# INTEGRATION OF A REGIME SWITCHING MODEL FOR INTEREST RATES WITH RANDOMIZED REGIME PARAMETERS

UCONN ANALYSIS & PROBABILITY SEMINAR

MARCH 23, 2007

(with some refinements and corrections March 30, 2007)

James Bridgeman

**OUTLINE**

- **THE MODEL PROCESS**
- **BACKGROUND**
  - **The application it comes from**
  - **Integration of a simpler model**
  - **Attempted integration of the full model**
  - **What's needed to proceed**
- **RESULTS**
  - **The set up**
  - **Some lemmata**
  - **Main result**
  - **Theorem that gets you there**
  - **Comments and proof of the main result**
  - **Dissection of the theorem**
  - **Proof of the theorem**

**THE MODEL PROCESS**

- **The discrete time-step version of the process (for Monte Carlo applications) is**

$$
\begin{aligned}
d\ln(r_t) \quad &= \quad \left[1-(1-F)^{dt}\right]\left[\sum_{j=0}^{\infty}\mathbf{1}_{[\mathbf{j},\mathbf{j+1})}(t)\ln(\mathbf{T}_j)-\ln(r_{t-dt})\right] \\
&\quad + (1-F)^{dt}D_t dt + (1-F)^{dt}\sigma\sqrt{dt}\mathbf{N}_t
\end{aligned}
$$

**Where**

$r_t$ **is the interest rate we want to model over time**

$F$ **is an annualized mean reversion factor between** $0$ **and** $1$

$dt$ **is a discrete time interval**

$\mathbf{1}_{[\mathbf{j},\mathbf{j+1})}(t)$ **is the indicator for** $t$ to be in a random interval $[\mathbf{t_j},\mathbf{t_{j+1}})$

$\{\mathbf{t_{j+1}}-\mathbf{t_j}\}_{1\leq j}$ **are i.i.d. random variables with common law** $gamma(\alpha,\beta)$

$\mathbf{t_1}-\mathbf{t}_0$ **is independent of** $\{\mathbf{t_{j+1}}-\mathbf{t_j}\}_{1\leq j}$ **and distributed as a randomly chosen point within a** $gamma(\alpha,\beta)$ **interval**

$\{\ln(\mathbf{T}_j)\}_{1\leq j}$ **are i.i.d. normal random variables, independent of** $\{\mathbf{t_{j+1}}-\mathbf{t_j}\}_{0\leq j}$**, making** $\{\mathbf{T}_j\}_{1\leq j}$ **i.i.d. lognormal mean reversion targets for the interest rate and** $T_0$ **is a fixed initial value of the target**

$D_t$ **is an annualized drift-compensation function (available t.b.d. up front)**

$\sigma$ **is an annualized volatility parameter**

$\{\mathbf{N}_t\}_{0\leq t}$ **are i.i.d standard normal random variables independent of all the other random variables in the process**

**For a continuous model just think of** $dt\to 0$**, replacing** $\sqrt{dt}\mathbf{N}_t$ **with a standard Wiener process** $\mathbf{dW}_t$ **and using the Taylor expansion of** $(1-F)^{dt}$**, ignoring** $dt\mathbf{dW}_t$**,** $dt^2$ **and higher.**

$$
d\ln(r_t)=\left\{-\ln(1-F)\left[\sum_{j=0}^{\infty}\mathbf{1}_{[j,j+1)}(t)\ln(T_j)-\ln(r_t)\right]+D_t\right\}dt+\sigma\mathbf{dW}_t
$$

## WHERE THIS COMES FROM

- **Monte Carlo stress-testing of financial positions**
- **Virtual realities (a.k.a. risk-neutral models) are OK for pricing financial positions but are not acceptable for stress-testing them**
- **Unconstrained lognormal models**

$$\begin{aligned} d\ln\left(r_t\right) &= D_t dt + \sigma\sqrt{dt}\mathbf{N}_t \\ d\ln(r_t) &= D_t dt + \sigma \mathbf{dW}_t \end{aligned}$$

  - **Too much probability in the tails of the resulting (integrated) interest rate distribution**

- **Mean-reverting lognormal models**

$$\begin{aligned} d\ln\left(r_t\right) &= \left[1-(1-F)^{dt}\right]\left[\ln(\mathbf{T}_0)-\ln(r_{t-dt})\right] \\ &\quad + (1-F)^{dt} D_t dt + (1-F)^{dt}\sigma\sqrt{dt}\mathbf{N}_t \\ &\quad \textbf{(actuarial folklore, unpublished, ?}\pm\textbf{1970? )} \\ d\ln(r_t) &= \left\{-\ln(1-F)\left[\ln(T_0)-\ln(r_t)\right]+D_t\right\}dt + \sigma\mathbf{dW}_t \\ &\quad \textbf{(Black-Karasinski, 1991)} \end{aligned}$$

  - **Too little probability of extended sequential runs at the shoulders of the resulting (integrated) interest rate distribution (which bothers financial engineers, too)**
  - **Too much probability in center of the integrated distribution**
  - **Empirical fit wants slightly smaller kurtosis (not well-appreciated)**

- **Starting about 1994, unpublished, and continuing to B, 2002 and 2007, we attacked this problem by randomizing the mean reversion target, now known as regime-switching**
  - **Lognormal distribution of the mean reversion target value was the easiest choice and we made it a random draw, not a sequential process**
  - **Inter-arrival times for switching to a new tarket ought look like an event-driven process**
  - **Erlang inter-arrival times seemed logical (gamma with integer $\alpha$ parameter, i.e. sum of exponentials, i.e. waiting time for $\alpha$ events to occur) but having chosen a gamma inter-arrival model, no reason to restrict to Erlang**

