Pricing Default Events : Surprise, Exogeneity and Contagion[*]

C., GOURIEROUX [(1)], A. MONFORT [(2)], and J.P., RENNE [(3)]

(January, 2013, revised October, 2013)

Abstract: In order to derive closed-form expressions of the prices of credit derivatives, standard credit-risk models typically price the default intensities, but not the default events themselves. The default indicator is replaced by an appropriate prediction and the prediction error, that is the default-event surprise, is neglected. Our paper develops an approach to get closed-form expressions for the prices of credit derivatives written on multiple names without neglecting default-event surprises. This approach differs from the standard one, since the default counts necessarily cause the factor process under the risk-neutral probability, even if this is not the case under the historical probability. This implies that the standard exponential pricing formula of default does not apply. Using U.S. bond data, we show that allowing for the pricing of default events has important implications in terms of both data-fitting and model-implied physical probabilities of default. In particular, it may provide a solution to the credit spread puzzle. Besides, we show how our approach can be used to account for the propagation of defaults on the prices of credit derivatives.

Keywords: Credit Derivative, Default Event, Default Intensity, Frailty, Contagion, Credit Spread Puzzle.
JEL Codes: E43, E47, G12.

[1]CREST, and University of Toronto.
[2]CREST, Banque de France, and University of Maastricht.
[3]Corresponding author. Banque de France. e-mail: jean-paul.renne@m4x.org

1 Introduction

Two alternative approaches are usually followed to price credit derivatives such as Credit Default Swaps (CDS). In the structural approach introduced by Merton the default of a corporation occurs when the asset side of the balance sheet becomes smaller than its liability side. Then the probability and the price of default are deduced from the historical and risk-neutral properties of these two underlying variables. Another approach is the reduced-form or intensity approach, in which the underlying phenomena are not explicitly modeled and the historical default intensity, assumed to exist, is directly analyzed [see e.g. Duffie, Singleton (1999)]. The latter approach is easily implemented in the framework of factor models, when the default intensity and the stochastic discount factor (s.d.f.) are exponential affine functions of these factors and when these factors feature an affine dynamics [see Duffie, Filipovic, Schachermayer (2003), Duffie (2005) in continuous time, Gouriéroux, Monfort, Polimenis (2006) in discrete time]. Indeed the term structure of riskfree as well as risky interest rates admit closed-form expressions and are affine functions of these factors.

However in order to derive these closed-form expressions of interest rates and prices, the reduced-form approach usually prices the default intensity, but not the default indicator itself. In other words, the default indicator does not appear in the s.d.f. and is replaced by an appropriate prediction. Thus the prediction error, that is the surprise on default event, is neglected.

There exist a few papers mentioning this approximation and trying to adjust for this practice [see e.g. Jarrow, Yu (2001), Bai, Collin-Dufresne, Goldstein, Helwege (2012)a, b]. However, in this context, pricing formulas have no longer closed forms and it seems much more difficult to account for default correlations. This explains why some of these analyses have considered the joint pricing of default for a small number of names, for instance two names in Jarrow, Yu (2001) while others focus on recursive credit events, which do not imply default [see e.g. Bai, Collin-Dufresne, Goldstein, Helwege (2012)a, eq (15)].

Our paper develops an approach that results in closed-form formula to price credit

derivatives written on any number of names, without neglecting default-event surprises. In Section 2 we review the standard reduced form approach and its limitations. In particular we carefully discuss the link between the assumption that the default count process does not cause the factor process and the existence of a default pre-intensity, both under the historical and risk-neutral dynamics. In Section 3, we consider a homogeneous pool of credits and introduce a pricing model, with a joint Compound AutoRegressive (CaR) dynamics for the factor and the default count. When the s.d.f. is exponential affine in both factor and default count, we get linear affine formulas for the term structures of riskfree and risky interest rates. The results are extended in Section 4 to account for a possible heterogeneity of the initial pool of credits. We consider that this pool can be partitioned into J homogeneous segments. The model allows for a common systematic factor (e.g. dynamic frailty), and also for contagion phenomena, where a default-event surprise of segment j may have an impact on the prices of credit derivatives written on another segment. Section 5 provides illustrations of our approach. The observation of a wide gap between Credit Default Swap (CDS) spreads, that can be seen as default-loss expectations under the risk-neutral measure, and expected default losses is usually called the *credit-spread puzzle* in the literature [see e.g. D'Amato, Remolona (2003), Hull, Predescu, White (2005)]. The standard credit-risk models, that do not price default-event surprises, deal with this puzzle by incorporating credit-risk premia. But these premia are too small for short maturities. By contrast, we show that pricing default-event surprises may solve the credit-puzzle for all maturities, including the shortest ones. To highlight this feature, we calibrate our model on U.S. banking-sector bond data covering the last twenty years. Our results suggest that neglecting the pricing of default events is likely to result in an overestimation of model-implied physical probabilities of defaults for short-term horizons. We also illustrate how our approach can be exploited to investigate the effect of the propagation of defaults on the prices of credit derivatives. Section 6 concludes. Proofs are gathered in appendices.

2 The standard reduced-form approach and its limitation

2.1 Basic assumptions

We consider a pool of I entities, indexed by $i = 1, \ldots, I$; these entities can be firms or credit contracts. We denote by $d_{i,t}$ the indicator of default of entity i, that is $d_{i,t} = 1$, if entity i is in default at time t (or before), and $d_{i,t} = 0$, otherwise.

We introduce the notations: $d_t = (d_{1,t}, \ldots, d_{I,t})'$, $\underline{d_t} = (d'_1, \ldots, d'_t)'$, $PaR_t = \{i | d_{i,t} = 0\}$, $n_t = \Sigma_{i \in PaR_{t-1}} d_{i,t}$, $N_t = \Sigma_{\tau=1}^t n_\tau$.

PaR_t is the Population-at-Risk that is the set of entities still alive at date t, n_t is the number of defaults occurring at date t and N_t is the number of defaults at date t or before.

Let us first assume that the pool is homogeneous and that the default dependence is driven by an exogenous (multivariate) factor F_t. Let us denote by $\Omega_t = (\underline{F_t}, \underline{d_t})$, $\Omega_t^* = (F_{t+1}, \underline{d_t}) = (F_{t+1}, \Omega_t)$ the information sets, where $\underline{F_t} = \{F_\tau, \tau \leq t\}$. Thus we have nested filtrations satisfying $\Omega_t \subset \Omega_t^* \subset \Omega_{t+1}$. The assumptions, which will be relaxed in Subsection 2.5 and in Section 4, can be formalized in the following way:

Assumption A_0: i) The variables $d_{i,t+1}, i = 1, \ldots, I$, are independent conditional on $\Omega_t^* = (F_{t+1}, \Omega_t)$, state 1 is an absorbing state, the variables $\{d_{i,t+1}, i \in PaR_t\}$ are identically distributed, conditional on[1] Ω_t^*, and this conditional distribution is only function of F_{t+1}. **ii)** The binomial distribution of n_{t+1} given Ω_t^* is approximated by a Poisson distribution depending on F_{t+1} only. **iii)** The conditional distribution of F_{t+1} given Ω_t is equal to the distribution of F_{t+1} given F_t.

Assumption A_0 ii) means that the number of entities still alive at t, that is $I - N_t$, remains large and that the probability of default is small [see Gagliardini, Gouriéroux (2013)]. This assumption of a small probability of default implies that the default risk events are not diversifiable even for large population conditionally on the driving factor since the number

of default is not exploding [compare with Jarrow, Lando, Yu (2005)].

Assumption A_0 iii) means that the process (d_t) does not cause the process (F_t), or equivalently that process (F_t) is exogenous, and that process (F_t) is Markov of order 1. Let us recall that the usual definition of noncausality introduced by Granger is characterized by the conditions:

$$f(F_{t+1}|\underline{F_t}, \underline{d_t}) = f(F_{t+1}|\underline{F_t}), \forall t, \tag{2.1}$$

where $f(.|.)$ denotes a conditional probability density function (p.d.f) [Granger (1980)]. Moreover, it is easily shown by projecting the noncausality condition (2.1) on the information $\underline{F_t}, \underline{n_t}$ that we also have:

$$f(F_{t+1}|\underline{F_t}, \underline{n_t}) = f(F_{t+1}|\underline{F_t}), \forall t,$$

that is, the count process (n_t) does not cause process (F_t).

Assumptions A_0 i) and A_0 ii) also implies that $P(d_{i,t+1} = 0|d_{i,t} = 0, \Omega_t)$ only depends on F_t and not on $\underline{n_t}$. In other words, there is no contagion.

In the following, we will also need a technical result stating that these conditions are also equivalent to:[2]

$$f(d_{t+1}|\underline{d_t}, \underline{F_T}) = f(d_{t+1}|\underline{d_t}, \underline{F_{t+1}}) = f(d_{t+1}|\Omega_t^*), \forall t, T, T \geq t. \tag{2.2}$$

This is the Sims' definition of noncausality [Sims (1972)].

Let us now consider three pricing situations.

2.2 Case 1: no default events in the s.d.f. and exogenous factors (standard approach)

Under the assumption of no arbitrage opportunity, derivatives can be priced by introducing stochastic discount factors (s.d.f.) from the issuing date t_0, say. The standard pricing

approach assumes that the short term s.d.f. is specified as a function of the current factor value only, that is, the s.d.f. for period $(t, t+1)$ is of the type: $\tilde{m}_{t,t+1} = \tilde{m}(F_{t+1})$, say.[3] Then the price at t_0 of the derivative written on the total number of defaults and paying $g(N_{t_0+h})$ at date $t_0 + h$ is:

$$\tilde{\Pi}(g, h) = E_{t_0}\left[\Pi_{k=1}^{h}\tilde{m}_{t_0+k-1,t_0+k}g(N_{t_0+h})\right] = E_{t_0}\left[\Pi_{k=1}^{h}\tilde{m}(F_{t_0+k})g(N_{t_0+h})\right], \qquad (2.3)$$

where E_{t_0} is the conditional expectation given $\Omega_{t_0} = (\underline{F_{t_0}}, d_{t_0} = 0)$, since all the entities are alive at the issuing of the pool.

By applying the iterated expectation theorem, we get:

$$\tilde{\Pi}(g, h) = E_{t_0}\left[\Pi_{k=1}^{h}\tilde{m}(F_{t_0+k})E(g(N_{t_0+h})|\underline{F_{t_0+h}})\right] = \tilde{\Pi}(\tilde{g}, h), \qquad (2.4)$$

where $\tilde{g}(\underline{F_{t_0+h}}) = E[g(N_{t_0+h})|\underline{F_{t_0+h}}]$.

In other words it is equivalent to price the derivative with payoff $g(N_{t_0+h})$ written on the cumulated number of defaults, or to price the derivative with payoff $\tilde{g}(\underline{F_{t_0+h}})$ written on the factor history. Thus the choice of a s.d.f. that is function of the latent factor only greatly simplifies the derivation of closed-form formulas for the prices of credit derivatives [see e.g. Lando (1998), Duffee (1999), Duffie, Singleton (1999), Duffie (2005) for pricing in continuous time, Gouriéroux, Monfort, Polimenis (2006) for pricing in discrete time].

2.3 Case 2: default events in the s.d.f. and exogenous factors

The standard practice described above may induce mispricing, since the default events themselves have not been included in the s.d.f.. Let us now consider a short term s.d.f. depending on both F_{t+1} and n_{t+1}:

$$m_{t,t+1} = m(F_{t+1}, n_{t+1}). \qquad (2.5)$$

The pricing formula becomes:

$$\Pi(g, h) = E_{t_0} \left\{ \Pi_{k=1}^h m(F_{t_0+k}, n_{t_0+k}) g(N_{t_0+h}) \right\}. \tag{2.6}$$

Since the pricing operator is linear, we get:

$$
\begin{aligned}
\Pi(g, h) &= E_{t_0} \left[\Pi_{k=1}^h m(F_{t_0+k}, n_{t_0+k}) \tilde{g}(\underline{F_{t_0+h}}) \right] \\
&+ E_{t_0} \left\{ \Pi_{k=1}^h m(F_{t_0+k}, n_{t_0+k}) [g(N_{t_0+h}) - \tilde{g}(\underline{F_{t_0+h}})] \right\} \\
&= E_{t_0} \left\{ E[\Pi_{k=1}^h m(F_{t_0+k}, n_{t_0+k}) | \underline{F_{t_0+h}}] \tilde{g}(\underline{F_{t_0+h}}) \right\} \\
&+ E_{t_0} \left\{ \Pi_{k=1}^h m(F_{t_0+k}, n_{t_0+k}) [g(N_{t_0+h}) - \tilde{g}(\underline{F_{t_0+h}})] \right\}.
\end{aligned}
\tag{2.7}
$$

By the iterated expectation theorem and by using the Sims' version (2.2) of the non-causality assumption A_0 ii), we get (see Appendix A.1):

$$E \left[\Pi_{k=1}^h m(F_{t_0+k}, n_{t_0+k}) | \underline{F_{t_0+h}} \right] = \Pi_{k=1}^h \tilde{m}(F_{t_0+k}), \tag{2.8}$$

with:

$$\tilde{m}(F_{t+1}) = E[m(F_{t+1}, n_{t+1}) | F_{t+1}]. \tag{2.9}$$

Let us now interpret Equations (2.8)-(2.9). The pricer $\tilde{\Pi}$ based on factor values only is obtained by considering the expectation $\tilde{m}(F_{t+1})$ of the s.d.f. of different maturities conditional on factor values. From (2.8)-(2.9), we see that the projection of the s.d.f. for maturity h is the product of the short term projections. This feature is needed for these approximated s.d.f.'s to be compatible with no dynamic arbitrage opportunity for an investor using information $(\underline{F_t})$ in his portfolio updating and interested in pricing derivatives written on the factor process. However, even if the "projected" s.d.f.'s are time-consistent, they differ from the initial s.d.f., and this implies mispricing for derivatives written on default

6

counts. Let us discuss this mispricing. From (2.7), we get:

$$\Pi(g, h) = \Pi(\tilde{g}, h) + \Pi(g - \tilde{g}, h),$$

with $g - \tilde{g} = g(N_{t_0+h}) - E[g(N_{t_0+h})|\underline{F_{t_0+h}}]$. Then, by applying (2.8)-(2.9):

$$\Pi(g, h) = \tilde{\Pi}(\tilde{g}, h) + \Pi(g - \tilde{g}, h) = \tilde{\Pi}(g, h) + \Pi(g - \tilde{g}, h). \qquad (2.10)$$

The true price is the standard one based on the projected s.d.f. plus an adjustment term. This adjustment term $\Pi(g - \tilde{g}, h)$ is the price of the *surprise* on default events: $g(N_{t_0+h}) - E[g(N_{t_0+h})|\underline{F_{t_0+h}}]$. When the s.d.f. does not depend on the default-event surprise of date $t + 1$, this adjustment term vanishes. Otherwise, there is a risk premium for the surprise and a need for price adjustment.

