n} (y_{it} - \mu)^2, \quad I_{nt} = \frac{1}{2 \hat{f}_{nt}^2}, \quad S_{nt} = \frac{2}{\hat{f}_{nt}^3}.
\]
The predictive density of $f_t$ at order $1/n$ is given by:

$$N \left( \hat{f}_{nt} + \frac{2}{n} \hat{f}_{nt} \frac{\partial \log g}{\partial f_t} \left( \hat{f}_{nt} \mid \hat{f}_{nt-1} \right) + \frac{4}{n} \hat{f}_{nt}, \frac{2}{n} \hat{f}_{nt}^2 \right).$$

**iii) One-factor corporate spread model**

Estimator $\hat{f}_{nt}$ is defined as the solution of the equation:

$$-\frac{d \log \Gamma}{df} \left( \hat{f}_{nt} \right) + \frac{1}{n} \sum_{i=1}^{n} \log y_{it} + \log \lambda = 0.$$

Further, we have:

$$I_{nt} = \frac{d^2 \log \Gamma}{df^2} \left( \hat{f}_{nt} \right), \quad S_{nt} = -\frac{d^3 \log \Gamma}{df^3} \left( \hat{f}_{nt} \right).$$

**iv) Homogenous class of CDS**

From the conditional density in (2.5) we deduce that $\hat{f}_{nt}$ is the solution of the equation:

$$\frac{d \log \Gamma}{ds} \left( \hat{f}_{nt} \right) - \alpha \frac{d \log \Gamma}{ds} \left( \alpha \hat{f}_{nt} \right) - (1 - \alpha) \frac{d \log \Gamma}{ds} \left[ (1 - \alpha) \hat{f}_{nt} \right]$$

$$+ \alpha \frac{1}{n} \sum_{i=1}^{n} \log y_{it} + (1 - \alpha) \frac{1}{n} \sum_{i=1}^{n} \log (1 - y_{it}) = 0.$$

Further, we have:

$$I_{nt} = \alpha ^2 \frac{d^2 \log \Gamma}{ds^2} \left( \alpha \hat{f}_{nt} \right) + (1 - \alpha)^2 \frac{d^2 \log \Gamma}{ds^2} \left[ (1 - \alpha) \hat{f}_{nt} \right] - \frac{d^2 \log \Gamma}{ds^2} \left( \hat{f}_{nt} \right),$$

and:

$$S_{nt} = -\alpha ^3 \frac{d^3 \log \Gamma}{ds^3} \left( \alpha \hat{f}_{nt} \right) - (1 - \alpha)^3 \frac{d^3 \log \Gamma}{ds^3} \left[ (1 - \alpha) \hat{f}_{nt} \right] + \frac{d^3 \log \Gamma}{ds^3} \left( \hat{f}_{nt} \right).$$

**v) Generalized one-factor firm value model [proof of (3.5)]**

In this model described in Section 3.3, we have:

$$\hat{f}_{nt} = \bar{y}_{nt}, \quad I_{nt} = \frac{1}{\bar{y}_{nt} (1 - \bar{y}_{nt})}, \quad S_{nt} = 2 \frac{1 - 2 \bar{y}_{nt}}{\bar{y}_{nt}^2 (1 - \bar{y}_{nt})^2}.$$

Then (3.5) follows from (3.2).
Appendix 3: Proof of Proposition 3

i) Stochastic expansion of \( \hat{f}_{n,t+h} \)

We first derive the higher order asymptotic expansion of the ML estimator 
\[
\hat{f}_{n,t+h} = \arg\max_{f_{t+h}} \sum_{i=1}^{n} \log h (y_{i,t+h}|f_{t+h}) \] 
at order \( 1/n \). From the results on ML estimation [see Gouriéroux, Monfort (1995), Chapter 23], this asymptotic expansion is given by:
\[
\hat{f}_{n,t+h} - f_{t+h} = \frac{1}{\sqrt{n}} I_{t+h}^{-1} A_{n,t+h} + \frac{1}{n} \left[ I_{t+h}^{-2} A_{n,t+h} B_{n,t+h} + I_{t+h}^{-3} S_{t+h} A_{n,t+h}^2 \right] + o(1/n), \tag{A.5}
\]