**WARM UP EXERCISE**

- **A busy decision-maker could understand the ordinary mean-reverting lognormal a lot better if you could tell her**

$$\mathbb{E}\left[r_t\right] = r_0^{(1-F)^t} T_0^{\left[1-(1-F)^t\right]}$$

- **So integrate the difference equation (just keep walking backwards) to get**

$$\begin{aligned}
\ln(r_t) &= \ln(r_0)\,(1-F)^{\frac{t}{dt}dt} + \sigma\sqrt{dt}\sum_{s=1}^{\frac{t}{dt}} \mathbf{N}_{t-(s-1)dt}\,(1-F)^{sdt} \\
&\quad + \ln(T_0)\left[1-(1-F)^{dt}\right]\sum_{s=1}^{\frac{t}{dt}}(1-F)^{(s-1)dt} \quad \Longleftarrow \textbf{ notice geom. series} \\
&\quad + dt\sum_{s=1}^{\frac{t}{dt}} D_{t-(s-1)dt}\,(1-F)^{sdt} \\
&= \ln(r_0)\,(1-F)^{t} + \sigma\sqrt{dt}\sum_{s=1}^{\frac{t}{dt}} \mathbf{N}_{t-(s-1)dt}\,(1-F)^{sdt} \\
&\quad + \ln(T_0)\left[1-(1-F)^{t}\right] + dt\sum_{s=1}^{\frac{t}{dt}} D_{t-(s-1)dt}\,(1-F)^{sdt}
\end{aligned}$$

- **This is just a sum of constants and constants times standard normal random variables, so it's a normal random variable. That makes $r_t$ a lognormal random variable so we know that**

$$\mathbb{E}\left[r_t\right] = e^{\mu+\frac{1}{2}\sigma^2}$$

**where the $\mu$ and $\sigma^2$ are some mess determined by the constants in the expression for $\ln(r_t)$. If you work that mess out and set it equal to $r_0^{(1-F)^t} T_0^{\left[1-(1-F)^t\right]}$, and require that it be true for all $t$, you can arrive at what the drift compensation function $D_t$ has to be to make our busy decision-maker happy**

$$\begin{aligned}
D_t &= -\frac{1}{2}\sigma^2 \frac{(1-F)^{dt}}{1+(1-F)^{dt}}\left[1+(1-F)^{2t-dt}\right], \textbf{ or} \\
D_t &= -\frac{1}{4}\sigma^2\left[1+(1-F)^{2t}\right] \textbf{ in continuous case}
\end{aligned}$$

- **You can calculate the variance** of **$r_t$ in a closed form essentially the same way to make it possible to calibrate the model against historical data directly, with no trial and error required**

## CAN WE FIND THE COMPARABLE RESULT FOR OUR REGIME-SWITCHING MODEL?

**(Or can we at least integrate it to get some insight?)**

- **The same approach (just walk backwards, but there's more book-keeping this time) gives**

$$\begin{aligned}\ln(r_t) \quad = \quad & \ln(r_0)\,(1-F)^t + \sigma\sqrt{dt}\sum_{s=1}^{\frac{t}{dt}} \mathbf{N}_{t-(s-1)dt}\,(1-F)^{sdt} \\ & + \left[1-(1-F)^{dt}\right]\sum_{s=1}^{\frac{t}{dt}}\sum_{j=0}^{\infty} \mathbf{1}_{[\mathbf{j},\mathbf{j+1})}(sdt)\ln(\mathbf{T}_j)\,(1-F)^{t-sdt} \\ & + dt\sum_{s=1}^{\frac{t}{dt}} D_{t-(s-1)dt}\,(1-F)^{sdt}\,.\end{aligned}$$

**which doesn't look so promising (no easy geometric series in sight). By the way, for notational sanity this expression and the rest of this presentation ignore the set of sample paths of measure $0$ that contain a regime-switching moment of time that coincides with a model time-step.**

- **But if we change the order of summation in the middle term and do some more book-keeping we at least get some telescoping across every inter-arrival interval for a regime-switch**

$$\begin{aligned}\ln(r_t) \quad = \quad & \ln(r_0)\,(1-F)^t + \sigma\sqrt{dt}\sum_{s=1}^{\frac{t}{dt}} \mathbf{N}_{t-(s-1)dt}\,(1-F)^{sdt} \\ & + \ln(T_0)\left[(1-F)^{(t-\mathbf{t}_1)_+} - (1-F)^t\right] \\ & + \sum_{j=1}^{\infty}\ln(\mathbf{T}_j)\left[(1-F)^{(t-\mathbf{t}_{j+1})_+} - (1-F)^{(t-\mathbf{t}_j)_+}\right] \quad \Longleftarrow \textbf{after telescoping} \\ & + dt\sum_{s=1}^{\frac{t}{dt}} D_{t-(s-1)dt}\,(1-F)^{sdt}\end{aligned}$$

- **Because of those random exponents, this isn't going to give us a lognormal $r_t$; in fact, it looks like a log-log-gamma, which won't be pretty. Even the log-gamma is ugly. But at least we have a cut-off at $t$ so things ought to converge**