When $m_{t,t+1}$ is proportional to $\exp(\delta'_F F_{t+1} + \delta_S n_{t+1})$, the projected s.d.f. $\tilde{m}_{t,t+1}$ will depend on δ_S. In particular, considering the payoff 1 at $t + h$, we see that the riskfree interest rate with residual maturity h at date t, denoted by $R(t, h)$, also depends on δ_S. Nevertheless, the price of the surprise is equal to zero.

2.4 Case 3: default events in the s.d.f. and non-exogenous factors

Let us now discuss how the results above are modified if we relax the noncausality assumption A_0 iii) under the historical distribution, that is, if we consider the new set of assumptions:

$$A_0^* = A_0 i) + A_0 ii) + A_0^* iii),$$

where:

Assumption A_0^* iii): The conditional historical distribution of F_{t+1} given Ω_t is equal to the distribution of F_{t+1} given F_t, n_t. Thus this conditional distribution can also depend on n_t.

When the noncausality, or exogeneity, condition is not satisfied, the decomposition (2.10) of the derivative price has to be modified. Indeed when the noncausality of process $(d_{i,t})$, that is Assumption A_0 iii), is no longer satisfied, we do not have equality (2.8) and $\Pi(\tilde{g}, h)$ becomes different from $\tilde{\Pi}(\tilde{g}, h)$. In this case we have the following decomposition of $\Pi(g, h)$:

$$\Pi(g, h) = \tilde{\Pi}(g, h) + [\Pi(\tilde{g}, h) - \tilde{\Pi}(\tilde{g}, h)] + \Pi(g - \tilde{g}, h), \tag{2.11}$$

using the fact that we still have $\tilde{\Pi}(g, h) = \tilde{\Pi}(\tilde{g}, h)$. The additional term between brackets is an adjustment term for causality of the count process and we have the decomposition:

Price=Standard Price + Causality Adjustment + Surprise Adjustment.

Of course, if $m_{t,t+1}$ is only function of F_{t+1}, the surprise adjustment disappears.

Finally let us discuss the expressions of the riskfree rate for an s.d.f. proportional to $\exp(\delta_F' F_{t+1} + \delta_S n_{t+1})$. We have seen in Subsection 2.3 that, if F_t is exogenous, the riskfree rate of residual maturity h is $R(t, h) \equiv r(h, F_t, \delta_S)$. Under Assumption A_0^*, the riskfree rate also depends on n_t: $R(t, h) \equiv r^*(h, F_t, n_t, \delta_S)$. We note two different effects of the introduction of defaults events in the s.d.f.. First the risk sensitivity δ_S still appears in the riskfree rate. Second, if the factor process is not exogenous, we observe "jumps" in the riskfree rate when default occurs in the sense that the formula valid in absence of default $r(h, F_t, \delta_S)$ is replaced by $r^*(h, F_t, n_t, \delta_S)$; moreover the magnitude of these jumps depends on the number of defaults.

2.5 Intensities, risk-neutral dynamics and exogeneity

Definition 1: The physical default intensity with respect to the filtration (Ω_t^*), denoted by $\lambda_{i,t}$, is defined by:

$$P(d_{i,t+1} = 0 | d_{i,t} = 0, \Omega_t^*) = \exp(-\lambda_{i,t+1}).$$

In Appendix A.2, we show how this definition of the *default intensity* is linked to the *intensity* of a point process in discrete-time [see Bremaud (1980)]. The notion of default pre-intensity

introduced by Duffie, Gârleanu (2001, p. 44) will also be useful.

Definition 2: The physical intensity $\lambda_{i,t}$ associated with the point process $(d_{i,t})$ is a default pre-intensity with respect to the filtration (Ω_t^*) if and only if:

$$P[\tau_i > t + h | \tau_i > t, \Omega_t^*] = E[\Pi_{k=1}^h \exp(-\lambda_{i,t+k}) | d_{i,t} = 0, \Omega_t^*], \forall t, h,$$

where $\tau_i = \inf\{t : d_{i,t} = 1\}$ is the lifetime of entity i.

In particular, it is clear, by taking $h = 1$, that if a default pre-intensity exists, it is equal to the default intensity. Note that the definition of default pre-intensity includes the exponential-type formula for the term structure of the probabilities of default.

Proposition 1: Under Assumption A_0 and under the historical probability, each point process $(d_{i,t})$ admits a default pre-intensity with respect to the filtration (Ω_t^*). This default pre-intensity does not depend on i and is denoted by λ_t.

Proof: In our framework, we have:

$$
\begin{aligned}
P[\tau_i > t + h | \tau_i > t, \Omega_t^*] &= E\{P[\tau_i > t + h | \tau_i > t, \Omega_t^*, \underline{F_{t+h}}] | \tau_i > t, \Omega_t^*\} \\
&= E\{\Pi_{k=1}^h P[\tau_i > t + k | \tau_i > t + k - 1, \Omega_t^*, \underline{F_{t+h}}] | \tau_i > t, \Omega_t^*\}.
\end{aligned}
$$

Using the Sims characterization of noncausality from $(d_{i,t})$ to (F_t), we get:

$$
\begin{aligned}
P[\tau_i > t + h | \tau_i > t, \Omega_t^*] &= E\{\Pi_{k=1}^h P[\tau_i > t + k | \tau_i > t + k - 1, \Omega_t^*, \underline{F_{t+k}}] | \tau_i > t, \Omega_t^*\} \\
&= E\{\Pi_{k=1}^h P[d_{i,t+k} = 0 | d_{i,t+k-1} = 0, \Omega_t^*, \underline{F_{t+k}}] | \tau_i > t, \Omega_t^*\} \\
&= E[\Pi_{k=1}^h \exp(-\lambda_{t+k}) | d_{i,t} = 0, \Omega_t^*].
\end{aligned}
$$

Thus the point process $(d_{i,t})$ admits the default pre-intensity λ_t.$\square$

If λ_{t+1} is small we have approximately: $P(d_{i,t+1} = 1 | d_{i,t} = 0, \Omega_t^*) \simeq \lambda_{t+1}$. This condition ($\lambda_{t+1}$ small) is usually satisfied if the time unit is small, that is, when the discrete time

approach tends to a continuous time approach and, then, $(1 - d_{i,t})\lambda_{t+1}$ is approximately the intensity of the point process $(d_{i,t})$ (see Appendix A.2).

Let us now consider the dynamics of the individual point processes under the risk-neutral distribution. For this purpose, we still assume A_0 and we assume that the s.d.f. $m_{t,t+1}$ is of the general form[4]:

$$m_{t,t+1} = \exp(\delta_0 + \delta_F' F_{t+1} + \delta_S n_{t+1}). \qquad (2.12)$$

Proposition 2: Under the risk-neutral dynamics,

i) The point processes $(d_{i,t})$, $i = 1, \ldots, I$, are still independent conditional on process (F_t);

ii) State 1 is still absorbing;

iii) We have: $Q[d_{i,t+1} = 0 | d_{i,t} = 0, \Omega_t^*] \equiv \exp(-\lambda_{t+1}^Q)$, where the risk-neutral default intensity is:

$$\lambda_{t+1}^Q = \lambda_{t+1} + \log\{\exp(-\lambda_{t+1}) + [1 - \exp(-\lambda_{t+1})]\exp(\delta_S)\}.$$

In particular, if $\delta_S = 0$, $\lambda_{t+1}^Q = \lambda_{t+1}$.

Proof: See Appendix A.3.

Since δ_S is expected to be nonnegative we have $\lambda_{t+1}^Q \geq \lambda_{t+1}$ $\forall t$. Moreover, $\lambda_{t+1}^Q = \lambda_t$, $\forall t$, if and only if $\delta_S = 0$. In other words, if the s.d.f. does not contain event variables, the historical and risk-neutral default intensities are the same functions of the factors.[5] If λ_{t+1} is small, it is easily checked that $\lambda_{t+1}^Q \simeq \lambda_{t+1}\exp(\delta_S)$.

Proposition 3: The risk-neutral p.d.f. of F_{t+1} given Ω_t is proportional to:

$$f_t^P(F_{t+1}|\Omega_t)\exp(\delta_F' F_{t+1})[\sum_{d_{i,t+1}=0}^{1} f^P(d_{i,t+1}|F_{t+1}, d_{i,t} = 0)\exp(\delta_S d_{i,t+1})]^{(I-N_t)},$$

where $f_t^P(F_{t+1}|\Omega_t)$ denotes the historical conditional p.d.f. of F_{t+1} given Ω_t.

Proof: see Appendix A.3.

Under the risk-neutral probability the distribution of F_{t+1} given Ω_t depends not only on F_t through $f_t^P(F_{t+1}|\Omega_t)$, but also on the cumulated default count N_t. Therefore the sequence of counts (n_t) will generally Granger cause the factor in the risk-neutral world. In particular, this implies that, conditional to Ω_t, the risk-neutral probability of default of entity i at date $t+1$ depends on (F_t, N_t).[6] This shows that contagion exists in the risk-neutral world. However, when $\delta_S = 0$, the sum appearing in the formula of Proposition 3 is equal to 1 and the dependency on N_t disappears. This discussion is summarized below.

Corollary 1: Under assumption A_0 and the exponential affine specification (2.12) of the s.d.f., the default count process (n_t) Q-causes the factor process (F_t) except if $\delta_S = 0$, that is, if the default-event surprise is priced at zero.

Proposition 4: If $\delta_S \neq 0$, a default pre-intensity does not exist in the risk-neutral world.

Proof: We know that, if the default pre-intensity exists, it is λ_t^Q. (This is obtained by setting $h = 1$ in Definition 2.) Let us now consider the quantity $Q[\tau_i > t + h|\tau_i > t, \Omega_t^*]$ and show that it cannot be equal to $E^Q[\Pi_{k=1}^h \exp(-\lambda_{t+k}^Q)|d_{i,t} = 0, \Omega_t^*]$. Indeed, since process (F_t) is no longer exogenous, we cannot replace $Q[\tau_i > t + k|\tau_i > t + k - 1, \Omega_t^*, \underline{F_{t+h}}]$ by $Q[\tau_i > t + k|\tau_i > t + k - 1, \Omega_t^*, \underline{F_{t+k}}]$ in the analogue of the proof of Proposition 1.□

The previous proposition has important consequences when pricing defaultable bonds. The price at date t of a defaultable bond with zero recovery rate and time-to-maturity h is:

$$B(t, h) = E_t^Q[\exp(-r_t - \ldots - r_{t+h-1})\mathbb{I}_{(d_{i,t+h}=0)}], \tag{2.13}$$

where r_t is the riskfree rate between t and $t + 1$, equal to $-\log E[m_{t,t+1}|\Omega_t]$. It is easily seen, using assumption A_0 and conditioning first on Ω_t^* in the previous expression, that r_t

11

is function of F_t only.

Proposition 5: If $\delta_S = 0$, we have:

$$B(t,h) = E_t^Q[\exp(-r_t - \ldots - r_{t+h-1} - \lambda_{t+1}^Q - \ldots - \lambda_{t+h}^Q)], \text{ with } \lambda_t^Q = \lambda_t. \qquad (2.14)$$

If $\delta_S \neq 0$, the previous formula is no longer valid. It could be replaced by:

$$B(t,h) = E_t^Q[\exp(-r_t - \ldots - r_{t+h-1} - \tilde{\lambda}_{t+1,t+h}^Q - \ldots - \tilde{\lambda}_{t+h,t+h}^Q)], \qquad (2.15)$$

where $\tilde{\lambda}_{t+k,t+h}^Q$ is defined by: $Q(d_{t+k} = 0 | d_{t+k-1} = 0, \underline{F_{t+h}}) = \exp(-\tilde{\lambda}_{t+k,t+h}^Q)$. $\tilde{\lambda}_{t+k,t+h}^Q$ is doubly indexed and function of $\underline{F_{t+h}}$, and thus is not a default intensity process. It can be seen as a *"forward" default intensity.*

Proof: We have:

$$
\begin{aligned}
B(t,h) &= E^Q[\exp(-r_t - \ldots - r_{t+h-1})\mathbb{I}_{(d_{i,t+h}=0)} | d_{i,t} = 0, \Omega_t] \\
&= E^Q[E^Q[\exp(-r_t - \ldots - r_{t+h-1})\mathbb{I}_{(d_{i,t+h}=0)} | d_{i,t} = 0, \Omega_t, \underline{F_{t+h}}] | d_{i,t} = 0, \Omega_t] \\
&= E^Q[\exp(-r_t - \ldots - r_{t+h-1})E^Q[\mathbb{I}_{(d_{i,t+h}=0)} | d_{i,t} = 0, \Omega_t, \underline{F_{t+h}}] | d_{i,t} = 0, \Omega_t] \\
&\quad \text{(using the fact that } r_t \text{ is function of } F_t) \\
&= E^Q[\exp(-r_t - \ldots - r_{t+h-1}) \\
&\quad \times E^Q\{\Pi_{k=1}^h Q[\tau_i > t + k | \tau_i > t + k - 1, \underline{F_{t+h}}] | \tau_i > t, \Omega_t\}] | d_{i,t} = 0, \Omega_t].
\end{aligned}
$$

If $\delta_S = 0$, the factor process (F_t) remains exogenous in the risk-neutral world and $Q[\tau_i > t + k | \tau_i > t + k - 1, \underline{F_{t+h}}]$ can be replaced by $Q[\tau_i > t + k | \tau_i > t + k - 1, \underline{F_{t+k}}]$, which is also equal to $\exp(-\lambda_{t+k}^Q)$ and to $\exp(-\lambda_{t+k})$ (Proposition 2 iii)). If $\delta_S \neq 0$, the expression $Q[\tau_i > t + k | \tau_i > t + k - 1, \underline{F_{t+h}}]$ is equal to $\exp(-\tilde{\lambda}_{t+k,t+h}^Q)$ and the result follows.$\square$

Proposition 4 shows that a default pre-intensity can exist in the historical world without existing in the risk-neutral world. Besides, Proposition 5 shows that by assuming the

existence of a default pre-intensity in the risk-neutral world we implicitly do not price the surprise events. It has already been shown in the literature that Equation (2.13) is typically not equal to Equation (2.14) [See e.g. Duffie, Schroeder, Skiadas (1996), Proposition 1] and that it may arise when there are contagious defaults under Q [See e.g. Bai, Collin-Dufresne, Goldstein, Helwege (2012)a]. Our results show that this is due to the non-exogeneity of the driving factors under Q, which can arise even if these factors are exogenous under P.