where:
\[
I_{t+h} = E \left[ -\frac{\partial^2 \log h (y_{i,t+h}|f_{t+h})}{\partial f^2} | f_{t+h} \right], \quad S_{t+h} = E \left[ \frac{\partial^3 \log h (y_{i,t+h}|f_{t+h})}{\partial f^3} | f_{t+h} \right],
\]

and:
\[
A_{n,t+h} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial \log h}{\partial f} (y_{i,t+h}|f_{t+h}) , \quad B_{n,t+h} = \frac{1}{n} \sum_{i=1}^{n} \left[ -\frac{\partial^2 \log h (y_{i,t+h}|f_{t+h})}{\partial f^2} - I_{t+h} \right].
\]

Conditionally on \( \Omega_t \) and \( (f_t) \), the random vector \( C_{n,t+h} := (A_{n,t+h}, B_{n,t+h})' \) is such that:
\[
E [C_{n,t+h}|\Omega_t, (f_t)] = 0, \tag{A.6}
\]

and:
\[
V[C_{n,t+h}|\Omega_t, (f_t)] = \begin{bmatrix} I_{t+h} & -S_{t+h} \\ -S_{t+h} & K_{t+h} \end{bmatrix}, \tag{A.7}
\]

where \( K_{t+h} = V \left[ \frac{\partial^2 \log h (y_{i,t+h}|f_{t+h})}{\partial f^2} | f_{t+h} \right]. \)

ii) Asymptotic expansion of the derivative price

The price at \( t \) of the derivative with payoff \( \exp \left( u \hat{f}_{n,t+h} \right) \) is given by:
\[
\pi_{n,t}(u, h) = E \left[ m(f_t)m(f_{t+1})\ldots m(f_{t+h-1}) e^{u\hat{f}_{n,t+h}} | \Omega_t \right] = E \left[ m(f_t)m(f_{t+1})\ldots m(f_{t+h-1}) E \left[ e^{u\hat{f}_{n,t+h}} | \Omega_t, (f_t) \right] | \Omega_t \right].
\]
By using stochastic expansion (A.5), we have:

\[
E \left[ e^{uf_{n,t+h}} \mid \Omega_t, (f_t) \right] = e^{uf_{t+h}} E \left[ \exp \left( \frac{u}{\sqrt{n}} I^{-1}_{t+h} A_n, t+h + \frac{u}{n} \left[ I^{-2}_{t+h} A_n, t+h B_n, t+h + I^{-3}_{t+h} S_{t+h} A^2_{n, t+h} \right] \right) \mid (f_t) \right] + o(1/n).
\]

By expanding the exponential function, and using (A.6)-(A.7), we get:

\[
E \left[ e^{uf_{n,t+h}} \mid \Omega_t, (f_t) \right] = e^{uf_{t+h}} \left( 1 + \frac{u^2}{2n} I^{-1}_{t+h} \right) + o(1/n) = \exp \left( uf_{t+h} + \frac{u^2}{2n} I^{-1}_{t+h} \right) + o(1/n).
\]

We conclude using the Law of Iterated Expectation and the Markov property of process \((f_t)\):

\[
\pi_{n,t}(u, h) = E \left[ m(f_t)m(f_{t+1})...m(f_{t+h-1}) \exp \left( uf_{t+h} + \frac{u^2}{2n} I^{-1}_{t+h} \right) \mid \Omega_t \right] + o(1/n)
\]

\[
= E \left[ E \left[ m(f_t)m(f_{t+1})...m(f_{t+h-1}) \exp \left( uf_{t+h} + \frac{u^2}{2n} I^{-1}_{t+h} \right) \mid f_t \right] \mid \Omega_t \right] + o(1/n).
\]