- **If we condition on the $\{\mathbf{t}_j\}$, aha!, we keep a normal distribution at least that far (notice that it's two different sets of i.i.d. normals in linear combination, $\{\mathbf{N}_t\}$ and $\{\ln(\mathbf{T}_j)\}$, but that's still just a normal random variable with a messy $\mu$ and $\sigma$)**

**How to Proceed?**

- **Tails of ln($r_t$) are supressed in favor of shoulders in our model. This suggests that $\mathbb{E}[r_t]$ might be approximated efficiently by an Edgeworth expansion for ln($r_t$). Multiply a normal density by the Taylor expansion of (the ratio of the Fourier transform of the ln($r_t$) density function to the normal density). Transform back to a sum of Hermite polynomials times constants times the normal density. $\mathbb{E}[r_t]$ will be a normal expectation of some exponential times a messy polynomial. Complete the square in the exponents (just like when you are calculating the mean of the lognormal) and the Hermite polynomials lead to massive simplifications downstairs:**

$$\mathbb{E}[r_t] \approx e^{\mu+\frac{1}{2}\sigma^2}\left\{1+\frac{1}{4!}\left[\mu_4-3\sigma^4\right]+\frac{1}{6!}\left[\mu_6-15\sigma^6\right]+...\right\}$$

**where $\sigma^2$, $\mu_4$, $\mu_6$, etc. are central moments of ln($r_t$) and $\mu$ is its mean (all hopefully calculable by conditioning on the $\{\mathbf{t}_j\}$, a mess, but hopefully calculable because ln($r_t$) is normal, conditional on the $\{\mathbf{t}_j\}$). The odd higher central moments are zero because, conditional on the $\{\mathbf{t}_j\}$, the distribution of ln($r_t$) is normal, hence symmetrical, making the conditional odd central moments all zero, hence, so are the unconditioned ones. A similar approach can give the variance or, indeed, higher moments of $r_t$**

- **The mean $\mu$ that we need for this, remembering that conditional on the $\{\mathbf{t}_j\}$ we are looking at sums of independent normals, is**

$$\begin{aligned}
\mathbb{E}[\ln(r_t)] &= \ln(r_0)(1-F)^t+\ln(T_0)\left\{\mathbb{E}\left[(1-F)^{(t-\mathbf{t}_1)_+}\right]-(1-F)^t\right\} \\
&\quad +\mu_T\mathbb{E}\left[\sum_{j=1}^{\infty}\left[(1-F)^{(t-\mathbf{t}_{j+1})_+}-(1-F)^{(t-\mathbf{t}_j)_+}\right]\right] \\
&\quad +dt\sum_{s=1}^{\frac{t}{dt}}D_{t-(s-1)dt}(1-F)^{sdt} \\
&= \ln(r_0)(1-F)^t+\ln(T_0)\left\{\mathbb{E}\left[(1-F)^{(t-\mathbf{t}_1)_+}\right]-(1-F)^t\right\} \\
&\quad \mu_T\left\{1-\mathbb{E}\left[(1-F)^{(t-\mathbf{t}_1)_+}\right]\right\} \Longleftarrow \textbf{telescoped} \\
&\quad +dt\sum_{s=1}^{\frac{t}{dt}}D_{t-(s-1)dt}(1-F)^{sdt}
\end{aligned}$$

**where $\mu_T$ is the common mean of the $\{\ln(\mathbf{T}_j)\}$, the telescoping comes because $\mathbf{t}_j \geq t$ for some $j$ and all thereafter, and using**

monotone convergence. The remaining expectation values in the expression work out to a constant times some value of a Laplace transform that we can calculate, related to the random variable $\mathbf{t}_1$(we'll see that later).

- The even central moments that we need, using what we know about the higher moments of the standard normal, conditioning on the $\{\mathbf{t}_j\}$, using the notation $(2n)? = (2n-1)(2n-3)\cdots 1$, and summing a geometric series that appears, are

$$\mathbb{E}\left[\{\ln(r_t) - \mathbb{E}\left[\ln(r_t)\right]\}^{2n}\right] =$$

$$= (2n)?\mathbb{E}\left[\left\{\sigma^2 dt\,(1-F)^{2dt}\,\frac{1-(1-F)^{2t}}{1-(1-F)^{2dt}} + \sigma_T^2\sum_{j=1}^{\infty}\mathbf{e}_j^2\right\}^n\right]$$

where $\sigma_T^2$ is the common variance of the $\{\ln(\mathbf{T}_j)\}$ and for each $j$

$$\mathbf{e}_j = \left\{(1-F)^{(t-\mathbf{t}_{j+1})_+} - (1-F)^{(t-\mathbf{t}_j)_+}\right\}$$

- We can expand the $n$th power of the bracket binomially and take the expectation, but that will still leave us needing to evaluate terms of the form

$$\mathbb{E}\left[\left(\sum_{j=1}^{\infty}\mathbf{e}_j^2\right)^m\right]$$

- It can become a combinatorial nightmare but if we keep the book-keeping straight, and note that (as we'll see later) $\mathbf{e}_j$ and $\mathbf{e}_i$ have correlation $0$ for $i \neq j$, the most general expectation we'll need to be able to calculate is

$$\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^{2n}\prod_{k=1}^{n}\left(\mathbb{E}\left[\mathbf{e}_j^{2k}\right]\right)^{n_k}\right]$$

where $\{n_k\}$ are arbitrary non-negative integers.

- An expression for this last expectation in terms of the underlying gamma distribution for interarrival times is what we propose to derive. Although not needed for this application, the expression we derive will work for odd powers, too.