If we replace assumption A_0 by assumption A_0^*, that is if F_t is no longer exogenous, a default pre-intensity exists neither in the historical world, nor in the risk-neutral world.

3   Homogeneous Pool

3.1   The dynamic Poisson model

Let us illustrate the discussion above by considering a Poisson regression model for the default counts with the exogenous factors as explanatory variables [see Cameron, Trivedi (1989) for Poisson regression models]. Moreover let us consider factors which follow a compound autoregressive (CaR) dynamic [Darolles, Gouriéroux, Jasiak (2006)].

Assumption A.1: i) The conditional distribution of n_{t+1} given $(F_t), \underline{n_t}$ is Poisson $\mathcal{P}(\beta' F_{t+1} + \gamma)$. **ii)** The conditional Laplace transform of F_{t+1} given $\underline{F_t}$ and $\underline{n_t}$ is exponential affine in F_t:

$$E_t[\exp(v'F_{t+1})] = E[\exp(v'F_{t+1})|\underline{F_t}, \underline{n_t}] \equiv \exp[A(1,v)'F_t + B(1,v)], \text{ for any } v \in \mathcal{V},$$

where $\mathcal{V}$ is the set of arguments v for which the Laplace transform exists and functions $A(1,.)$, $B(1,.)$ characterize the dynamics of factor F.[7]

The conditional Poisson model defined in Assumption A.1 i) is the aggregate of a microeconomic model in which the individual defaults are independent conditional on the factor path, the conditional individual default probabilities are the same for all non-defaulted en-

tities and are equal to $\beta' F_{t+1} + \gamma$ divided by the number of alive entities $I - N_t$. This is the doubly stochastic model or model with stochastic intensity [see Cox (1955)] written under its macroeconomic version. In this respect the model extends the model considered in Collin-Dufresne, Goldstein, Helwege (2010), in which the common frailty $F_t \equiv S$ is assumed time independent. As seen below the introduction of a dynamic frailty is needed to get an appropriate dynamic treatment of the information available to investors. Indeed, even if the dynamic frailty is observed up to time t by the investor, the investor will not know perfectly its future values; this creates a dependence between the future individual defaults, jump in the default intensities when a default occurs, and this dependence changes with the prediction horizon. This specification allows to manage the term structure of default dependence in a flexible way.[8]

For a CaR process, the Laplace transform of the cumulated process is also exponential affine at any prediction horizon h, and we can write:

$$E_t[\exp(v' \sum_{k=1}^{h} F_{t+k})] = \exp[A(h, v)' F_t + B(h, v)],$$

where functions $A(h, v), B(h, v)$ are defined recursively (see Appendix A.4).

Proposition 6: Under Assumption A.1, process (F_t, n_t) is jointly compound autoregressive and, for any horizon h, we can write:

$$E_t[\exp(u_F' \sum_{k=1}^{h} F_{t+k} + u_S \sum_{k=1}^{h} n_{t+k})] = \exp[a_F'(h, u_F, u_S)' F_t + b(h, u_F, u_S)],$$

where u_F, u_S are the arguments of the Laplace transform and functions a_F and b are given by:

$$a_F(h, u_F, u_S) = A[h, u_F + \beta(\exp u_S - 1)]$$
$$b(h, u_F, u_S) = B[h, u_F + \beta(\exp u_S - 1)] + h\gamma(\exp u_S - 1).$$

Proof: See Appendix A.5.

Let us now compare the different pricing formulas, when the s.d.f. is exponential affine

in both the factor and the default count:[9]

$$m_{t,t+1} = \exp(\delta_0 + \delta_F' F_{t+1} + \delta_S n_{t+1}).$$

(3.1)

The price of the payoff $\exp(uN_{t_0+h}) = \exp(u\sum_{k=1}^{h} n_{t_0+k}) \equiv N(u)_{t_0+h}$ (say) is given by:

$$
\begin{aligned}
\Pi(N(u), h) &= E_{t_0}[\Pi_{k=1}^h m_{t_0+k-1,t_0+k} \exp(u\sum_{k=1}^{h} n_{t_0+k})] \\
&= E_{t_0}\{\exp[h\delta_0 + \delta_F' \sum_{k=1}^{h} F_{t_0+k} + (\delta_S + u)\sum_{k=1}^{h} n_{t_0+k}]\} \\
&= \exp\{A[h, \delta_F + \beta(\exp(\delta_S + u) - 1]'F_{t_0} + B[h, \delta_F + \beta(\exp(\delta_S + u) - 1)] \\
&+ h[\delta_0 + \gamma[\exp(\delta_S + u) - 1]]\}.
\end{aligned}
$$

(3.2)

When the payoff is replaced by its expectation given $\underline{F_{t_0+h}}$, we get:

$$
\begin{aligned}
\tilde{N}(u)_{t_0+h} \equiv E[\exp(uN_{t_0+h})|\underline{F_{t_0+h}}] &= \Pi_{k=1}^h E[\exp(un_{t_0+k})|F_{t_0+k}] \\
&= \exp[\beta'\sum_{k=1}^{h} F_{t_0+k}(\exp u - 1) + h\gamma(\exp u - 1)].
\end{aligned}
$$

The price of this expected payoff is given by:

$$
\begin{aligned}
\Pi(\tilde{N}(u), h) &= E_{t_0}\{\Pi_{k=1}^h m_{t_0+k-1,t_0+k} E_{t_0}[\exp(uN_{t_0+h})|\underline{F_{t_0+h}}]\} \\
&= E_{t_0}\{\exp[h\delta_0 + \delta_F' \sum_{k=1}^{h} F_{t_0+k} + \delta_S \sum_{k=1}^{h} n_{t_0+k} \\
&+ \beta'\sum_{k=1}^{h} F_{t_0+k}(\exp u - 1) + h\gamma(\exp u - 1)]\} \\
&= \exp\{A[h, \delta_F + \beta(\exp u - 1) + \beta(\exp \delta_S - 1)]'F_{t_0} \\
&+ B[h, \delta_F + \beta(\exp u - 1) + \beta(\exp \delta_S - 1)] \\
&+ h[\delta_0 + \gamma(\exp u - 1) + \gamma(\exp \delta_S - 1)]\}.
\end{aligned}
$$

(3.3)

Let us finally consider how the pricing formulas are modified when the s.d.f. depends on the factor only and is given by:

$$
\begin{aligned}
\tilde{m}_{t,t+1} &= E[\exp(\delta_0 + \delta_F' F_{t+1} + \delta_S n_{t+1}) | F_{t+1}] \\
&= \exp\{\delta_0 + \gamma(\exp \delta_S - 1) + [\delta_F + \beta(\exp \delta_S - 1)]' F_{t+1}\}.
\end{aligned} \tag{3.4}
$$

We easily derive the price of $N(u)_{t_0+h} = \exp(u N_{t_0+h})$ and of its expectation given $\underline{F_{t_0+h}}$ based on this projected s.d.f. We get:

$$
\tilde{\Pi}(N(u), h) = \tilde{\Pi}(\tilde{N}(u), h) = \Pi(\tilde{N}(u), h). \tag{3.5}
$$

We deduce the following proposition:

Proposition 7: Under Assumption $A.1$, the term structures of the prices given in (3.2), (3.3) and (3.5) are exponential affine in F_t. The factor sensitivities are all based on the $A(h, .)$ function and derived by changing the argument u, and the risk sensitivity coefficients associated with the factor and default count, that are δ_F and δ_S, respectively, according to the derivative to be priced.

In particular for $u = 0$, we get the term structure of the riskfree zero-coupon prices:

$$
\begin{aligned}
B_f(t_0, h) &= \Pi(1, h) = \tilde{\Pi}(1, h) \\
&= \exp\{A[h, \delta_F + \beta(\exp \delta_S - 1)]' F_{t_0} + B[h, \delta_F + \beta(\exp \delta_S - 1)] \\
&+ h[\delta_0 + \gamma(\exp \delta_S - 1)]\}.
\end{aligned} \tag{3.6}
$$

3.2 Pricing individual and joint defaults

We have derived above the closed-form expression of the price of an exponential transformation of the default count: $\Pi(N(u), h)$, with $N(u) = \exp(uN)$. It is known that the price of any derivative written on N_{t_0+h} can be deduced from the prices of the derivatives with

exponential payoff [see Duffie, Pan, Singleton (2000)]. Let us now explain how the pricing formula (3.2) can be used to deduce a closed-form expression for the price of the joint default of K individual contracts, that is $\Pi(d_1 \dots d_K, h)$. We have the following result proved in Appendix A.6:

Lemma 1: If $N = d_1 + \dots + d_I$ and the indicator variables $d_i, i = 1, \dots, I$, are exchangeable,

$$E(d_1 \dots d_K) = \frac{E[N(N-1) \dots (N-K+1)]}{I(I-1) \dots (I-K+1)}, \text{ for } K \le I.$$

This lemma can be applied to the forward-neutral probability to get the similar relationship written in term of prices.

Corollary 2: $\Pi(d_1 \dots d_K, h) = \dfrac{\Pi[N(N-1) \dots (N-K+1), h]}{I(I-1) \dots (I-K+1)}.$

Moreover the price of $g(N) = N(N-1) \dots (N-K+1)$ can be deduced from the prices of exponential transforms of N.

Corollary 3: $\Pi(d_1 \dots d_K, h) = \dfrac{1}{I(I-1) \dots (I-K+1)} \left(\dfrac{d^K}{dv^K} \Pi[\exp(N \log(v)), h] \right)_{v=1}.$

The standard approaches for credit derivative pricing assume the existence of default intensities under the risk-neutral probability in order to derive closed-form expressions of the derivative prices and they cannot be applied when the factor process is not exogenous under Q. Corollary 3 explains how to deal with this difficulty when the pool is homogeneous with a sufficiently large size. We first derive the price of exponential functions of default counts, which admit closed-form expressions [see pricing formula (3.2)]. Then the prices of individual and joint defaults are deduced by an appropriate differentiation. In particular, the price of a CDS or of a defaultable bond is easily derived although the standard exponential affine pricing formula no longer applies.

4   Heterogenous pools

The approach of Section 3 can be extended to a heterogenous pool composed of J homogeneous segments with different risk characteristics. We will also authorize an influence of the number of past defaults on present defaults under the physical measure. This extension allows to disentangle the effect on price of the common factor and the effect of contagion. It is also appropriate for the analysis of the default correlations within and between segments, both under the historical and risk-neutral probabilities. We will also relax the noncausality assumption from the count process to the factor process under the historical distribution.

4.1   The model

Let us consider a pool which is segmented into J segments of initial size $I_j, j = 1, \ldots, J$. We denote by $d_{i,j,t}$ the default indicator at date t of the entity i belonging to segment j, $i = 1, \ldots, I_j$, $j = 1, \ldots, J$, by $n_{j,t}, j = 1, \ldots, J$, the default counts by segment and $n_t = (n_{1,t}, \ldots, n_{J,t})'$. We assume that a given entity belongs to the same segment at all dates. For instance if the entities are firms, the segment can be e.g. defined by the industrial sector, by the size or by the domicile country. The extension of the model introduced in Subection 3.1 is given below.

Assumption A.2: Model for heterogenous pools.
i) Conditional on $\Omega_t^* = (F_{t+1}, \Omega_t)$ the counts $n_{j,t+1}, j = 1, \ldots, J$ are independent with Poisson distributions: $n_{j,t+1} \sim \mathcal{P}(\beta_j' F_{t+1} + c_j' n_t + \gamma_j), j = 1, \ldots, J$.
ii) The conditional Laplace transform of F_{t+1} given Ω_t is exponential affine in F_t, n_t:

$$E_t[\exp(v' F_{t+1}) | \underline{F_t}, \underline{n_t}] = \exp[A_F(1, v)' F_t + A_S(1, v)' n_t + B(1, v)], \text{ for any } v \in \mathcal{V}.$$

Thus, according to A.2.i), the conditional distribution of future default counts depends on both a dynamic frailty component and lagged default counts, the latter variables introducing contagion effects. This approach extends the specifications with dynamic frailty

only, introduced to reproduce the observed default clustering and default dependence [see e.g. Gouriéroux, Monfort and Polimenis (2006), Duffie, Eckner, Horel, Saita (2009)], as well as the specifications with contagion only. For instance, the introduction of the lagged default counts in the conditional distribution of $n_{j,t+1}$ given in i), is in line with Lang, Stulz (1992), Jarrow, Yu (2001), Billio, Getmansky, Lo, Pellizon (2012), or with the Hawkes' (1971) specification of the mutually exciting point processes in a continuous-time framework [see e.g. Lando, Nielsen (1998), Errais, Giesecke, Goldberg (2010) for applications to credit risk]. Note that the total number of defaults N_t might also be introduced as a component of F_{t+1} (see the applications in Section 5).