**Appendix 4: Numerical illustration to basket default derivatives**

**i) Computation by simulation of the theoretical price of exponential derivatives**

The theoretical price \(\pi_{n,t}(u, 1)\) in (4.2) can be computed with Monte-Carlo simulation by drawing first a sample \(\{f_t^s\}_{s=1,...,S}\) from the density of \(f_t\) given \(\Omega_t\), and then drawing a value \(f_{t+1}^s\) from the density of \(f_{t+1}\) given \(f_t^s\) for any \(s = 1, ..., S\). To simulate from the density of \(f_t\) given \(\Omega_t\) we can use the acceptance-rejection algorithm [e.g., Robert, Casella (2004)]. We have from (4.4):

\[
g(f_t \mid \Omega_t) \propto (f_t)^{n_{t-1} - n_t} (1 - f_t)^{n_t} g(f_t \mid f_{t-1}),
\]

where \(\propto\) denotes equality of the densities up to a scale factor (depending on the conditioning variables only). By definition of the ML estimator \(\hat{f}_{n,t}\), we have:

\[
(f_t)^{n_{t-1} - n_t} (1 - f_t)^{n_t} \leq (\hat{f}_{n,t})^{n_{t-1} - n_t} \left( 1 - \hat{f}_{n,t} \right)^{n_t}, \text{ for any value of } f_t.
\]
Thus, the density of $f_t$ given $\Omega_t$ is upper bounded, and the density of $f_t$ given $f_{t-1}$ is a majorizing density. The acceptance-rejection algorithm to draw $f_t^s$ works as follows:

i) Generate a random draw $\tilde{f}_t$ from the density of $f_t$ given $f_{t-1}$.

ii) Generate a uniform variable $U \sim U(0, 1)$, independent of $\tilde{f}_t$.

iii) If:

$$U \leq \frac{(\tilde{f}_t)^{n_{t-1}-n_t} (1 - \tilde{f}_t)^{n_t}}{(\hat{f}_{n,t})^{n_{t-1}-n_t} (1 - \hat{f}_{n,t})^{n_t}},$$

then $f_t^s = \tilde{f}_t$. Otherwise, return to i).

ii) Approximation of the price of exponential derivatives

Approximation (4.6) follows by applying Corollary 2 for $n_{t-1} \to \infty$ with:

$$n_{t-1} = \sum_{i=1}^n (1 - y_{i,t-1}), \quad \hat{f}_{nt} = \frac{1}{n_{t-1}} \sum_{i=1}^n y_{i,t}, \quad I_{nt} = \frac{1}{\hat{f}_{nt} (1 - \hat{f}_{nt})}, \quad S_{nt} = 2 \frac{1 - 2 \hat{f}_{nt}}{\hat{f}_{nt}^2 (1 - \hat{f}_{nt})^2},$$

[see Appendix 2 v]). Function $\varphi(f_t; u)$ and its first and second order derivatives can be computed by Monte-Carlo simulation, by using:

$$\frac{d\varphi(f_t; u)}{df_t} = E \left[ \left\{ 1 + \left( \exp \left( \frac{u}{n_t} \right) - 1 \right) f_{t+1} \right\}^{n_t} \frac{\partial \log g}{\partial f_t} (f_{t+1}|f_t) \mid f_t \right],$$

and similarly for the second order derivative.
Table 1: Percentage pricing errors: Gaussian approximation

<table>
<thead>
<tr>
<th>$\rho = 0.01$</th>
<th>n=100</th>
<th>n=1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>.025</td>
<td>.05</td>
</tr>
<tr>
<td>Mean</td>
<td>.0254</td>
<td>.0068</td>
</tr>
<tr>
<td>Low 5% Qu</td>
<td>-.0497</td>
<td>-.5362</td>
</tr>
<tr>
<td>Median</td>
<td>.0029</td>
<td>.0009</td>
</tr>
<tr>
<td>Up 5% Qu</td>
<td>.1808</td>
<td>.5757</td>
</tr>
<tr>
<td>Mean abs</td>
<td>.0460</td>
<td>.2682</td>
</tr>
<tr>
<td>Median abs</td>
<td>.0201</td>
<td>.2179</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\rho = 0.10$</th>
<th>n=100</th>
<th>n=1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>.0260</td>
<td>.0287</td>
</tr>
<tr>
<td>Mean</td>
<td>-.0065</td>
<td>-.0023</td>
</tr>
<tr>
<td>Low 5% Qu</td>
<td>-.4735</td>
<td>-.8361</td>
</tr>
<tr>
<td>Median</td>
<td>-.0462</td>
<td>-.0983</td>
</tr>
<tr>
<td>Up 5% Qu</td>
<td>.5385</td>
<td>.7743</td>
</tr>
<tr>
<td>Mean abs</td>
<td>.1701</td>
<td>.2758</td>
</tr>
<tr>
<td>Median abs</td>
<td>.0792</td>
<td>.1662</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\rho = 0.30$</th>
<th>n=100</th>
<th>n=1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>.025</td>
<td>.05</td>
</tr>
<tr>
<td>Mean</td>
<td>-.0573</td>
<td>-.0750</td>
</tr>
<tr>
<td>Low 5% Qu</td>
<td>-.4988</td>
<td>-.6628</td>
</tr>
<tr>
<td>Median</td>
<td>-.0027</td>
<td>-.0029</td>
</tr>
<tr>
<td>Up 5% Qu</td>
<td>.0925</td>
<td>.1408</td>
</tr>
<tr>
<td>Mean abs</td>
<td>.0972</td>
<td>.1361</td>
</tr>
<tr>
<td>Median abs</td>
<td>.0358</td>
<td>.0505</td>
</tr>
</tbody>
</table>