**THE SETUP**

**Definitions and Notation**

$$
\begin{array}{rcl}
 & & \mathbf{d}_1, \mathbf{d}_2, ..., \mathbf{d}_j, ...\textbf{are i.i.d with common law } \mathbf{d} \\
\textbf{define } 0 & \leq & \bar{\mathbf{d}}_0 \ \leq \ \mathbf{d}_1 \textbf{ with } \bar{\mathbf{d}}_0 \simeq (\mathbf{d}_1 - \bar{\mathbf{d}}_0) \simeq \bar{\mathbf{d}} \\
\textbf{where} & \simeq & \textbf{means "in law"} \\
\textbf{define } \bar{\mathbf{d}}_1 & = & (\bar{\mathbf{d}}_0 + t) \wedge \mathbf{d}_1 - \bar{\mathbf{d}}_0 \\
\textbf{and } \mathbf{t}_0 & = & 0,\ \mathbf{t}_1 = \bar{\mathbf{d}}_1\textbf{, ..., } \mathbf{t}_j = \bar{\mathbf{d}}_1 + \mathbf{d}_2 + ... + \mathbf{d}_j \\
\textbf{so } \mathbf{t}_j - \mathbf{t}_{j-1} & = & \mathbf{d}_j \textbf{ for } j > 1 \\
\mathbf{t}_1 - \mathbf{t}_0 & = & \mathbf{t}_1 = \bar{\mathbf{d}}_1 \\
\textbf{define } \mathbf{J} & = & \min\{j : \mathbf{t}_j > t\} \\
\textbf{define } \mathbf{1}_{j<\mathbf{J}} & = & 0 \textbf{ for } j \geq \mathbf{J} \\
 & = & \mathbf{1} \textbf{ for } j < \mathbf{J}\textbf{, a random variable for each } j \\
\textbf{define } \bar{\mathbf{d}}_{\mathbf{J}} & = & t - \mathbf{t}_{\mathbf{J}-1} \\
\textbf{define } \bar{\mathbf{d}}_{\mathbf{J+1}} & = & \mathbf{t}_{\mathbf{J}} - t
\end{array}
$$

**Key Lemmata**

$$
\begin{array}{rcl}
\bar{\mathbf{d}}_0 + \bar{\mathbf{d}} & \simeq & \mathbf{d} \\
\bar{\mathbf{d}}_{\mathbf{J}} + \bar{\mathbf{d}}_{\mathbf{J+1}} & \simeq & \mathbf{d} \\
\textbf{the density for } \bar{\mathbf{d}} \textbf{ is} & = & \dfrac{\mathbb{P}[\mathbf{d} \geq \mathbf{t}]}{\mathbb{E}[\mathbf{d}]} \textbf{ a.k.a. "equilibrium density"} \\
\bar{\mathbf{d}}_{\mathbf{J}} & \simeq & \bar{\mathbf{d}}_1 \ \simeq \ \bar{\mathbf{d}} \ \wedge t \\
\mathbb{E}[\mathbf{1}_{1<\mathbf{J}}] & = & \mathbb{P}\left[\bar{\mathbf{d}} < \mathbf{t}\right] \\
 & & \left\{\mathbf{J}, \bar{\mathbf{d}}_1, \mathbf{d}_2, ..., \mathbf{d}_{\mathbf{J-1}}\right\} \textbf{ are independent} \\
 & & \left\{\mathbf{J}, \mathbf{d}_2, \mathbf{d}_3, ..., \mathbf{d}_{\mathbf{J-1}}, \bar{\mathbf{d}}_{\mathbf{J}}, \bar{\mathbf{d}}_{\mathbf{J+1}}\right\} \textbf{ are independent} \\
t - \mathbf{t}_j & \in & \sigma(\bar{\mathbf{d}}_1, \mathbf{d}_2, ..., \mathbf{d}_{\mathbf{j}}) \textbf{ for all } j \\
t - \mathbf{t}_j & \in & \sigma(\mathbf{d}_{j+1}, ..., \mathbf{d}_{\mathbf{J-1}}, \bar{\mathbf{d}}_{\mathbf{J}}, \bar{\mathbf{d}}_{\mathbf{J+1}}) \textbf{ for } \mathbf{1} \leq j \leq \mathbf{J} \\
t_{\mathbf{J}-1} - \mathbf{t}_j & \in & \sigma(\mathbf{d}_{j+1}, ..., \mathbf{d}_{\mathbf{J-1}}) \textbf{ for } \mathbf{1} \leq j < \mathbf{J} \\
\mathbf{1}_{j<\mathbf{J}} & \in & \sigma(\bar{\mathbf{d}}_1, \mathbf{d}_2, ..., \mathbf{d}_j)
\end{array}
$$

**We can discuss the proofs of the non-trivial ones at the end or at another time**

**THE MAIN RESULT**

$$
\mathbb{E}\left[\sum_{j=1}^{\infty} \mathbf{e}_j^n\right] = \mathbb{E}\left[\left(1-(1-F)^{\mathbf{d}}\right)^n\right] \times
$$

$$
\left\{\mathbb{P}\left[\bar{\mathbf{d}} < \mathbf{t}\right] \frac{\mathbb{E}\left[\left(1-(1-F)^{\bar{\mathbf{d}}\wedge\mathbf{t}}\right)^n\right]}{\mathbb{E}\left[\left(1-(1-F)^{\mathbf{d}}\right)^n\right]} + \mathbb{P}\left[\bar{\mathbf{d}} \geq \mathbf{t}\right] \frac{1-\mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge\mathbf{t}\right)}\right]}{1-\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]}\right\}
$$

