RECORD, Volume 28, No. 1*
Colorado Springs Spring Meeting
May 30-31, 2002

Session 86L
Cox, Ingersoll And Ross Models Of Interest Rates

Track: Education and Research

Moderator: PETER D. TILLEY
Lecturer: WOJCIECH SZATZSCHNEIDER†

Summary: Our lecturer presents findings from a research project relating to the Cox, Ingersoll & Ross (CIR) model of interest rates. The findings address the extended CIR model and easiest construction thereof; an effective way of pricing general interest rate derivatives within this model; and pricing of bonds using the Laplace transform of functionals of the "elementary process," which is squared Bessell.

MR. PETER D. TILLEY: Our guest speaker today is Wojciech Szatzschneider. Dr. Szatzschneider received his Ph.D. from the Polish Academy of Sciences, and he's been at Anahuac University in Mexico City since 1983. He started the graduate program in actuarial science and financial mathematics at that university. He's also been a visiting professor at universities in Canada, Germany, Switzerland, Spain and Poland and has participated in conferences around the world. He's going to be speaking on the Cox, Ingersoll, Ross (CIR) interest rate model and other interest rate models in general; calibration techniques for this model; and he's going to look at the CIR model as part of a financial market.

DR. WOJCIECH SZATZSCHNEIDER: Why CIR? I'm a mathematician, and the CIR model is a very difficult model to analyze correctly. I tried to solve some problems. I saw many solutions to these problems, which I couldn't solve, but these solutions were wrong. So there are many, many published results, but they're incorrect results. So this serves as a great warning—don't believe the papers or textbooks

† Mr. Wojciech Szatzschneider, not a member of the sponsoring organizations, is a professor at Escuela de Actuaria, Universidad Anahuac in Mexico City, Mexico.

too much, because there are many mistakes, and I will try to explain why. In some parts this must be highly technical because the model is a little bit difficult. I will try to explain why it is difficult and what kind of difficulties could result.

Figure one shows the price of a zero coupon bond. The *r* is the instantaneous interest rate, but usually it is written as E^*. What does star mean? Star means that it is some kind of risk-neutral world, so we have some kind of risk premium, but we don't know what kind of risk premium, so usually the price of the zero coupon bond is written with E^*. So, the risk-neutral measure means that all prices, when discounted, are at least local martingales—let's say martingales. If you work only with interest rates, the only observed world is the risk-neutral world. It's not the case of assets, because if we deal with assets, we have the physical world, the observed world, and we construct a risk-neutral world under this risk-neutral measure, the martingale measure.

Figure 1

Preface

$$B(0,t) = E\left(e^{-\int_0^t r(s)ds} \right)$$

Usually wr itten as

$$E^*\left(e^{-\int_0^t r(s)ds} \right)$$

E^* - under risk neutral measure

But for interest rates only observed world is risk neutral world

Not as in assets case.

Risk premiums for interest rates ALONE can not be introduced.

Many, many textbooks and papers consider risk premiums!

My first disagreement with most researchers is that risk premiums for interest rates alone may be introduced, but not observed. Many textbooks and papers consider risk premiums working with interest rates alone. I don't understand this.

I was talking to some top specialists in the field, and they couldn't explain what the risk premium is for interest rates if you work in markets in which the only assets are interest rate derivatives, bond options on future bonds or many others. So, what do we want from interest rates? Pi means that it is okay. O means that something goes wrong. It should be a flexible interest rate. It should be simple. We should make a decision. We could be able to make a decision in a short period of time.

What does well-specified mean in actuarial modeling? In all modeling, well-specified means that we can observe, or at least estimate, inputs. So, it should be realistic. For example, the Vasicek model is not realistic enough because there is the possibility of a negative interest rate, but maybe that is not a big problem. Of course, it should be a good fit to the data; otherwise we would probably like to have some kind of equilibrium derivation of the model.

So, this equilibrium derivation comes from Cox, Ingersoll and Ross themselves, but it is a little bit different. Let me talk about the CIR model. This is still the preface. It is not my exposition about this. It was constructed in 1980. The relevant mathematics have been known since 1950. Analysis was made by William Fellow, so it is not very new.

Figure 2 is the extended CIR model from 1990 by Hull and White. It is simply defined by the diffusion process. It is not under a very strong assumption, and from sigma, alpha and β, we can conclude that a unique, strong solution to this process exists. If this parameter is a functional constant, we have the CIR model itself. It is sometimes considered to be an extended version of the CIR model. Now we understand that by extending the CIR model, we have the Hull and White extension.