Assumption A.2.ii) allows for the nonexogeneity of F_t, when $A_S(.)$ is different from zero.

The dynamic model described in Assumption A.2 is easily interpretable. Factor (F_t) represents the shocks with joint effect on the probabilities of default, whereas the matrix C with rows $c'_j, j = 1, \ldots, J$ characterizes a contagion channel. This matrix gives the segments connected by possible contagion effects, but also the direction and magnitude of the contagion [for such an interpretation, see e.g. Billio, Getmansky, Lo, Pellizon (2012) for a model with contagion only, Darolles, Gagliardini, Gouriéroux (2013) with a model including both dynamic frailty and contagion]. Note that a second contagion channel is introduced through A.2.ii) if A_S is different from 0, since the conditional distribution of $n_{j,t+1}$ given Ω_t is obtained by marginalizing $\mathcal{P}(\beta'_j F_{t+1} + c'_j n_t + \gamma_j)$ with respect to the conditional distribution of F_{t+1} given Ω_t, which also depends on n_t.

We deduce the following Proposition, which extends Proposition 6.

Proposition 8: Under Assumption A.2 the process (F_t, n_t) is jointly compound autore-

gressive. For any horizon h, we can write:

$$E_t[\exp(u_F' \sum_{k=1}^{h} F_{t+k} + u_S' \sum_{k=1}^{h} n_{t+k})]$$
$$= \exp[a_F'(h, u_F, u_S)F_t + a_S'(h, u_F, u_S)n_t + b(h, u_F, u_S)],$$

where

$$a_F(1, u_F, u_S) = A_F[1, u_F + \sum_{j=1}^{J} \beta_j(\exp u_{jS} - 1)],$$

$$a_S(1, u_F, u_S) = A_S[1, u_F + \sum_{j=1}^{J} \beta_j(\exp u_{jS} - 1)] + \sum_{j=1}^{J} c_j(\exp u_{jS} - 1),$$

$$b(1, u_F, u_S) = B[1, u_F + \sum_{j=1}^{J} \beta_j(\exp u_{jS} - 1)] + \sum_{j=1}^{J} \gamma_j(\exp u_{jS}' - 1),$$

and similar functions for other horizons h are deduced by recursion (see Appendix A.4).

Proof: See Appendix A.7.

4.2 Pricing formulas

This subsection extends Proposition 7. Let us introduce the vector $N_{t+h} = \sum_{k=1}^{h} n_{t+k}$ and consider the s.d.f. function incorporating the surprises of default events in each segment:

$$m_{t,t+1} = \exp[\delta_0 + \delta_F' F_{t+1} + \delta_S' n_{t+1}].$$

Proposition 9: Under Assumption A.2, the price at date t_0 of the exponential affine payoff $N(u)_{t_0+h} = \exp(u'N_{t_0+h})$ is given by:

$$\begin{aligned}\Pi(N(u), h) &= \exp\{a_F(h, \delta_F, \delta_S + u)'F_{t_0} + a_S(h, \delta_F, \delta_S + u)'n_{t_0} \\ &+ b(h, \delta_F, \delta_S + u) + h\delta_0]. \end{aligned} \tag{4.1}$$

Proof: See Appendix A.8.

Thus the prices of derivatives, including riskfree or defaultable zero-coupon bonds, can be derived from formula of Corollary 3 and depend on the surprise on credit events in two ways: i) by means of the risk premium components of vector δ_S, and ii) by the current default counts n_{j,t_0}, $j = 1, \ldots, J$, in the different segments. These effects can be more or less important according to the form of functions a_F, a_S and b, that is, according to the sensitivity parameters β_j and the contagion parameters c_j.

5   Illustrations

In this section, we illustrate the relevance of the models introduced in the previous sections. First, considering riskfree bonds and bonds issued by banks, we show that part of the credit-spread puzzle might be due to the omission of the pricing of default-event surprises in the standard credit-risk models. Then we analyze the propagation of the effect of default events in a model with several segments.

5.1   Surprise pricing and the credit-spread puzzle

In this subsection, we use the model of Section 3 to jointly price U.S. banking-sector bonds and Treasury bonds (T-bonds). The latter bonds are considered riskfree. The defaultable entities are U.S. BBB-rated banks and are assumed to constitute a homogeneous pool of credits.[10] The sample covers the period from February 1995 to May 2013. The sources and the preliminary treatments of data are presented in Appendix B. This exercise illustrates the relevance of the model pricing default-event surprises. In particular, we show that this feature is needed to capture the risk premia in the short-end of the term structure of spreads and therefore to extract the physical default probabilities in an appropriate way. In that sense, pricing default-event surprises might contribute to solve the credit-spread puzzle.

5.1.1 Model specification

The factor F_t is given by $[F_{1,t}, F_{1,t-1}, F_{2,t}, F_{2,t-1}]$, where processes $(F_{1,t})$ and $(F_{2,t})$ are independent autoregressive gamma (ARG) processes with parameters μ_i, ρ_i and ν_i, $i \in \{1, 2\}$ (see Appendix A.4 for the definition of an ARG process). By introducing lagged values of $F_{1,t}$ and $F_{2,t}$ in the factor F_t, we get more flexible specifications of the s.d.f. and hence of the term structure of yields and spreads.

We set $\beta = [0, 0, 1, 0]'$ and $\gamma = 0$, implying that $F_{2,t}$ is the expectation of default count n_t conditional on F_t, since in this case $n_t \sim \mathcal{P}(F_{2,t})$.

We denote by $y_{t,h}$ the yield-to-maturity of a zero-coupon riskfree bond with residual maturity h. In order to facilitate the calibration procedure, we transform observed yields-to-maturity of defaultable bond into (synthetic) Credit Default Swap (CDS) spreads. These spreads are $s_{t,h} \equiv (12/h) \times \Pi(d_1, h)/\Pi(1, h) = (12/h) \times E_t^{Q^*}(d_{1,t+h})$, where Q^* denotes the forward-neutral measure.[11] $s_{t,h}$ is such that the price of the payoff $d_{1,t+h} - (h/12) \times s_{t,h}$ at date t is zero (see Appendix B for details about the computation of $s_{t,h}$ from the data). In the following, we refer to the $y_{t,h}$'s and to the $s_{t,h}$'s as "yields" and "spreads", respectively. These yields and spreads are affine functions of factor F_t.[12] Thus, by gathering yields and spreads of different maturities in vectors Y_t and S_t, respectively, we have:

$$[\ Y_t'\quad S_t'\ ]' \equiv M F_t + m. \tag{5.1}$$

Since the moments of F_t are available in closed form (see Appendix A.4), the same is true for $[Y_t', S_t']'$. This is exploited by our calibration procedure.

5.1.2 Calibration procedure

The parameters of the model are calibrated to reproduce a set of unconditional moments derived from observed data. Nine parameters are estimated, that are: ν_1, ν_2, μ_1, μ_2, three entries of δ_F (the first one being set to one for sake of identification of $F_{1,t}$), and δ_S and δ_0.[13]

Four types of moments are used for calibration: (i) the means of yields and spreads, (ii)

the standard deviations of the same yields and spreads, (iii) the correlations between yields and spreads of the same maturity and (iv) the average default frequency. Regarding the latter, the unconditional (physical) default rate is set to 0.4%, that is the average annual default rates across Moody's-rated bank issuers [Moody's (2010), Exhibit 39]. The first three types of moments are the sample moments computed on our yield and spread dataset: we use yields and spreads of maturities 1 year, 3 years and 5 years. For T-bonds, we add the short-term 1-month rate (this maturity is not available for banks' yields). Specifically, 17 moments of types i) to iii) have to be fitted: 7 means, 7 standard deviations and 3 correlations. Weighted squared deviations between sample and model-implied moments are minimized using a numerical procedure.[14]

To show the need for pricing default-event surprises, we calibrate a baseline version of our model (M1) and an alternative version (M2), which does not allow for the pricing of default-event surprises. Only eight parameters have to be calibrated in M2 (in which $\delta_S = 0$). Calibrated parameters for both models are reported in Table 1. Panel A of Table 2 compares the sample moments with the model-implied ones. Overall, Model M1 provides a better fit of sample moments than model M2, especially at the short-end of the term structure of spreads (this will be discussed further in Subsection 5.1.3). From a time-series perspective, Panel B of Table 2 reports the ratios of the mean squared errors to the variances of yields and spreads at different maturities.[15] The pricing errors are relatively small; they are larger for shorter and longer maturities. Although the model with pricing of default events features no more factors than the alternative one (but only one additional parameter), it entails a much better fit of the data.

5.1.3   Results and interpretation

Figure 1 compares the model-implied unconditional means of riskfree yields (upper panel) and spreads (lower panel) with the sample-based ones. For both yields and spreads, the model that prices default events (model M1) provides a better adjustment to sample means.

The difference is especially significant for short-term spreads; this stems from the ability of model M1 to generate sizeable credit-risk premia for short maturities. Let us elaborate on this.

[Insert Figure 1: Sample vs. Model-Implied Averages of Yields and Spreads]

The *credit-risk premia* are defined as the differences between priced spreads and the ones that would prevail if investors were risk-neutral, that is between the expectations of $\frac{12}{h} d_{1,t}$ under the forward-neutral and the physical measures [see e.g. Berndt, Douglas, Duffie, Ferguson, Schranz (2008)]. At this stage, it is important to distinguish the *observed* credit-risk premia, that are directly deduced from a comparison of the observed spreads with the historical probability of default, from *model-based* risk premia. The latter ones, that we call implied risk premia, are calibrated from corporate bonds but depend on the selected credit-risk model.

As stated by D'Amato, Remolona (2003), "while (observed) credit spreads are often generally understood as the compensation for credit risk, it has been difficult to explain the precise relationship between (these) spreads and such (credit) risk." More precisely, the observed spreads tend to be "many times wider than what would be implied by expected default loss", but also significantly larger than implied credit-risk premia derived from standard credit-risk models. This is the so-called credit-spread puzzle [Altman (1989), D'Amato, Remolona (2003), Hull, Predescu, White (2005), Huang and Huang (2012)].

In the lower panel of Figure 1, the implied credit-risk premia (resp. the observed credit-risk premia) are the differences between solid and dashed lines (resp. between the circles and the dashed lines). Model M1 entails a flexible modeling of risk premia resulting from the facts that: (a) the risk-neutral forward default intensities and the physical default intensity are not equal[16] and (b) physical and risk-neutral default intensities depend on factor F_t, whose dynamics are not the same under the two probability measures. For model M2, point (a) does not apply (see Proposition 2). This tends to limit the size of the implied credit-risk premia for short maturities under M2, because the differences in the historical

24

and risk-neutral dynamics of factor F_t translate into implied credit-risk premia only in a progressive way with respect to maturities. This limitation of models that do not allow for the pricing of credit-event surprises explains why empirical studies based on this standard approach obtain ratios of credit-risk premia to total spreads that are strongly increasing across maturities [see e.g. Doshi, Ericsson, Jacobs, Turnbull (2012), Table 5]. Because of the inability of structural credit-risk models [à la Black, Scholes (1973), Merton (1974)], but also of basic intensity credit-risk models, to account for large credit-risk premia at the short-end of the yield curve, alternative considerations are often invoked to explain the large bond or CDS spreads at short maturities such as liquidity [Feldhütter (2012)], taxes [Elton, Gruber, Agrawal, Mann (2001) and Driessen (2005)] and incomplete accounting information [Duffie, Lando (2001)]. By accommodating sizeable risk premia also for short maturities and not for long maturities only, our framework provides an alternative solution to the credit-spread puzzle. This feature is compatible with the results in Bai, Collin-Dufresne, Goldstein, Helwege (2012)a, since our specification includes the effect of contagion within the segment under the risk-neutral measure (see the comment following Proposition 3 in Subsection 2.5).

Figure 2 illustrates a time-series implication of the inability of basic intensity credit-risk models to account for large credit-risk premia for short maturities. On this figure, we consider the one-year probability of default (PD) of a U.S. BBB-rated bank. The dashed line corresponds to the observed spread $s_{t,12}$. In each model, there is a one-to-one relationship between $s_{t,12}$ and the factor value $F_{2,t}$.[17] From the implied factor path $(F_{2,t})$, we can deduce the implied physical PDs for both models. Model-implied physical PDs are also displayed in Figure 2; they are substantially lower with model M1. Therefore, in that context, using a model that does not price default-event surprises may tend to overestimate physical probabilities of default at short horizons or, equivalently, to underestimate the credit-risk premia.

[Insert Figure 2: 1-Year Probability of default]

5.2 Recursive contagion

This subsection provides an illustration of the model for heterogenous pools introduced in Section 4. For expository purpose, we consider a simple setting, but the approach remains tractable with larger systems and more complicated exposure setups.