- **The Main Result will follow from a Theorem that doesn't have as pleasant a form**

**THE THEOREM**

$$\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^n\right] = \frac{\mathbb{E}\left[\left(1-(1-F)^{\mathbf{d}}\right)^n\right]}{1-\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]}\mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge\mathbf{t}\right)}\right]\times$$

$$\mathbb{P}\left[\bar{\mathbf{d}}<\mathbf{t}\right]\left\{1-\mathbf{K}\mathbb{E}\left[\left(\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]\right)^{\mathbf{J-1}} \mid \mathbf{J}>\mathbf{1}\right]\right\}$$
$$+\mathbb{E}\left[\left(1-(1-F)^{\bar{\mathbf{d}}\wedge\mathbf{t}}\right)^n\right]\left(1-\mathbb{E}\left[(1-F)^{n\left(t-\bar{\mathbf{d}}\wedge\mathbf{t}\right)}\right]\right)$$

**where**

$$\mathbf{K}=1-\left(\frac{1-\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]}{\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]}\right)\left(\frac{\mathbb{E}\left[\left(1-(1-F)^{\bar{\mathbf{d}}\wedge\mathbf{t}}\right)^n\right]}{\mathbb{E}\left[\left(1-(1-F)^{\mathbf{d}}\right)^n\right]}-1\right)$$

**COMMENTS AND CONNECTION TO MAIN RESULT**

- **Note that in both the Theorem and the Main Result we have something that we can calculate**

  $\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]$ **is a Laplace transform value for the random variable d;** $\mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge\mathbf{t}\right)}\right]$ **is a Laplace transform value for the random variable $\bar{\mathbf{d}}\wedge\mathbf{t}$; everything else is of the form of a constant times one of these (expand the binomials) except for** $\mathbb{E}\left[\left(\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]\right)^{\mathbf{J-1}} \mid \mathbf{J}>\mathbf{1}\right]$ **which is in terms of a value of a conditional Laplace transform for the random variable J (that last doesn't sound very promising).**

## MORE COMMENTS ON THEOREM AND CONNECTION TO MAIN RESULT

- **But notice that when $n = 1$, inserting the definition of $\mathbf{e}_j$, the left hand side of the Theorem reads**

$$\begin{aligned}
\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j\right] &= \mathbb{E}\left[\sum_{j=1}^{\infty}\left\{(1-F)^{(t-\mathbf{t}_{j+1})_+} - (1-F)^{(t-\mathbf{t}_j)_+}\right\}\right] \\
&\quad \text{which telescopes to} \\
\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j\right] &= 1-\mathbb{E}\left[(1-F)^{(t-\mathbf{t}_1)_+}\right]\text{, which by one of the key lemmata} \\
&= 1-\mathbb{E}\left[(1-F)^{t-\bar{\mathbf{d}}\wedge\mathbf{t}}\right]\text{, which can be expressed in terms} \\
&\quad \text{of one of our Laplace transform values.}
\end{aligned}$$

**The telescoping follows (using monotone convergence) from the fact that $\mathbf{t}_j \geq t$ for some $j$ and all thereafter**

- **Now set this equal to the right hand side of the Theorem when $n = 1$ and rearrange to get**

$$\mathbb{E}\left[\left(\mathbb{E}\left[(1-F)^{\mathbf{d}}\right]\right)^{\mathbf{J-1}} \mid \mathbf{J} > \mathbf{1}\right] =$$

$$= \frac{\mathbb{P}\left[\bar{\mathbf{d}} < \mathbf{t}\right] - \left(1-\mathbb{E}\left[(1-F)^{\left(t-\bar{\mathbf{d}}\wedge\mathbf{t}\right)}\right]\right)}{\mathbb{P}\left[\bar{\mathbf{d}} < \mathbf{t}\right]\mathbb{E}\left[(1-F)^{\bar{\mathbf{d}}\wedge\mathbf{t}}\right]}\mathbb{E}\left[(1-F)^{\mathbf{d}}\right]$$

- **But the distributions of J, d, and $\bar{\mathrm{d}}$ depend only on the inter-arrival structure, not at all on $(1-F)$. So choose $(1-F)^n$ instead of $(1-F)$:**

$$\mathbb{E}\left[\left(\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]\right)^{\mathbf{J-1}} \mid \mathbf{J} > \mathbf{1}\right] =$$

$$= \frac{\mathbb{P}\left[\bar{\mathbf{d}} < \mathbf{t}\right] - \left(1-\mathbb{E}\left[(1-F)^{n\left(t-\bar{\mathbf{d}}\wedge\mathbf{t}\right)}\right]\right)}{\mathbb{P}\left[\bar{\mathbf{d}} < \mathbf{t}\right]\mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge\mathbf{t}\right)}\right]}\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]$$

**which is in terms of our Laplace Transform values and a probability that we can calculate (see key lemmata)**

- **Substituting this value into the Theorem, and simplifying, produces the Main Result**

## MORE COMMENTS ON THEOREM AND CONNECTION TO MAIN RESULT

- **For d a gamma random variable, and using the equilibrium density expression for $\bar{\mathbf{d}}$, integration by parts will reduce the Laplace transforms in question and the probability $\mathbb{P}\left[\bar{\mathbf{d}} < \mathbf{t}\right]$ to closed forms involving incomplete gamma functions. Let's spare ourselves the details today.**