Figure 2

- Cox, Ingersoll & Ross Model
- Constructed in 1980
- Extension 1990 by Hull & White

$$dr(t) = s_t \sqrt{r(t)}\, dW(t) + (a_t - b_t r(t))dt$$

if s_t, a_t, b_t constant $\Rightarrow$ CIR

$\frac{4a}{s^2}$ dimension, if $\geq$ 2 then r(t)>0

So, if alpha and sigma are constant, then we can talk about fixed dimension. If it is greater than or equal to two, then the interest rate is strictly greater than zero if initially it is greater than zero. I don't know of many papers that have written that this must be the condition (that dimension must be greater than or equal to two) because if it comes to the zero interest rate, it is instantaneously reflected. So, let's say local time zero is zero, or, in other words, the Lebesgue measure of zeroes is zero. Of course there are many other extensions for negative dimensions. There is a beautiful paper by Anja Going and Marc Yor about the negative dimensions of the process, that it can be negative in this case, but in this case we have to put an absolute value inside the square root.

There are two parts to my talk. It's probably a great surprise that there are positive results about the extended CIR model, but there are negative results about the CIR model in the sense of working with it. In the financial market, we have assets and interest rates that are very difficult problems and there are many errors in financial literature.

Part 1 is, I believe, the easiest way to define and apply the extended CIR model. Part 2 is what is too difficult, and this is the warning I was talking about (Figure 3). Let me define this once again. This is the extended CIR model. There is a function locally greater than zero just to produce the mean reversion. They are locally bounded. Locally bounded means that they are bounded in any finite interval, but I will not be using this kind of parameterization because it is completely irrelevant. I will be using another parameterization that will produce a better and easier valuation formula.

Figure 3

ECIR Model:

$$dr(t) = \tilde{s}\sqrt{r(t)}dW(t) + (\tilde{a} - \tilde{b}r(t))dt$$

$\tilde{a}, \tilde{b}, \tilde{s} > 0$, locally bounded, different parametrization "better".

Apply change of measure (Girsanov) to BESQ^{d}

$$dX_t = 2\sqrt{X_t}dW(t) + d\,dt$$

and

$$dY_t = 2\sqrt{Y_t}dW(t) + (2b_t Y_t + d)dt, Y_0 = X_0.$$

.. Multiply by a deterministic function s_t

Finally if $s_t \in C^1$

$$dr(t) = 2\sqrt{s_t r(t)}dW(t) + \left\{\left[2b_t + \frac{s'_t}{s_t}\right]r(t) + ds_t\right\}dt$$

We start from squared Bessell processes of the dimension delta. Delta is greater than or equal to zero. This has been well studied for many years, and the most important model results are due to Pitman and Yor 1982, '78, '82 and after. So, if you apply a change of measure using the Girsanov theorem, we can introduce the drift term.

I will be talking a little bit more about this change of measure, but let me review first. We apply a change of measure just to add the risk, and after that we multiply by a deterministic function, sigma t. I will define the CIR model and the extended CIR model by this equation. Why? Because the results are easier if we simply parameterize the model in this way. It is similar to the famous paper by Chris Rogers. Which model for interest rates should one use? They say that one should use the change of time of the original squared Bessell process, but this change of time is a little bit more difficult, and a parameterization is a little bit different and it is a little more difficult to evaluate the formula. So, it is not exactly what we need. We need the simplest way to try to apply and eventually calibrate this model.

Let's talk about this change of measure (Figure 4). Change of measure is simply the Girsanov theorem. So we consider this exponential martingale, which is a true martingale. Usually, the exponential martingale is only the local martingale. It is not sufficiently bounded, so we have to define this exponential like this. Take away minus one half of the quadratic valuation of the process. Everybody believes it is a martingale, but check to be sure. If we can't apply some easy criterion to check it, usually there is difficulty. If we can't simply apply this kind of usual criterion like Novikoff's, it becomes hard to check whether this is a martingale or not, but to change the measure we need a true martingale. Sometimes it's just a little bit tricky. So, this is still a martingale, and the change of drift is justified.

Figure 4

Change of measure

$$X_t \sim BESQ^d$$

$$\overline{Z}_t = e\left(\int_0^{\bullet} b_s \sqrt{X_s}\, dW_s\right) = \exp\left\{\frac{1}{2}\int_0^t b_s dM_s - \frac{1}{2}\int_0^t b_s^2 X_s ds\right\}$$

$$M_s = X_s - d \cdot s$$

$$\overline{Z}_t = \exp\left\{\frac{1}{2}\left[b_t X_t - b_0 X_0 - d\int_0^t b_s ds - \int_0^t \left(b_s^2 + b_s'\right) X_s ds\right]\right\}$$

Fact : $\overline{\mathrm{Z}}_{\mathrm{t}}$ is true martingale $\Rightarrow$ change of drift is justified

These are the examinations that we usually use in these kinds of processes. It is a true martingale. We can change the drift. It means that we have equivalent measure and equivalent probability. Equivalent probability means that it is a derivative or simply that the sets of probability zero are the same. It's the same family of probability zero or probability zero sets in both cases. So now it is extremely easy to price bonds.

We have to price bonds (Figure 5). We need to solve some kind of deterministic equation, which is $F^2(s) + F'(s)$. It is equal to the $\beta_s^2 + \beta'_s + 2\sigma_s$. With some initial condition, it is the end of the interval. So, the solution to this equation is tricky, but if the right side of this equation results in a constant, the solution is trivial.