Six homogeneous segments are involved, each of them being constituted of 100 entities. The factor F_t is equal to $[F_{B,t}, F'_{N,t}]'$; $(F_{B,t})$ is a sequence of i.i.d. Bernoulli variables with parameter $\nu = 0.05$. The process $(F_{N,t})$, of dimension 6, keeps memory of past default counts in the different segments. Specifically, we have:

$$F_{N,j,t} = \rho F_{N,j,t-1} + n_{j,t-1}, \quad j = 1, \ldots, 6,$$

where the smoothing parameter ρ is chosen independent of the segment. If ρ is equal to one and $F_{N,t_0} = 0$, then $F_{N,t}$ gives the cumulated number of defaults between t_0 and $t-1$ in the different segments. When $0 < \rho < 1$, $F_{N,t}$ keeps track of the number of past defaults, but underweights the oldest ones. We use $\rho = 0.8$ in the numerical example presented below.

Conditional on $\Omega_t^* = (F_{t+1}, \Omega_t)$, the counts $n_{j,t+1}$, $j = 1, \ldots, 6$, follow independent Poisson distributions:

$$n_{1,t+1} \sim \mathcal{P}(0.4 \times F_{N,6,t} + F_{B,t}) \text{ and } n_{j,t+1} \sim \mathcal{P}(0.4 \times F_{N,j-1,t}) \text{ if } j > 1. \qquad (5.2)$$

This structure defines a circular network of segments where the probability of experiencing defaults in segment j depends on the number of recent defaults in segment $j - 1$ (or in segment 6 for $j = 1$).

[Insert Figure 3: Evolutions of Factors and Default Counts]

Figure 3 displays simulated trajectories of the processes (F_t) and (n_t). We initialize the simulation with $F_{N,1} = 0$. At date 5, we get the high value of factor F_B ($F_{B,5} = 1$) that generates two defaults in segment 1. This implies an increase in factor $F_{N,1,t}$, which induces

one default in segment 2 at date 6, and so on. Even in the absence of new shock on $F_{B,t}$, defaults occur again in segment 1 at date 17 because of propagation across segments till segment 6 (recall that segment 1 is exposed to segment 6, see Equation 5.2). After the 30th period, default intensities fade and the $F_{N,j,t}$'s are all back to small values. In the absence of a new shock on F_B, there is no additional default. A new default wave is triggered after the 40th period, due to a shock on F_B that translates into three defaults in segment 1 and so on.

[Insert Figure 4: CDS prices and Probabilities of Default]

Figure 4 illustrates the implications of the model with contagion in terms of forecasting and pricing. We focus on two dates ($t = 1$ and $t = 45$) and two segments (1 and 4). For each segment and date, two charts are provided:

- The upper chart presents cumulated probabilities of default: more precisely, the black solid line indicates the probabilities that entity i defaults between t and $t + h$ (where t is the current date, i.e. either 1 or 45); the grey solid line is a forward CDS price, given by $\Pi(d_{j,i}, h)/\Pi(1, h)$;[18] the dotted line is the forward CDS price computed with the s.d.f. $\tilde{m}_{t,t+1}$ that ignores the pricing of default events.

- The lower chart presents the first differences of the previous curves (with respect to horizon h). Therefore, this chart focuses on the event of a default of entity i at specific future dates $t + h$, for h between 1 and 15: the black solid line indicates the probabilities of default of entity i at date $t + h$, the grey solid line reflects the cost of insuring against a default of entity i exactly at date $t + h$ and the dotted line shows the cost that would prevail if $m_{t,t+1}$ was replaced by $\tilde{m}_{t,t+1}$.

The prices are obtained with a s.d.f. $m_{t,t+1}$ defined by Equation (3.1), with $\delta_0 = 0$, $\delta_F = [-0.1, 0, 0, 0, 0, 0]$ and $\delta_S = [0.1, 0, 0, 0, 0, 0, 0]$. As in the previous example, the spread between the grey solid line and the dotted solid line is accounted for by credit-event risk premia.

At date 1, the default probabilities for segment-1 entities in any of the next seven periods is equal 0.05% (see Panel A in Figure 3). To understand that, recall that the number of defaults in segment 1, conditional on F_t, follows a Poisson distribution $\mathcal{P}(0.4 \times F_{N,6,t} + F_{B,t})$. Therefore, we can have a default in segment 1 at date t only if either $F_{B,t} > 0$, or $F_{N,6,t} > 0$. Further, since there cannot be any default in segment j $(j > 1)$ without previous defaults in segment $j-1$, $F_{N,6,t}$ necessarily remains at zero for at least 6 periods when $F_{N,t} = 0$. In the latter case, the default probability for any entity in segment 1 at dates $t + h$ for $h < 7$ is constant and equal to 0.05%. Beyond that horizon, the probability of default increases because of possible contagion along the lines described above.

The other plots in Figure 5 show that various profiles of expected probabilities of default can be obtained in that framework. Let us look at Panel D, that corresponds to the probabilities of default of segment-4 entities in future dates $t + h$, as expected from date $t = 45$. This chart suggests that the probabilities of default are decreasing in the next 6 periods, but increase beyond that horizon. This stems from the fact that the expectation of $n_{4,t}$ conditional on future F_t is equal to $F_{N,3,t}$ and that, based on the information available at date $t = 45$, $F_{N,3,t+h}$ is expected to decrease in the next six periods. However, one default occurs in segment 4 at date 45 and this default could propagate across the different segments and generates a new wave of defaults that would take 6 periods before affecting segment 4 again. This contributes to the increase (with respect to h) in the expected probabilities of default in segment 4 beyond $t + 6$.

The propagation schemes are summarized in the left-hand-side plots of Figure 5. They provide the direction of propagation and indicate the number of defaults, when defaults occur.

This illustration shows how the model for heterogenous pool with both dynamic frailty and recursive contagion is able to reproduce stylized facts highlighted in the literature such as the increase in CDS spreads and the increase in default correlation responding to a borrower bankruptcy [see e.g. Jorion, Zhang (2009)]. This model is even more flexible since

it is able to analyze these responses in a dynamic way.

6   Concluding remarks

In order to derive closed-form expressions of interest rates and prices, standard reduced-form credit-risk models usually price the default intensity, but not the default indicator itself. In other words, the default indicator does not appear in the s.d.f. and is replaced by an appropriate prediction. Thus the prediction error, that is the surprise on default event, is neglected. This paper develops an approach that results in closed-form formula to price credit derivatives written on any number of names, without neglecting default-event surprises.

An empirical analysis based on U.S. bond data highlights the importance of pricing default-event surprises. We show that this feature adds flexibility to the modelling of credit-risk premia. Models pricing default-event surprises can generate sizeable credit-risk premia at the short end of the yield curve and, hence, can solve the credit-risk puzzle.

The analysis is extended to heterogenous pools of credits, where defaults can be driven by a dynamic frailty as well as by past default counts in the different segments. This model is appropriate for disentangling the effects of exogenous shocks from contagion effects. The illustration shows how shocks propagate in the system and the implications of this propagation on derivative prices.

Finally, note that even if traded volumes of these derivatives have decreased in the aftermath of the recent financial crisis, credit-derivative pricing is still an important topic.[19] First, this volume remains significant. Second, coherent pricing formulas are also useful from a regulating point of view, in particular to compute the required capital for financial institutions. Indeed, for rather illiquid assets, the usual mark-to-market (fair value) approach is progressively replaced by mark-to-model values. The model considered in this paper can serve this purpose.

R E F E R E N C E S

Altman, E. (1989): "Measuring Corporate Bond Mortality and Performance", Journal of Finance, 44, 902-922.

Bai, J., Collin-Dufresne, P., Goldstein, R., and J., Helwege (2012)a: "On Bounding Credit Event Risk Premia", Federal Reserve, Bank of New-York, Report 577.

Bai, J., Collin-Dufresne, P., Goldstein, R., and J., Helwege (2012)b: "Is Credit Event Risk Priced?", Columbia University DP.

Berndt, A., R. Douglas, D. Duffie, M. Ferguson, and D. Schranz (2008): "Measuring Default Risk Premia from Default Swap Rates and EDFs", Working Paper, Carnegie Mellon University.

Billio, M., Getmansky, M., Lo, A., and L., Pellizon (2012): "Econometric Measures of Correlation and Systematic Risk in the Finance and Insurance Sectors", Journal of Financial Economics, 104, 535-539.

Black, F., and Scholes, M. (1973): "The Pricing of Options and Corporate Liabilities", Journal of Political Economy, 81:3, 637-654.

Bremaud, P. (1980): "Point Processes and Queues", Springer Verlag, 350 p.

Cameron, A., and P., Trivedi (1989): "Regression Analysis of Count Data", Cambridge University Press, Cambridge.

Christensen, J.H.E., and J.A., Lopez (2012): "Common Risk Factors in the US Treasury and Corporate Bond Markets: An Arbitrage-free Dynamic Nelson-Siegel Modeling Approach", Federal Reserve Bank of San Francisco Working Paper.

Collin-Dufresne, P., Goldstein, R., and J., Helwege (2010): "Is Credit Event Risk Priced ? Modeling Contagion via the Updating of Beliefs", NBER Working Paper 15733.

Cox, D. (1955): "Some Statistical Methods Connected with Series of Events", Journal of the Royal Statistical Society, 17, 12-164.

Cox, J., Ingersoll, J., and S., Ross (1985): "A Theory of the Term Structure of Interest Rates", Econometrica, 53, 385-407.

CreditMetrics (1997): "CreditMetrics – Technical Document", J.P. Morgan, April 2, 1997.

Crédit Suisse (2007): "Credit Derivatives Handbook", Fixed Income Strategy, February 2007.

D'Amato, J., and M., Luisi (2006): "Macro Factors in the Term Structure of Credit Spreads", BIS Working Papers No 203.

D'Amato, J., and E.M., Remolona (2003): "The Credit Spread Puzzle", BIS Quarterly Review, December 2003.

Darolles, S., Gagliardini, P., and C., Gouriéroux (2013): "Survival of Hedge Funds : Frailty vs Contagion", CREST DP.

Darolles, S., Gouriéroux, C., and J., Jasiak (2006): "Structural Laplace Transform and Compound Autoregressive Models", Journal of Time Series Analysis, 27, 477-503.

Das, S., Duffie, D., Kapadia, N. and L., Saita (2007): "Common Failings: How Corporate Defaults Are Correlated", Journal of Finance, 62, 93-117.

Doshi, H., Ericsson, J., Jacobs, C., and J., Turnbull (2012): "On Pricing Credit Default Swaps With Observable Covariates", Review of Financial Studies, forthcoming.

Driessen, J. (2005): "Is Default Event Risk Priced in Corporate Bonds", Review of Financial Studies, 18:1, 165-195.

Duffee, G. (1999): "Estimating the Price of Default Risk", The Review of Financial Studies, 12, 197-226.

Duffie, D. (2005): "Credit Risk Modelling with Affine Processes", Journal of Banking and Finance, 29, 2751-2802.

Duffie, D., Eckner, A., Horel, G., and L., Saita (2009): "Frailty Correlated Default", Journal of Finance, 64, 2089-2183.

Duffie, D., Filipovic, D., and W., Schachermayer (2003): "Affine Processes and Application in Finance", Annals of Applied Probability, 13, 984-1053.

Duffie, D., and N., Gârleanu (2001): "Risk and Valuations of Collaterized Debt Obligations", Financial Analysts Journal, 51, 41-59.

Duffie, D., and D., Lando (2001): "Term Structures of Credit Spreads with Incomplete Accounting Information", Econometrica, 69:3, 633-654.

Duffie, D., Pan, J., and K., Singleton (2000): "Transform Analysis for Affine Jump Diffusions", Econometrica, 68, 1343-1376.

Duffie, D., Schroder, M., and C., Skiadas (1996): "Recursive Valuation of Defaultable Securities and the Timing of Resolution of Uncertainty", The Annals of Applied Probability, 6, 1075-1090.

Duffie, D., and K., Singleton (1999): "Modelling Term Structure of Defaultable Bonds", Review of Financial Studies, 12, 687-720.

Elton, E.J, Gruber, M.J., Agrawal, D., and C. Mann (2001): "Explaining the Rate Spread on Corporate Bonds", Journal of Finance, 56:1, 247-277.

Errais, E, Giesecke, K., and L., Goldberg (2010): "Affine Point Processes and Portfolio Credit Risk", SIAM Journal of Financial Mathematics, 1, 642-665.

Esscher, F. (1932): "On the Probability Function in the Collective Theory of Risk". Skandinavisk Aktuarietidskrift, 15, 175-195.

Feldhütter, P. (2012): "Short-term corporate bond yield spreads". London Business School Working Paper.

Feller, W. (1971): "An Introduction to Probability Theory and its Applications", Wiley, 704p.

Florens, J.P., and M., Mouchart (1982): "A Note on Noncausality", Econometrica, 50, 583-591.

Gagliardini, P. and C., Gouriéroux (2013): "Correlated Risks vs. Contagion in Stochastic Transition Models", Journal of Economic Dynamics and Control, forthcoming.

Gerber, H., and E., Shiu (1994): "Option Pricing by Esscher Transforms". Transactions of the Society of Actuaries, 46, 99-140.

Gouriéroux, C., and J., Jasiak (2006): "Autoregressive Gamma Process", Journal of Forecasting, 25, 129-152.

Gouriéroux, C., and A., Monfort (1989): "Statistics and Econometric Models", Vol1, Cambridge Univ. Press, 504p.

Gouriéroux, C., Monfort, A., and V., Polimenis (2006): "Affine Models for Credit Risk Analysis", Journal of Financial Econometrics, 4, 494-530.

Granger, C. (1980): "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods", Econometrica, 37, 424-438.

Hawkes, A. (1971): "Point Spectra of Some Self-Exciting and Mutually-Exciting Point Processes", Journal of the Royal Statistical Society, Series B, 33, 438-443.

Huang, J., and M., Huang (2012): "How Much of the Corporate-Treasury Yields Spread is Due to Credit Risk?" Review of Asset Pricing Studies, 2, 153-202.