## DISSECTION OF THE THEOREM

- **If it started at $j = 0$, had identical terms, and never terminated you'd expect something like $\left(1 - (1-F)^{\mathbf{d}}\right)^n$ multiplied by (and independent of) an infinite geometric series of independent terms of form $\left[(1-F)^{n\mathbf{d}}\right]^j$. The expectation would sum to (montone convergence)**

$$\frac{\mathbb{E}\left[\left(1 - (1-F)^{\mathbf{d}}\right)^n\right]}{1 - \mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]}$$

- **But it terminates after a random number J of terms:**

$$\frac{\mathbb{E}\left[\left(1 - (1-F)^{\mathbf{d}}\right)^n\right]}{1 - \mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]}\left\{1 - \mathbb{E}\left[\left(\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]\right)^{\mathbf{J}}\right]\right\}$$

- **And the last term $\bar{\mathbf{d}}_{\mathbf{J}} \neq \mathbf{d}_{\mathbf{J}}$:**

$$\frac{\mathbb{E}\left[\left(1 - (1-F)^{\mathbf{d}}\right)^n\right]}{1 - \mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]}\mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge\mathbf{t}\right)}\right] \times$$

$$\mathbb{P}\left[\bar{\mathbf{d}} < \mathbf{t}\right]\left\{1 - \mathbb{E}\left[\left(\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]\right)^{\mathbf{J-1}} \mid \mathbf{J} > \mathbf{1}\right]\right\}$$
$$+\mathbb{E}\left[\left(1 - (1-F)^{\bar{\mathbf{d}}\wedge\mathbf{t}}\right)^n\right]$$

- **And it starts at $j = 1$, not $0$**

$$\frac{\mathbb{E}\left[\left(1 - (1-F)^{\mathbf{d}}\right)^n\right]}{1 - \mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]}\mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge\mathbf{t}\right)}\right] \times$$

$$\mathbb{P}\left[\bar{\mathbf{d}} < \mathbf{t}\right]\left\{1 - \mathbb{E}\left[\left(\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]\right)^{\mathbf{J-1}} \mid \mathbf{J} > \mathbf{1}\right]\right\}$$
$$+\mathbb{E}\left[\left(1 - (1-F)^{\bar{\mathbf{d}}\wedge\mathbf{t}}\right)^n\right]\left(1 - \mathbb{E}\left[(1-F)^{n\left(t-\bar{\mathbf{d}}\wedge\mathbf{t}\right)}\right]\right)$$

**DISSECTION OF THE THEOREM (continued)**

- **And the first term $\bar{\mathbf{d}}_{\mathbf{1}} \neq \mathbf{d}_1$**

$$\frac{\mathbb{E}\left[\left(1-(1-F)^{\mathbf{d}}\right)^n\right]}{1-\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]}\mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge\mathbf{t}\right)}\right]\times$$
$$\mathbb{P}\left[\bar{\mathbf{d}}<\mathbf{t}\right]\left\{1-\mathbf{K}\mathbb{E}\left[\left(\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]\right)^{\mathbf{J-1}}\mid \mathbf{J}>\mathbf{1}\right]\right\}$$
$$+\mathbb{E}\left[\left(1-(1-F)^{\bar{\mathbf{d}}\wedge\mathbf{t}}\right)^n\right]\left(1-\mathbb{E}\left[(1-F)^{n\left(t-\bar{\mathbf{d}}\wedge\mathbf{t}\right)}\right]\right)$$

**PROOF OF THE THEOREM**
**(with thanks to de la Peña and Giné's Decoupling monograph)**

**The whole trick is that $\{\mathbf{J}, \bar{\mathrm{d}}_1, \mathrm{d}_2, ..., \mathbf{d_{J-1}}, \bar{\mathbf{d}}_\mathbf{J}\}$ has a dependence structure owing to the cut-off at $t$, but subsets not including both $\bar{\mathrm{d}}_1$and $\bar{\mathbf{d}}_\mathbf{J}$ are independent. In what follows, whenever we use $0$ correlation, or replace one random variable by another, refer to the key lemmata. Whenever the reason for a step is unclear, it is trying either to set up the conditions to apply a lemma a step or two later or to introduce an indicator function a step or two later.**

$$\mathbb{E}\left[\sum_{j=1}^{\infty} \mathbf{e}_j^n\right] =$$

$$\begin{aligned}
&= \mathbb{E}\left[\sum_{j=0}^{\infty} \mathbf{e}_j^n\right] - \mathbb{E}\left[\left\{(1-F)^{(t-\mathbf{t}_1)_+} - (1-F)^t\right\}^n\right], \textbf{ mechanically} \\
&= \mathbb{E}\left[\sum_{j=0}^{\infty} \left\{(1-F)^{t-\mathbf{t}_{j+1}} - (1-F)^{t-\mathbf{t}_j}\right\}^n \mathbf{1}_{j+1<\mathbf{J}}\right] \\
&\quad + \mathbb{E}\left[\left\{1-(1-F)^{t-\mathbf{t}_{\mathbf{J}-1}}\right\}^n\right] - \mathbb{E}\left[\left\{(1-F)^{t-\mathbf{t}_1} - (1-F)^t\right\}^n\right]
\end{aligned}$$