45
Table 2: Percentage pricing errors: Approximation based on formula (3.3)

<table>
<thead>
<tr>
<th>$\rho = 0.01$</th>
<th>$n=100$</th>
<th>$n=1000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>.025 .05 .10 .125</td>
<td>.025 .05 .10 .125</td>
</tr>
<tr>
<td>Mean</td>
<td>-.0112 -.0034 price below tick</td>
<td>-.0035 .0074 price below tick</td>
</tr>
<tr>
<td>Low 5% Qu</td>
<td>-.1705 -.3895 price below tick</td>
<td>-.0496 -.3358 price below tick</td>
</tr>
<tr>
<td>Median</td>
<td>.0052 .0139 price below tick</td>
<td>.0005 .0096 price below tick</td>
</tr>
<tr>
<td>Up 5% Qu</td>
<td>.1468 .4727 price below tick</td>
<td>.0295 .3501 price below tick</td>
</tr>
<tr>
<td>Mean abs</td>
<td>.0799 .2146 price below tick</td>
<td>.0175 .1659 price below tick</td>
</tr>
<tr>
<td>Median abs</td>
<td>.0519 .1471 price below tick</td>
<td>.0126 .1415 price below tick</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\rho = 0.10$</th>
<th>$n=100$</th>
<th>$n=1000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>.025 .05 .10 .125</td>
<td>.025 .05 .10 .125</td>
</tr>
<tr>
<td>Mean</td>
<td>-.0205 -.0247 -.0094 -.0062</td>
<td>-.0055 -.0079 -.0000 -.0089</td>
</tr>
<tr>
<td>Low 5% Qu</td>
<td>-.2765 -.4429 -.6233 -.6892</td>
<td>-.0525 -.0969 -.1116 -.0944</td>
</tr>
<tr>
<td>Median</td>
<td>-.0027 -.0044 .0061 .0001</td>
<td>.0001 .0006 .0050 .0106</td>
</tr>
<tr>
<td>Up 5% Qu</td>
<td>.1109 .2191 .4401 .5393</td>
<td>.0239 .0506 .0987 .1048</td>
</tr>
<tr>
<td>Mean abs</td>
<td>.0741 .1314 .2206 .2629</td>
<td>.0157 .0305 .0455 .0468</td>
</tr>
<tr>
<td>Median abs</td>
<td>.0483 .0915 .1626 .1990</td>
<td>.0077 .0176 .0324 .0353</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\rho = 0.30$</th>
<th>$n=100$</th>
<th>$n=1000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>.025 .05 .10 .125</td>
<td>.025 .05 .10 .125</td>
</tr>
<tr>
<td>Mean</td>
<td>-.0300 -.0358 -.0400 -.0406</td>
<td>-.0086 -.0102 -.0130 -.0139</td>
</tr>
<tr>
<td>Low 5% Qu</td>
<td>-.3511 -.4382 -.5310 -.5564</td>
<td>-.0631 -.0846 -.1207 -.1477</td>
</tr>
<tr>
<td>Median</td>
<td>.0060 .0097 .0131 .0147</td>
<td>.0003 .0003 .0002 .0000</td>
</tr>
<tr>
<td>Up 5% Qu</td>
<td>.0835 .1205 .1743 .1946</td>
<td>.0171 .0260 .0435 .0542</td>
</tr>
<tr>
<td>Mean abs</td>
<td>.0709 .0952 .1261 .1375</td>
<td>.0164 .0219 .0315 .0369</td>
</tr>
<tr>
<td>Median abs</td>
<td>.0271 .0388 .0549 .0628</td>
<td>.0048 .0070 .0114 .0141</td>
</tr>
</tbody>
</table>
Figure 1: Transition density of the factor