Hull, J., Predescu, M. and A. White (2005): "Bond Prices, Default Probabilities and Risk Premiums", Journal of Credit Risk 1:2, pp. 53-60.

Jacobs, K., and X., Li (2008): "Modeling the Dynamics of Credit Spreads with Stochastic Volatility", Management Science, 54:6, 1176-1188.

Jarrow, R., Lando, D., and F., Yu (2005): "Default Risk and Diversification: Theory and Empirical Implications." Mathematical Finance, 15, 1-26.

Jarrow, R., and F., Yu (2001): "Counterparty Risk and the Pricing of Defaultable Securities", Journal of Finance, 56, 1765-1799.

Johnson, N., Kemp, A., and S., Kotz (2005): "Univariate Discrete Distributions", Wiley, 646p.

Jorion, P., and G., Zhang (2009): "Credit Contagion from Counterparty Risk", Journal of Finance, 64, 2053-2087.

Lando, D. (1998): "On Cox Processes and Credit Risky Securities", Review of Derivatives Research, 2, 99-120.

Lando, D., and M., Nielsen (2010): "Correlation in Corporate Defaults: Contagion or Conditional Independence ?", Journal of Fixed Income, 19, 355-372.

Lang, L., and R., Stulz (1992): "Contagion and Competitive Intra-Industry Effects of Bankruptcy Announcements", Journal of Financial Economics, 32, 45-60.

Merril Lynch (2006): "Credit Derivatives Handbook, Vol.1: A Guide to Single-Name and Index CDS Products". Global Securities Research & Economics Group, February 2006.

Merton, R.C. (1974): "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates", Journal of Finance, 29:2, 449:470.

Monfort A., and J.P., Renne (2013): "Default, Liquidity and Crises : an Econometric Framework", Journal of Financial Econometrics, 11:2, 221-262. forthcoming.

Sims, C. (1972): "Money, Income and Causality", American Economic Review, 62, 540-562.

Wu, L., and F.X., Zhang (2008): "A No-Arbitrage Analysis of Macroeconomic Determinants of the Credit Spread Term Structure", Management Science, 54:6, 1160-1175.

Notes

[1]Note that the population-at-risk PaR_t is measurable with respect to Ω_t.

[2]See e.g. Florens, Mouchart (1982), and Gouriéroux, Monfort (1989) Property 1.2 for a simpler proof.

[3]This specification includes the specifications with lags, i.e. with $\tilde{m}(\tilde{F}_{t+1}, \tilde{F}_t, \ldots, \tilde{F}_{t-p})$, by setting $F_{t+1} = (\tilde{F}'_{t+1}, \tilde{F}'_t, \ldots, \tilde{F}'_{t-p})'$.

[4]The results that follow (Propositions 2 to 5) remain valid when δ_0 is replaced by $-r(F_t) - \Psi_t(\delta_F, \delta_S)$, where r is a function of F_t, that defines the riskfree short-term rate, and Ψ_t denotes the conditional log-Laplace transform of (F_t, n_t).

[5]See also Monfort, Renne (2013). In a continuous-time setup, Jarrow, Lando, Yu (2005) show that both kinds of default intensities are equal when default-event risk is perfectly diversifiable.

[6]Indeed $Q(d_{i,t+1} = 1|d_{i,t} = 0, \Omega_t) = E^Q(Q(d_{i,t+1} = 1|d_{i,t} = 0, \Omega_t^*)|d_{i,t} = 0, \Omega_t) = E^Q(\exp(-\lambda_{t+1}^Q)|d_{i,t} = 0, \Omega_t)$. Since (a) λ_{t+1}^Q depends on F_{t+1} and (b) the process (N_t) Q-Granger causes the process (F_t), the latter expression is a function of both F_t and N_t.

[7]The factors are often assumed nonnegative as well as the components of β and parameter γ to ensure the positivity of the default intensity. In this case $\mathcal{V} \supset (-\infty, 0)^L$, where $L = \dim F_t$.

[8]To keep a nondenegerate default dependence, Collin-Dufresne, Goldstein, Helwege (2008) assumed that the common static frailty S is not observed by the investor. However, when time goes on, the investor updates in a Bayesian way his knowledge about S, which becomes known after a sufficiently long time. In our framework the investor cannot get the asymptotic knowledge of the future frailty values, since the frailty receives independent shocks at any future date.

[9]The following propositions remain valid if δ_0 is replaced by $\delta_0 + \delta_1' F_t$ in Equation (3.1).

[10]Modeling the term structures of interest rates by rating class is usual in the literature [see e.g. D'Amato, Luisi (2006), Jacobs, Li (2008), Wu, Zhang (2008), or Christensen, Lopez (2012)]. The assumption that the entities of a same industrial sector, a same country and a same rating class form a homogeneous pool is for instance made in J.P. Morgan's CreditMetrics model (1997) and in the standard Basel regulation.

[11]The forward-neutral measure is such that the present value of a payoff g_{t+h} (settled at date $t+h$) is given by $\Pi_t(1, h) \times E_t^{Q^*}(g_{t+h})$. The Radon-Nikodym derivative of Q^* with respect to the historical distribution is $m_{t,t+1} \times \cdots \times m_{t+h-1,t+h}/\Pi_t(1, h)$, where $m_{t,t+1}$ is the s.d.f. between t and $t + 1$.

[12]Regarding riskfree yields, this is easily seen from Equation (3.6). For spreads, it follows from Corollary 3 that $\Pi(d_1, h) = (1/I) \times \Pi(1, h) \times (C'(0, h)F + D'(0, h))$, where the functions C and D are such that $\Pi(1, h) = \exp(C(0, h)F + D(0, h))$; therefore $s_{t,h} = (12/(Ih)) \times (C'(0, h)F_t + D'(0, h))$ is an affine function of F_t.

[13]In order to minimize the number of parameters to estimate, some parameters are fixed: ρ_1 and ρ_2 are taken equal to 0.95 (a value that is consistent with the persistence of yields and spreads) and the number of banks is set to 100. Our results are robust to these choices.

[14]The weights are presented in line ω of Table 2; qualitative results are fairly robust to the choice of these weights.

[15] Model-implied spreads are based on estimates $\hat{F}_t$ of the factor path. At each date, we assume that two linear combinations of yields or spreads are measured without errors by the model. We select the linear combinations as being those that minimize the squared measurement errors across yields and spreads for

each month; these combinations are simply given by the OLS formula: $\hat{F}_t = (MM')^{-1}M'([Y_t', S_t']' - m)$, using the notations of Eq. 5.1.

[16]see Proposition 5 and Equation (2.15) for the definition of the forward default intensity. Note that the physical forward-default intensity is equal to the physical default intensity if (F_t) does not cause n_t under P, which is the case here.

[17]In the present context, $F_{1,t}$ does not intervene in spreads. This stems from the fact that $(F_{1,t})$ and $(F_{2,t}, n_t)$ are independent under both P and Q, which implies that $\Pi(N(u), h)$ is of the form $\exp(G_{1,h}F_{1,t}) \times \exp(G_{2,h}(u)F_{2,t})$. Therefore, since $s_{t,h} = (1/I)\left(d\Pi(N(u), h)/du\right)_{u=0}/\Pi(N(0), h)$, factor $F_{1,t}$ does not appear in $s_{t,h}$.

[18]This standardization avoids the discounting effects that are implicitly present in the CDS pricing formula given in Corollary 3 where $\Pi(d_{j,i}, h)$ is paid upfront at date t, while the payoff $d_{j,i,t+h}$ is implicitly settled at date $t + h$.

[19]According the BIS (http://www.bis.org/statistics/otcder/dt21.csv), the amounts outstanding of over-the-counter traded CDS was larger than \$45tr in mid-2008. Between mid-2009 and the end of 2012, this amount lied between \$20tr and \$25tr.

Appendix A: Proofs and Technical Results

Appendix A.1: Conditional Independence

Lemma A.1: The process (n_t) admits independent components conditional on the factor process (F_t).

Proof: Let us consider the Sims' characterization of the noncausality of process (n_t). We get:

$$f(n_t|\underline{n_{t-1}}, \underline{F_T}) = f(n_t|\underline{n_{t-1}}, \underline{F_t}), \ \forall t \leq T.$$

Moreover by Assumption A_0 ii), we get:

$$f(n_t|\underline{n_{t-1}}, \underline{F_t}) = f(n_t|F_t).$$

Thus we deduce that $f(n_t|\underline{n_{t-1}}, \underline{F_T})$ does not depend on $\underline{n_{t-1}}$, which characterizes the independence of $n_1, \ldots, n_T$ given $\underline{F_T}$. $\square$

The same approach can be followed to prove the conditional independence of the individual point processes $d_i = (d_{i,t}), i = 1, \ldots, I$, conditional on the factor process.

Let us now consider the projected s.d.f. We get:

$$
\begin{aligned}
& E[\Pi_{k=1}^h m(F_{t_0+k}, n_{t_0+k})|\underline{F_{t_0+h}}] \\
= \ & \Pi_{k=1}^h E[m(F_{t_0+k}, n_{t_0+k})|\underline{F_{t_0+h}}] \text{ (using the conditional independence)} \\
= \ & \Pi_{k=1}^h E[m(F_{t_0+k}, n_{t_0+k})|\underline{F_{t_0+k}}] \text{ (using the noncausality from } (n_t) \text{ to } (F_t)).
\end{aligned}
$$

Appendix A.2: Default intensity and intensity of a point process in discrete time

Let us consider the point process $(d_{i,t})$. We have:

$$
\begin{aligned}
E[d_{i,t+h} - d_{i,t}|\Omega_t^*] & = \sum_{k=1}^h E(d_{i,t+k} - d_{i,t+k-1}|\Omega_t^*) \\
& = \sum_{k=1}^h E\left[E(d_{i,t+k} - d_{i,t+k-1}|\Omega_{t+k-1}^*)|\Omega_t^*\right] \\
& = E\left[\sum_{k=1}^h \lambda_{t+k}^*|\Omega_t^*\right],
\end{aligned}
$$

with $\lambda_{t+k}^* = E(d_{i,t+k} - d_{i,t+k-1}|\Omega_{t+k-1}^*)$. The process $d_{i,t} - \Sigma_{\tau=0}^t \lambda_\tau^*$ is a Ω_t^*-martingale since $E[d_{i,t+h} - \Sigma_{\tau=0}^{t+h} \lambda_\tau^*|\Omega_t^*] = d_{i,t} - \Sigma_{\tau=0}^t \lambda_\tau^*$. Thus, by definition, λ_t^* is the intensity of $d_{i,t}$ [see Bremaud (1980), Section 2]. This intensity is linked to the default intensity introduced in Definition 1 by:

$$\lambda_{t+1}^* = (1 - d_{i,t})[1 - \exp(-\lambda_{t+1})].$$

Appendix A.3: Proofs of Propositions 2 and 3

i) Proof of Proposition 2

We have:

$$\begin{aligned}
f^P(F_{t+1}, d_{t+1}|\Omega_t) &= f^P(F_{t+1}|\Omega_t) f^P(d_{t+1}|F_{t+1}, \Omega_t) \\
&= f^P(F_{t+1}|\Omega_t) \prod_{i \in PaR_t} f^P(d_{i,t+1}|F_{t+1}, d_{i,t} = 0) \prod_{i \notin PaR_t} d_{i,t+1},
\end{aligned}$$

by the independence of the individual point processes given (F_t).

The risk-neutral conditional p.d.f of (F_{t+1}, d_{t+1}) given Ω_t is proportional to:

$$f^P(F_{t+1}|\Omega_t) \exp(\delta_F' F_{t+1}) \prod_{i \in PaR_t} \left[f^P(d_{i,t+1}|F_{t+1}, d_{i,t} = 0) \exp(\delta_S d_{i,t+1}) \right] \prod_{i \notin PaR_t} d_{i,t+1}.$$

Therefore, we have:

$$f^Q(d_{t+1}|F_{t+1}, \Omega_t) \propto \prod_{i \in PaR_t} \left[f^P(d_{i,t+1}|F_{t+1}, d_{i,t} = 0) \exp(\delta_S d_{i,t+1}) \right] \prod_{i \notin PaR_t} d_{i,t+1},$$

where the proportionality coefficient depends on the conditioning variables. This implies that under the risk-neutral probability:

i) the point processes $(d_{i,t})$, $i = 1, \ldots, I$ are independent conditional on (F_t), due to the multiplicative decomposition of the joint density;

ii) state 1 is still absorbing;

iii) the risk-neutral p.d.f. $f^Q(d_{i,t+1}|F_{t+1}, d_{i,t} = 0)$ is the Esscher transform [Esscher (1932), Gerber, Shin (1994)] of the historical p.d.f. $f^P(d_{i,t+1}|F_{t+1}, d_{i,t} = 0)$ associated with parameter δ_S.

In particular $f^Q(d_{i,t+1} = 0|F_{t+1}, d_{i,t} = 0) \equiv \exp(-\lambda_{t+1}^Q)$ is proportional to $f^P(d_{i,t+1} = 0|F_{t+1}, d_{i,t} = 0) \equiv \exp(-\lambda_{t+1})$ and $f^Q(d_{i,t+1} = 1|F_{t+1}, d_{i,t} = 0)$ is proportional to $[1 - \exp(-\lambda_{t+1})] \exp(\delta_S)$.