**where the sum and next term are purely mechanical, and definition $\mathbf{t}_1 = \bar{\mathrm{d}}_1$ and lemma $\bar{\mathrm{d}}_1 \le t$ justify the last term. Note that when $\mathbf{J} = 1$ the indicators kill the whole sum.**

$$\begin{aligned}
&= \sum_{j=0}^{\infty} \mathbb{E}\left[\left\{(1-F)^{t-\mathbf{t}_{j+1}} - (1-F)^{t-\mathbf{t}_j} \mid j+1<\mathbf{J}\right\}^n\right] \mathbb{E}\left[\mathbf{1}_{j+1<\mathbf{J}}\right] \\
&\quad + \textit{same}, \textbf{ using monotone convergence} \\
&= \mathbb{E}\left[\sum_{j=0}^{\infty} \mathbb{E}\left[\left\{(1-F)^{t-\mathbf{t}_{j+1}} - (1-F)^{t-\mathbf{t}_j}\right\}^n \mid j+1<\mathbf{J}\right] \mathbf{1}_{j+1<\mathbf{J}}\right] \\
&\quad + \textit{same}, \textbf{ using monotone convergence. Note that the} \\
&\quad \textbf{indicators kill the whole sum when } \mathbf{J} \textbf{ is } 1
\end{aligned}$$

$$\begin{aligned}
&= \mathbb{E}\left[\mathbf{1}_{1<\mathbf{J}} \sum_{j=0}^{\mathbf{J-2}} \mathbb{E}\left[\left\{(1-F)^{t-\mathbf{t}_{j+1}} - (1-F)^{t-\mathbf{t}_j}\right\}^n \mid j+1<\mathbf{J}\right]\right] \\
&\quad + \mathbb{E}\left[\left\{1-(1-F)^{t-\mathbf{t}_{\mathbf{J}-1}}\right\}^n\right] - \mathbb{E}\left[\left\{(1-F)^{t-\mathbf{t}_1} - (1-F)^t\right\}^n\right]
\end{aligned}$$

**purely mechanically.**

**PROOF OF THEOREM (continued)**

$$\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^n\right] =$$

$$\begin{aligned}
= \quad & \mathbb{E}\left[\mathbf{1}_{1<\mathbf{J}}\sum_{j=0}^{\mathbf{J-2}}\ \mathbb{E}\left[(1-F)^{n(t-\mathbf{t_{J-1}})}\left\{(1-F)^{\mathbf{t_{J-1}}-\mathbf{t}_{j+1}}-(1-F)^{\mathbf{t_{J-1}}-\mathbf{t}_j}\right\}^n \mid j+1<\mathbf{J}\right]\right] \\
& +\mathbb{E}\left[\left\{1-(1-F)^{t-\mathbf{t_{J-1}}}\right\}^n\right]-\mathbb{E}\left[(1-F)^{n(t-\mathbf{t}_1)}\left\{1-(1-F)^{\mathbf{t}_1}\right\}^n\right] \\
= \quad & \mathbb{E}\left[\mathbf{1}_{1<\mathbf{J}}\sum_{j=0}^{\mathbf{J-2}}\ \mathbb{E}\left[(1-F)^{n(t-\mathbf{t_{J-1}})}\right]\mathbb{E}\left[\left\{(1-F)^{\mathbf{t_{J-1}}-\mathbf{t}_{j+1}}-(1-F)^{\mathbf{t_{J-1}}-\mathbf{t}_j}\right\}^n \mid j+1<\mathbf{J}\right]\right] \\
& +\mathbb{E}\left[\left\{1-(1-F)^{t-\mathbf{t_{J-1}}}\right\}^n\right]-\mathbb{E}\left[(1-F)^{n(t-\mathbf{t}_1)}\left\{1-(1-F)^{\bar{\mathbf{d}}_1}\right\}^n\right]
\end{aligned}$$

$$\begin{aligned}
= \quad & \mathbb{E}\left[\mathbf{1}_{1<\mathbf{J}}\sum_{j=0}^{\mathbf{J-2}}\ \mathbb{E}\left[(1-F)^{n\bar{\mathbf{d}}_\mathbf{J}}\right]\mathbb{E}\left[(1-F)^{n(\mathbf{t_{J-1}}-\mathbf{t}_{j+1})}\left\{1-(1-F)^{\mathbf{t_{j+1}}-\mathbf{t}_j}\right\}^n \mid j+1<\mathbf{J}\right]\right] \\
& +\mathbb{E}\left[\{1-(1-F)^{\bar{\mathbf{d}}_\mathbf{J}}\}^n\right]-\mathbb{E}\left[(1-F)^{n(t-\mathbf{t}_1)}\right]\mathbb{E}\left[\left\{1-(1-F)^{\bar{\mathbf{d}}_1}\right\}^n\right] \\
= \quad & \mathbb{E}\left[\begin{array}{c}\mathbf{1}_{1<\mathbf{J}}\displaystyle\sum_{j=0}^{\mathbf{J-2}}\ \mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge t\right)}\right]\mathbb{E}\left[(1-F)^{n(\mathbf{t_{J-1}}-\mathbf{t}_{j+1})}\mid j+1<\mathbf{J}\right]\times \\ \mathbb{E}\left[\left\{1-(1-F)^{\mathbf{t}_{j+1}-\mathbf{t}_j}\right\}^n \mid j+1<\mathbf{J}\right]\end{array}\right] \\
& +\mathbb{E}\left[\left\{1-(1-F)^{\bar{\mathbf{d}}\wedge t}\right\}^n\right]-\mathbb{E}\left[(1-F)^{n\left(t-\bar{\mathbf{d}}_1\right)}\right]\mathbb{E}\left[\left\{1-(1-F)^{\bar{\mathbf{d}}_1}\right\}^n\right] \\
= \quad & \mathbb{E}\left[\mathbf{1}_{1<\mathbf{J}}\sum_{j=0}^{\mathbf{J-2}}\ \mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge t\right)}\right]\left(\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]\right)^{\mathbf{J-2}-j}\mathbb{E}\left[\left\{1-(1-F)^{\mathbf{d}}\right\}^n\right]\right] \\
& -\mathbb{E}\left[\mathbf{1}_{1<\mathbf{J}}\ \frac{\mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge t\right)}\right]}{\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]}\left(\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]\right)^{\mathbf{J-1}}\right]\mathbb{E}\left[\left\{1-(1-F)^{\mathbf{d}}\right\}^n\right] \\
& +\mathbb{E}\left[\mathbf{1}_{1<\mathbf{J}}\ \frac{\mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge t\right)}\right]}{\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]}\left(\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]\right)^{\mathbf{J-1}}\right]\mathbb{E}\left[\left\{1-(1-F)^{\bar{\mathbf{d}}_1}\right\}^n\right] \\
& +\mathbb{E}\left[\left\{1-(1-F)^{\bar{\mathbf{d}}\wedge t}\right\}^n\right]-\mathbb{E}\left[(1-F)^{n\left(t-\bar{\mathbf{d}}\wedge t\right)}\right]\mathbb{E}\left[\left\{1-(1-F)^{\bar{\mathbf{d}}\wedge t}\right\}^n\right]
\end{aligned}$$