The Figure plots the conditional distribution of factor $f_t$ given $f_{t-1}$, for different values of $f_{t-1}$. The conditioning values of $f_{t-1}$ are given in terms of their corresponding Gaussian factor values $F_{t-1}$; they are $F_{t-1} = 0$, $F_{t-1} = 2$, $F_{t-1} = -2$, respectively. Asset correlation is $\rho = 0.10$. 
Figure 2: Predictive distribution of the factor given investors’ information
The Figure plots the conditional distribution of factor $f_t$ given $f_{t-1}$, $n_{t-1} = 1000$, and $n_t = 960$, for different values of $f_{t-1}$. The conditioning values of $f_{t-1}$ are given in terms of their corresponding Gaussian factor values $F_{t-1}$; they are $F_{t-1} = 0$, $F_{t-1} = 2$, and $F_{t-1} = -2$ respectively. Asset correlation is $\rho = 0.10$. 
Figure 3: Gaussian approximation of the factor predictive density
The Figure plots the approximate conditional distribution of $f_t$ given past default history $n_{t-1} = 1000, n_t = 960, \hat{f}_{n,t-1}$, for different values of $\hat{f}_{n,t-1}$. Asset correlation is $\rho = 0.10$. 

49
Figure 4: Price of the $\alpha$-to-default swap for different values of the asset correlation
The Figure plots the price of $\alpha$-to-default swap at time-to-maturity 1 year as a function of $\alpha$, for three different values of asset correlation $\rho = 0.01$ (dotted line), $\rho = 0.10$ (solid line) and $\rho = 0.30$ (dashed line). The lagged Gaussian factor value is $F_{t-1} = 0$. 
Figure 5: Price of the $\alpha$-to-default swap for different values of the lagged factor

The Figure plots the price of $\alpha$-to-default swap at time-to-maturity 1 year as a function of $\alpha$, for three different values of the lagged factor $F_{t-1} = 0$, $F_{t-1} = -2$, and $F_{t-1} = 2$. Asset correlation is $\rho = 0.10$. 
Figure 6: Approximate price of the $\alpha$-to-default swap for different values of asset correlation
The Figure plots the approximate price of $\alpha$-to-default swap at time-to-maturity 1 year as a function of $\alpha$, for three different values of asset correlation $\rho = 0.01$ (dotted line), $\rho = 0.10$ (solid line) and $\rho = 0.30$ (dashed line). The past default history is such that $\hat{f}_{n,t-1} = 0.04$. 
Figure 7: Approximate price of the $\alpha$-to-default swap for different values of the lagged factor
The Figure plots the approximate price of $\alpha$-to-default swap at time-to-maturity 1 year as a function of $\alpha$, for different values of $\hat{f}_{n,t-1}$. Asset correlation $\rho = 0.10$. 
Figure 8: **Percentage pricing errors for default frequency** $\alpha = 5\%$

The upper Panels display scatter plots of percentage pricing errors vs. theoretical prices (left), and vs. estimated factor $\hat{f}_{n,t}$ (right). The lower Panels display the quantile of percentage pricing errors conditional on theoretical price (left), and conditional on estimated factor $\hat{f}_{n,t}$ (right), at levels 5%, 50%, 95%. The frequency is $\alpha = 0.025$ and the initial size of the pool is $n = 1000$. Asset correlation is $\rho = 0.10$. 
Figure 9: Percentage pricing errors for default frequency $\alpha = 12.5\%$

The upper Panels display scatter plots of percentage pricing errors vs. theoretical prices (left), and vs. estimated factor $\hat{f}_{n,t}$ (right). The lower Panels display the quantile of percentage pricing errors conditional on theoretical price (left), and conditional on estimated factor $\hat{f}_{n,t}$ (right), at levels 5%, 50%, 95%. The frequency is $\alpha = 0.125$ and the initial size of the pool is $n = 1000$. Asset correlation is $\rho = 0.10$. 