Therefore we get:

$$\exp(-\lambda_{t+1}^Q) = \frac{\exp(-\lambda_{t+1})}{\exp(-\lambda_{t+1}) + [1 - \exp(-\lambda_{t+1})]\exp(\delta_S)}$$

and Proposition 2 follows.$\square$

ii) Proof of Proposition 3

The risk-neutral conditional p.d.f of F_{t+1} given Ω_t is obtained by summing on d_{t+1} the joint risk-neutral conditional p.d.f. of (F_{t+1}, d_{t+1}) given Ω_t, therefore the p.d.f. is proportional to:

$$f^P(F_{t+1}|\Omega_t)\exp(\delta_F' F_{t+1})[\sum_{d_{i,t+1}=0}^{1} f^P(d_{i,t+1}|F_{t+1}, d_{i,t}=0)\exp(\delta_S d_{i,t+1})]^{(I-N_t)}.\square$$

Appendix A.4: Recursive Formulas for CaR Processes

We recall in this appendix the recursive formulas for computing the Laplace transform of a (multidimensional) CaR process at the different prediction horizons [see e.g. Darolles, Gouriéroux, Jasiak (2006)]. We write these formulas for a process (Y_t), which will be either $Y_t = F_t$, or $Y_t = (F_t', n_t')'$ in our applications.

Proposition A.1: For a CaR process such that:

$$E_t[\exp(u'Y_{t+1})] = \exp[a(1,u)'Y_t + b(1,u)],$$

we also have:
$$E_t[\exp(u'\sum_{k=1}^{h} Y_{t+k})] = \exp[a(h,u)'Y_t + b(h,u)],$$

where the functions $a(h,u), b(h,u)$ satisfy the recursive equations:

$$a(h,u) = a[1, u + a(h-1,u)], \tag{a.1}$$
$$b(h,u) = b[1, u + a(h-1,u)] + b(h-1,u). \tag{a.2}$$

Proof: The recursive formulas are easily derived by applying the iterated expectation

theorem. We get:

$$E_t[\exp(u' \sum_{k=1}^{h} Y_{t+k})]$$

$$= E_t\{\exp(u'Y_{t+1})E_{t+1}[\exp(u' \sum_{k=1}^{h-1} Y_{t+1+k})]]\}$$

$$= E_t\{\exp[u'Y_{t+1} + a(h-1,u)'Y_{t+1} + b(h-1,u)]\}$$

$$= \exp[a[1,u+a(h-1,u)]'Y_t + b[1,u+a(h-1,u)] + b(h-1,u)].$$

The recursive formulas of the Proposition are deduced by identification.□

The recursive formulas (a.1) - (a.2) can also be used to deduce recursively the expressions of the derivatives w.r.t. argument u. We get:

$$\frac{\partial a(h,u)}{\partial u'} = \frac{\partial a}{\partial u'}[1,u+a(h-1,u)][Id + \frac{\partial a}{\partial u'}(h-1,u)], \qquad (a.3)$$

$$\frac{\partial b(h,u)}{\partial u'} = \frac{\partial b}{\partial u'}[1,u+a(h-1,u)](Id + \frac{\partial a}{\partial u'}(h-1,u)] + \frac{\partial b}{\partial u'}(h-1,u). \qquad (a.4)$$

Example: The autoregressive gamma (ARG) process.

This process is the time-discretized Cox, Ingersoll, Ross process [Cox, Ingersoll, Ross (1985)]. Z_t follows an ARG of parameters (μ,ρ,ν) iff the conditional distribution of Z_t/ν given $\underline{Z_{t-1}}$ is $\gamma(\mu,\rho Z_{t-1}/\nu)$ where γ is the noncentral gamma distribution. The conditional Laplace transform of an ARG process of parameters (μ,ρ,ν) is:

$$E_t[\exp(uZ_{t+1})] = \exp[\frac{\rho u}{1-\nu u}Z_t - \mu \log(1-\nu u)].$$

Thus we have: $a(1,u) = \dfrac{\rho u}{1-\nu u}, b(1,u) = -\mu \log(1-\nu u)$. Besides, conditional mean and variance of Z_t are given by:

$$E_t(Z_{t+h}) = \frac{1-\rho^h}{1-\rho}\nu\mu + \rho^h Z_t; \text{ and } V_t(Z_{t+h}) = \left(\frac{1-\rho^h}{1-\rho}\nu\right)^2 \mu + 2\rho^h\frac{1-\rho^h}{1-\rho}\nu Z_t$$

Appendix A.5: Proof of Proposition 6

$$E_t[\exp(u_F' \sum_{k=1}^{h} F_{t+k} + u_S \sum_{k=1}^{h} n_{t+k})]$$

$$= E_t\{\exp(u_F' \sum_{k=1}^{h} F_{t+k}) E_t[\exp(u_S \sum_{k=1}^{h} n_{t+k}) | \underline{F_{t+h}}]\} \text{ (by iterated expectation)}$$

$$= E_t\{\exp(u_F' \sum_{k=1}^{h} F_{t+k}) \Pi_{k=1}^{h} E[\exp(u_S n_{t+k}) | F_{t+k})]\} \text{ (by Assumption A.1)}$$

$$= E_t\{\exp(u_F' \sum_{k=1}^{h} F_{t+k}) \Pi_{k=1}^{h} \exp[(\beta' F_{t+k} + \gamma)(\exp u_S - 1)]\} \text{ (by using the}$$

expression of the Laplace transform of a Poisson variable)

$$= E_t\{\exp([u_F + \beta(\exp u_S - 1)]' \sum_{k=1}^{h} F_{t+k})\} \exp|h\gamma(\exp u_S - 1)]$$

$$= \exp\{A[h, u_F + \beta(\exp u_S - 1)]' F_t + B[h, u_F + \beta(\exp u_s - 1)] + h\gamma(\exp u_S - 1)\}.$$

This is an exponential affine function of F_t. This proves that the process (F_t, n_t) is jointly affine and the expressions of a_F and b follow.$\square$

Appendix A.6: Characterization of the joint probability of defaults

Lemma A.2: If $N = d_1 + \ldots + d_I$, where the variables $d_i, i = 1, \ldots, I$ are exchangeable, we have, for $K \leq I$:

$$P[d_1 = \ldots = d_K = 1] = E(d_1 \ldots d_K) = \frac{E[N(N-1)\ldots(N-K+1)]}{I(I-1)\ldots(I-K+1)}.$$

Proof: i) Let us first consider the case of independent defaults. Then $N \sim \mathcal{B}(I, p)$, where $p = P[d_1 = 1]$. It is easily deduced from the moment generating function of the binomial distribution [see e.g. Johnson, Kemp, Kotz (2005)] that.

$$E[N(N-1)\ldots(N-K+1)] = I(I-1)\ldots(I-K+1)p^K$$
$$= I(I-1)\ldots(I-K+1)E(d_1 \ldots d_K).$$

ii) In the general framework with possible default dependence, we know by de Finetti's theorem [see e.g. Feller (1971)], that any exchangeable sequence $d_1, \ldots, d_I$ of $\{0, 1\}$ variable is such that there exists a latent variable Z, say, such that $d_1, \ldots, d_I$ are i.i.d. Bernoulli variables with probability $p(Z)$ conditional on Z.

We deduce that:

$$E[N(N-1)\dots(N-K+1)|Z] = I(I-1)\dots(I-K+1)E[d_1\dots d_K|Z],$$

and by taking the expectation of both sides:

$$E[N(N-1)\dots(N-K+1)] = I(I-1)\dots(I-K+1)E(d_1\dots d_K).\ \square$$

Appendix A.7: Proof of Proposition 8

We have:

$$E_t[\exp(u_F' F_{t+1} + u_S' n_{t+1})]$$
$$= E_t[\exp(u_F' F_{t+1})\exp(\sum_{j=1}^{J} u_{jS} n_{j,t+1})]$$
$$= E_t\{\exp[u_F' F_{t+1} + \sum_{j=1}^{J}[\beta_j' F_{t+1} + c_j' n_t + \gamma_j][\exp(u_{jS}) - 1]]\}$$
$$= E_t \exp\{[u_F + \sum_{j=1}^{J}\beta_j(\exp u_{jS} - 1)]' F_{t+1}\}\exp[\sum_{j=1}^{J} c_j' n_t(\exp u_{jS} - 1)]$$
$$\exp(\sum_{j=1}^{J}\gamma_j[\exp(u_{jS}) - 1]).$$

Therefore:

$$E_t[\exp(u_F' F_{t+1} + u_S' n_{t+1})]$$
$$= \exp\{A_F'[1, u_F + \sum_{j=1}^{J}\beta_j(\exp u_{jS} - 1)]F_t + A_S'[1, u_F + \sum_{j=1}^{J}\beta_j(\exp u_{jS} - 1)]n_t$$
$$+ \sum_{j=1}^{J} c_j'(\exp u_{jS} - 1)n_t + B[1, u_t + \sum_{j=1}^{J}\beta_j \exp(u_{jS} - 1)]$$
$$+ \sum_{j=1}^{J}\gamma_j[\exp(u_{jS}) - 1]\}.$$

The result follows by identification.$\square$

Appendix A.8: Proof of Proposition 9

We have:

$$
\begin{aligned}
\Pi(N(u), h) &= E_{t_0}[\Pi_{k=1}^{h} m_{t_0+k-1, t_0+k} \exp(u' N_{t_0+h})] \\
&= E_{t_0}\{\exp[h\delta_0 + \delta'_F \sum_{k=1}^{h} F_{t_0+k} + (\delta_S + u)' \sum_{k=1}^{h} n_{t_0+k}]\} \\
&= \exp\{a_F(h, \delta_F, \delta_S + u)' F_{t_0} + a_S(h, \delta_F, \delta_S + u)' n_{t_0} \\
&\quad + b(h, \delta_F, \delta_S + u) + h\delta_0\}. \square
\end{aligned}
$$

Appendix B: The dataset

Bank and Treasury yields-to-maturity are end-of-month data extracted from Bloomberg (Tickers C070 for bank yields and C082 for Treasury yields). Bank yields are generic yields computed as averages of yields-to-maturity of bonds issued by BBB-rated banks. After extraction, bootstrap techniques are applied on these coupon-bond yields so as to get zero-coupon yields. The defaultable-bond yields have to be corrected for non-zero recovery rates. Assuming that potential recovery payments take place at maturity, and denoting by R the recovery payment, one can reconstruct a riskfree bond with a defaultable bond issued by debtor i and a CDS written on the same entity:

$$
\Pi(1, h) = \exp(-h \times y_{i,h}^{NZ}) + (1 - R)\Pi(d_i, h),
$$

where $y_{i,h}^{NZ}$ is the yield-to-maturity of a defaultable zero-coupon bond issued by entity i with a residual maturity of h and R is the recovery rate. Market prices of $\Pi(d_i, h)$ are then obtained using the previous formula. For that, we fix the recovery rate at 40%, consistent with the industry practice [see e.g. Merril Lynch (2006) or Crédit Suisse (2007)].

Finally, $s_{t,h}$ is defined through $\frac{h}{12} s_{t,h} \Pi_t(1, h) = \Pi_t(d_1, h)$; it is such that the credit swap whose payoff at date $t + h$ is $d_{i,t+h} - h/12 s_{t,h}$ is worth zero at date t. The reason why we consider annualized value (introducing the multiplicative factor $12/h$) is to make $s_{t,h}$ commensurate with usual CDS or bond spreads, that are expressed in annualized terms.

Figure 1: **Sample vs. Model-Implied Averages of Yields and Spreads**
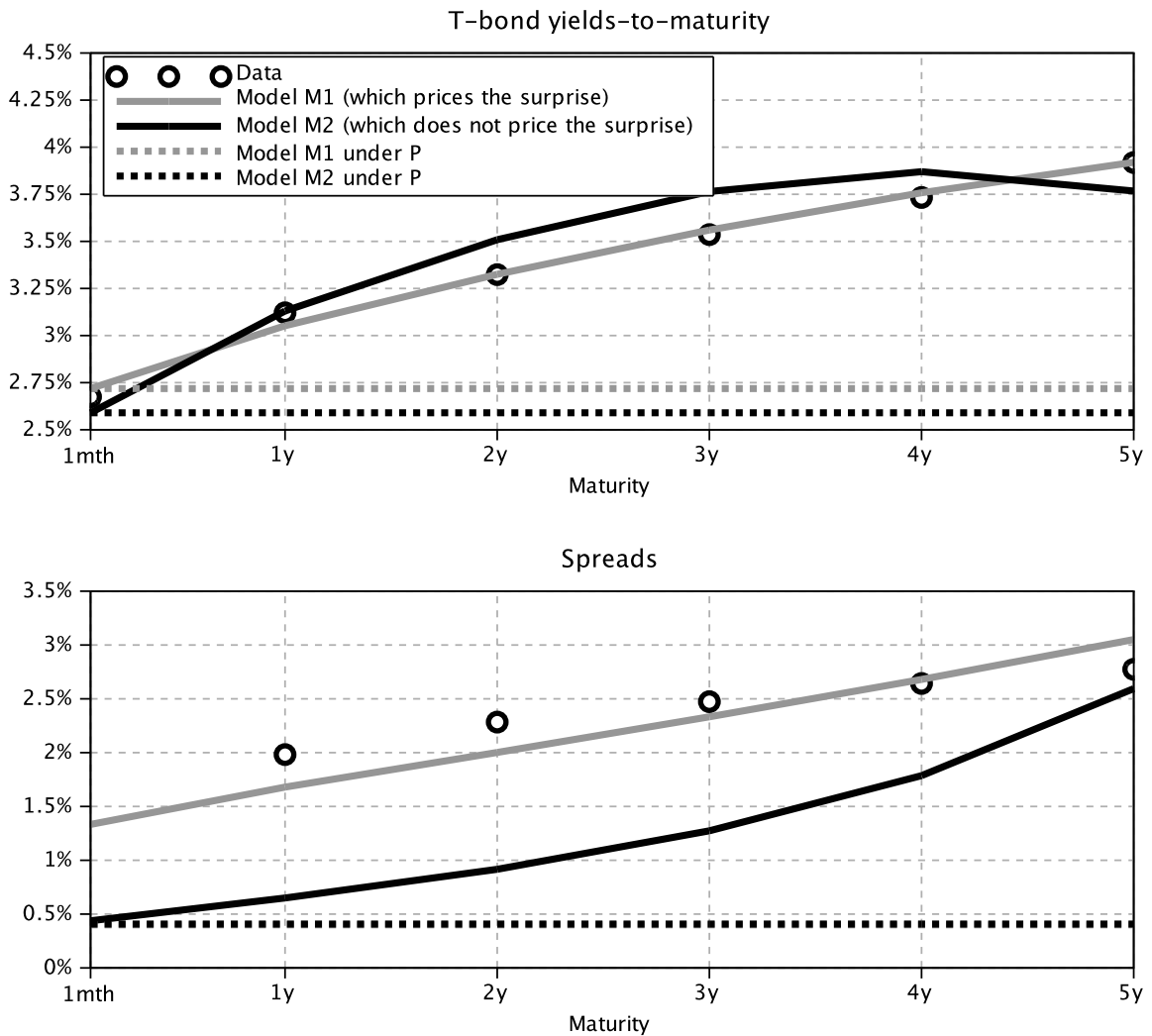
Bond issued by the U.S. Treasury are considered as riskfree. Circles correspond to sample means of Treasury yields (upper plot) or spreads (lower plot). Solid lines are under Q, dashed lines are under P. On the lower chart, the grey and black dashed lines are confounded since both models are consistent with the same unconditional default frequency (of 0.4%).