**where the two new terms correct for the $j=0$ term in the sum (since $\mathbf{t_1}-\mathbf{t}_0=\bar{\mathbf{d}}_1\simeq\bar{\mathbf{d}}\wedge\mathbf{t}$, not $\mathbf{d}$)**

**PROOF OF THEOREM (continued)**

$$\mathbb{E}\left[\sum_{j=1}^{\infty} \mathbf{e}_j^n\right] =$$

$$\begin{aligned}
= \quad & \mathbb{E}\left[\left\{1-(1-F)^{\mathbf{d}}\right\}^n\right] \mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge t\right)}\right] \mathbb{E}\left[\mathbf{1}_{1<\mathbf{J}} \sum_{j=0}^{\mathbf{J-2}} \left(\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]\right)^{\mathbf{J-2}-j}\right] \\
& -\frac{\mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge t\right)}\right]\mathbb{E}\left[\mathbf{1}_{1<\mathbf{J}}\right]}{\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]} \mathbb{E}\left[\left(\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]\right)^{\mathbf{J-1}} \mid 1<\mathbf{J}\right] \mathbb{E}\left[\left\{1-(1-F)^{\mathbf{d}}\right\}^n\right] \\
& +\frac{\mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge t\right)}\right]\mathbb{E}\left[\mathbf{1}_{1<\mathbf{J}}\right]}{\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]} \mathbb{E}\left[\left(\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]\right)^{\mathbf{J-1}} \mid 1<\mathbf{J}\right] \mathbb{E}\left[\left\{1-(1-F)^{\bar{\mathbf{d}}\wedge t}\right\}^n\right] \\
& +\mathbb{E}\left[\left(1-(1-F)^{\bar{\mathbf{d}}\wedge \mathbf{t}}\right)^n\right]\left(1-\mathbb{E}\left[(1-F)^{n\left(t-\bar{\mathbf{d}}\wedge \mathbf{t}\right)}\right]\right)
\end{aligned}$$

**A geometric series at last! Now it looks like a generalization of Wald's equations (see de la Peña and Giné)**

$$\begin{aligned}
= \quad & \mathbb{E}\left[\left\{1-(1-F)^{\mathbf{d}}\right\}^n\right] \mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge t\right)}\right] \mathbb{E}\left[\frac{\mathbf{1}_{1<\mathbf{J}} - \mathbf{1}_{1<\mathbf{J}}\left(\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]\right)^{\mathbf{J-1}}}{1-\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]}\right] \\
& +same \\
= \quad & \mathbb{E}\left[\left\{1-(1-F)^{\mathbf{d}}\right\}^n\right] \mathbb{E}\left[(1-F)^{n\left(\bar{\mathbf{d}}\wedge t\right)}\right] \mathbb{E}\left[\mathbf{1}_{1<\mathbf{J}}\right] \frac{1-\mathbb{E}\left[\left(\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]\right)^{\mathbf{J-1}} \mid 1<\mathbf{J}\right]}{1-\mathbb{E}\left[(1-F)^{n\mathbf{d}}\right]} \\
& +same
\end{aligned}$$

**Remembering the lemma that $\mathbb{E}\left[\mathbf{1}_{1<\mathbf{J}}\right] = \mathbb{P}\left[\bar{\mathbf{d}} < \mathbf{t}\right]$, the theorem is now just algebra from here.**

**THE FULL THEOREM**

- **It turns out that the foregoing proof goes through exactly in the more general case that we need**

$$\mathbb{E}\left[\sum_{j=1}^{\infty} \mathbf{e}_j^{2n} \prod_{k=1}^{n} \left(\mathbb{E}\left[\mathbf{e}_j^{2k}\right]\right)^{n_k}\right]$$

**but it isn't worth the notational white-out to write down even the statement of it here. Every Laplace transform that our**

**simpler statement contains will be multiplied by a product of powers of the same Laplace transform. The only exception is the Laplace transform of J which keeps the same value as in the simpler statement. The whole thing now fails to simplify into anything as clean as the Main Result**