Figure 2: **1-year Probability of Default**
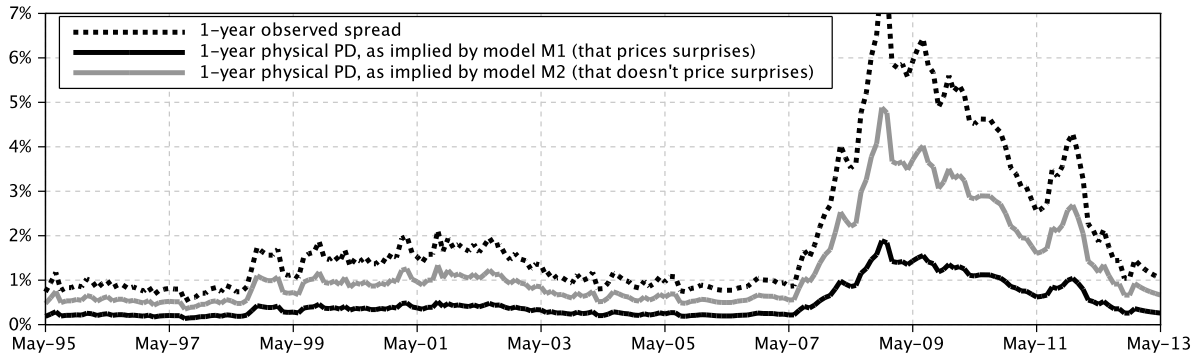
The dashed line shows the time series of the observed spread over a one-year horizon; formally, for date t, it shows $E^{Q^*}(d_{i,t+1year} = 1 | d_{i,t} = 0, \Omega_t) = \Pi(d_1, 1year)/\Pi(1, 1year)$. Solid lines show implied physical probabilities of default.

Figure 3: **Heterogenous Pool: Simulated Paths of Factors and Default Counts**
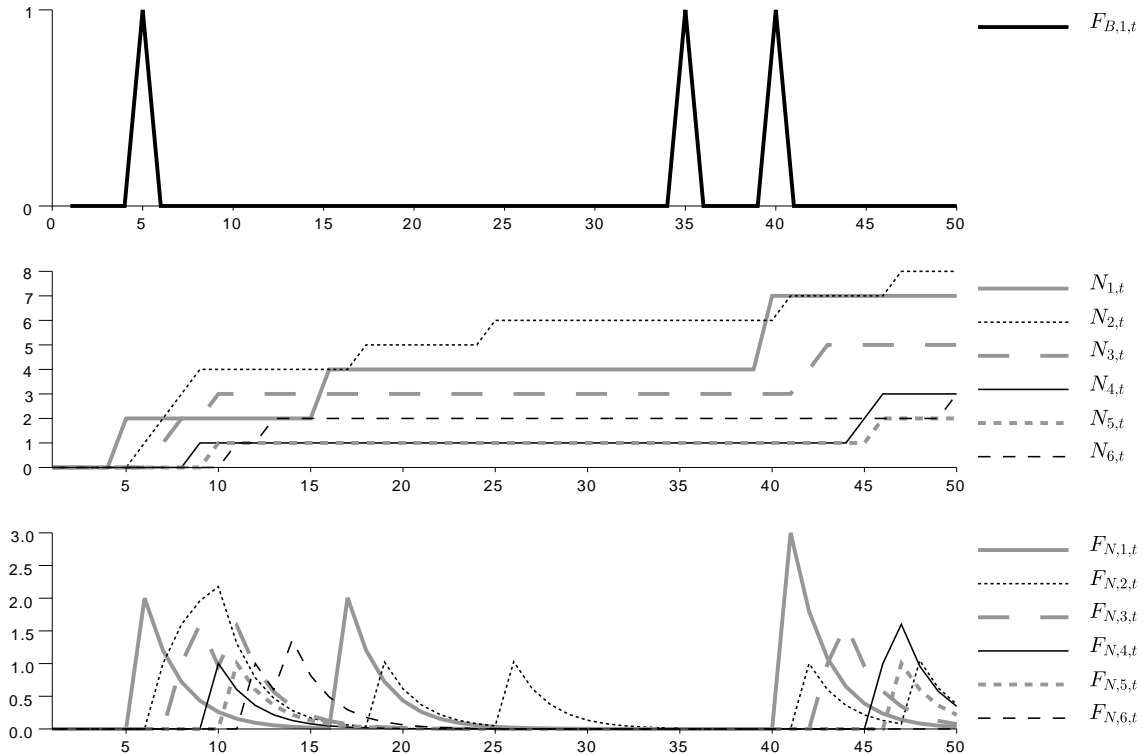
$F_{B,t}$ is drawn from a Bernoulli distribution with parameter $\nu = 5\%$. Conditional on $\Omega_t^* = (F_{t+1}, \Omega_t)$, $n_{1,t+1} \sim \mathcal{P}(0.4 \times F_{N,6,t} + F_{B,t})$ and, for $i > 1$, $n_{i,t+1} \sim \mathcal{P}(0.4 \times F_{N,i-1,t})$. In addition, $F_{N,i,t} = 0.8 \times F_{N,i,t-1} + n_{i,t-1}$. A high value of factor F_B may immediately generate defaults in segment 1, and these defaults propagate to the other segments by contagion.

43

Figure 4: **Heterogenous Pool: CDS Prices and Probabilities of Default**
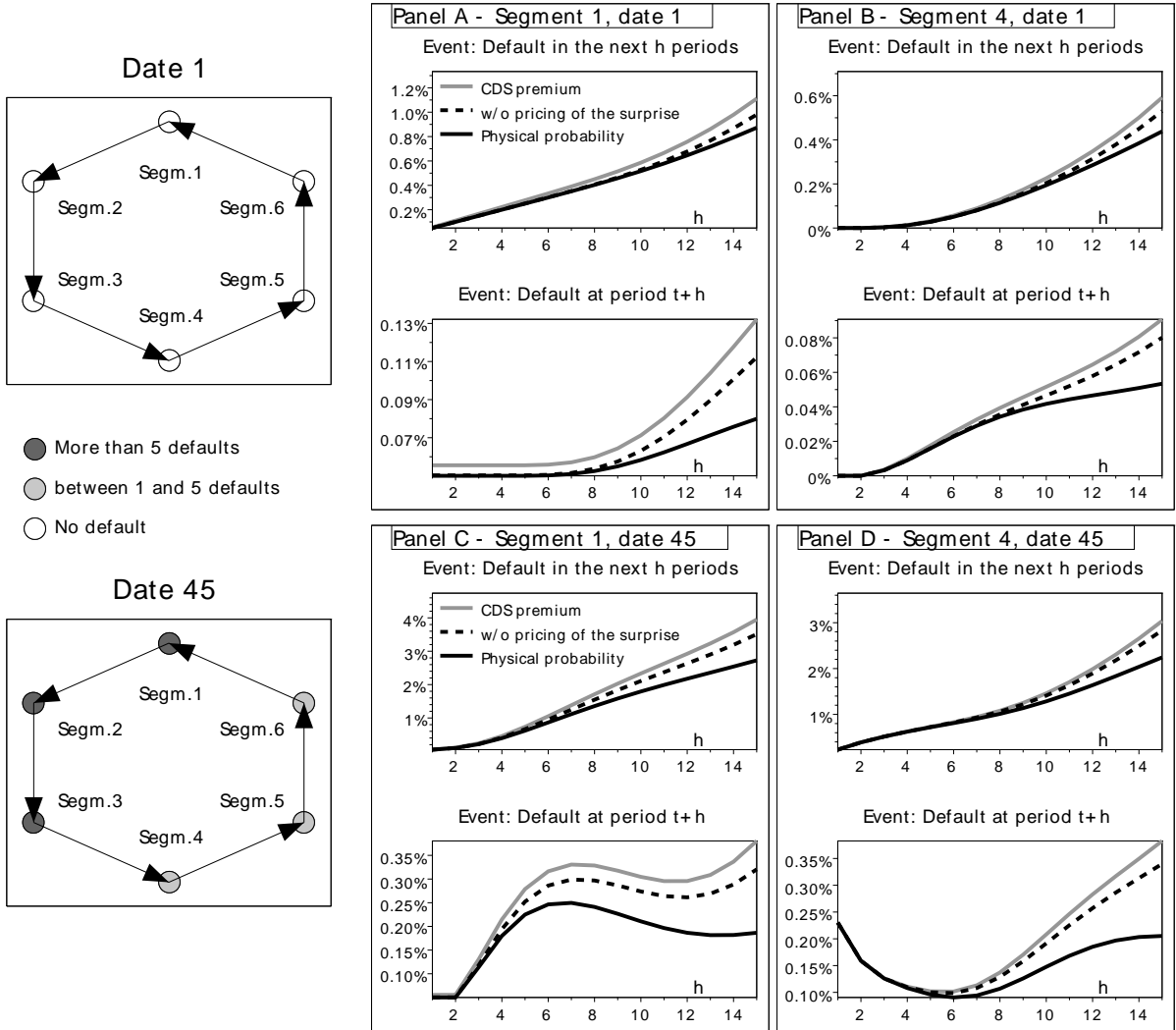
In the upper charts of Panel A to D, the black solid line indicates the probabilities that entity i defaults between t and $t+h$ (where t is either 1 or 45); the grey solid line plots the forward prices of CDS (that are the same probabilities under the forward-neutral measure) and the dotted line corresponds to the forward price of CDS computed with the s.d.f. $\tilde{m}_{t,t+1}$ that ignores the pricing of default events. The lower charts in Panel A to D display the first differences of the curves plotted on the upper chart and is related to the event of a default of entity i at specific future dates $t+h$.

Table 1: Model calibrations

	μ_1	ν_1	ρ_1	μ_2	ν_2	ρ_2	$\delta_{F,1}$	$\delta_{F,2}$	$\delta_{F,3}$	$\delta_{F,4}$	δ_S	δ_0
M1	1.55	0.022	0.95	0.428	0.004	0.95	1	-0.974	3.045	-5.063	1.163	-0.044
M2	3.26	0.021	0.95	0.267	0.006	0.95	1	-0.972	5.681	-5.589	-	-0.081

M1 (resp. M2) is the model pricing the default-event surprise, i.e. with $\delta_S \neq 0$ (resp. $\delta_S = 0$). $F_{1,t}$ and $F_{2,t}$ follow independent autoregressive gamma processes of respective parameters (μ_1, ρ_1, ν_1) and (μ_2, ρ_2, ν_2); the sdf is given by $m_{t,t+1} = \exp(\delta_0 + \delta_F' F_{t+1} + \delta_S n_{t+1})$ where $F_t = [F_{1,t}, F_{1,t-1}, F_{2,t}, F_{2,t-1}]'$ and the conditional distribution of n_t given $\underline{F_t}, \underline{n_{t-1}}$ is Poisson $\mathcal{P}(F_{2,t})$.

Table 2: Fitting properties of the model

	Panel A - Unconditional moments									
	Treasuries (riskfree) yields				Spreads (Banks vs. Treas.)			Correlations		
	1 mth	1y	3y	5y	1y	3y	5y	1y	3y	5y
ω	50	50	50	50	100	100	100	0.05	0.05	0.05
S	2.7/2.1	3.1/2.2	3.5/2.0	3.9/1.8	2.0/1.6	2.5/1.8	2.8/2.0	-60	-70	-65
M1	2.7/2.2	3.1/2.1	3.6/1.9	3.9/1.8	1.7/1.9	2.3/1.8	3.1/1.8	-65	-65	-65
M2	2.6/1.7	3.1/1.7	3.8/1.8	3.8/2.3	0.6/0.9	1.3/1.3	2.6/2.4	-47	-54	-70

	Panel B - Time-series fit (MSE divided by series variances, in %)						
	Treasuries (riskfree) yields				Spreads (Banks vs. Treas.)		
	1 mth	1y	3y	5y	1y	3y	5y
M1	8.6	2.8	0.2	3.1	11.1	1.2	7.2
M2	16.3	9.2	1.0	36.3	57.8	19.5	24.2

Panel A reports sample moments (row S) as well as the unconditional moments implied by the two calibrated models (rows M1 and M2); M1 (resp. M2) is the model pricing (resp. not pricing) default-event surprises. Rows S, M1 and M2 show the *means / standard deviations* of yields and spreads. These models are estimated by weighted-moment methods with weights provided in row ω (while imposing an unconditional monthly default rate of 0.4%/12). Panel B reports the ratios of mean squared pricing errors (MSE) to the sample variances of corresponding yields/spreads. The computation of pricing errors is based on estimates of the factor path (F_t) (see Footnote 15).