Figure 5

Pricing bonds:

$$P(0,u) = E\left\{ e^{\frac{1}{2}\left[b_u X_u - \int_0^u X_s (b_s{}^2 + b_s' + 2s_s)\, ds \right]} \right\}$$

Look for

$$F_u(s) = F(s) \text{ that } F^2(s) + F'(s) = b_s^2 + b_s' + 2s_s,$$

$$\text{for } s \in [0,u).\ F(u) = b_u.$$

So, pricing bonds in the original CIR model are really trivial. The numerical solution to this equation is easy, but it is not the best way to try to solve this equation. The best way is to change it into a differential equation, and we have phi prime divided by phi. It's equal, exactly the same as before, $\beta_s^2 + \beta'_s + 2\sigma_s$. And then we can solve this equation at least numerically. In some cases we can solve this equation analytically. This is a very easy equation.

So, we have some extremely elementary calculations (Figure 6). We can compute the price of a zero coupon bond by this formula. And if $\beta_s^2 + \beta'_s + 2\sigma_s$ is constant, the solution is trivial. It's some constant, $Ae^{cs} + Be^{-cs}$. We just have to choose *A* and *B* that satisfies the initial condition, and we have a well-known formula of pricing bonds in the original CIR model. I think it is important that we don't need constant parameters to explicitly price bonds. To price explicitly in a very easy way we need the sum to be constant. There are many functions that satisfy this condition. We have a whole family of functions in which the original bond price is in the CIR model paper. We can price in a very easy way. But practically, what can we do? We usually write too many χ^2 distributions. If you see papers by Cox, Ingersoll and Ross on pricing bonds or pricing options, they are so complicated that nobody can understand where the numbers came from. If you just try to see this equation and put it like this, it is simply trivial. Of course we have some complicated functions, but I believe that it helps to understand what is behind them.

Figure 6

Easy to solve numericall y. Now, "inverse procedure"

Writing $F(s) = \frac{j\,'(s)}{j\,(s)}$, $s \in [0,u), j\,(0) = 1$ we have

$$\frac{j\,''(s)}{j\,(s)} = h(s) = b_s^2 + b\,'_s + 2s_s \text{ in } [0,\mathrm{u}), \frac{j^{\,-}(u)}{j\,(u)} = b_u$$

Elementary calculatio ns

$$P(0,u) = \exp\frac{1}{2}\left\{\left[\ln j_u(u) - \int_0^u b_s ds\right]d + x(-b_0 + j_u\,'(0))\right\}$$

If $b_s^2 + b_s' + 2s_s = c^2$

$j\,(s) = Ae^{cs} + Be^{-cs}$

Usually "too many different C^2".

Let's talk about a practical extension (Figure 7). The first time, it was proposed by two German mathematicians, Schlogl and Schlogl, but they use different parameterization, and they had to solve some very complicated equations. Here it is simple and totally explicit. So, what can we do? For example, in Mexico we have four-week bonds, and everything is done on a weekly basis. We can change parameters every week, consider piece-wise constant functions, these parameters.

Figure 7

Practical extension.

d fixed, $b_s = b_i$, $s_s = s_i$ for $s \in [t_i, t_{i+1})$

$i = 0,1,\ldots,n-1.$ $t_n = t,$ $b_t = b_n.$

Solving for $\tilde{j}$ in each interval $[t_i, t_{i+1})$.

$$P(0,t) = \exp\frac{1}{2}\left\{d\left[\ln\prod_{i=1}^{n}\tilde{j}(t_i^-) - \int_0^t b_s ds\right] + x\left[-b_0 + \tilde{j}(0)\right]\right\}$$

$$x = \frac{r_0}{s_0}.$$

Calibration from the history of bond prices.

$o(i, i+j)$, $i = 1,\ldots,n;$ $j = 1,\ldots,d$ observed prices, d - maturities

choose parameters $\rightarrow$ recover $r_i^{(j)}$

Once again, we have a completely explicit formula for pricing bonds. This is good because if I want to do some statistics, some kind of calibration, it is very important to have an explicit formula for pricing bonds. The Hull and White paper said that we can get these functions using present market specifications, but it is impossible. Why is it impossible? We have present market specifications. We have some interest rate derivatives from now to almost infinity. So, we can try to calibrate the model, but to find this function we have to take derivatives of the third order, but if we don't have the whole price of whole bonds maturing any time, we have to use some type of approximation. If you use some type of approximation, taking derivatives of the third order means nothing. It simply means that we can't do it.

I believe that we can go to the history of the bond prices and calibrate the model using historical calibration. After that we can use a time-series analysis of these parameters. At time *i* we have an observed history of bond prices, which are *o* of *i*, *i* plus *j*, maturities observed prices. We have *d* maturities, and we choose parameters, and from these parameters we can recover this initial, at time i, interest rate.

But if we do it, there is of course some inconsistency because we consider a d-dimensional problem, and we try to calibrate one-dimensional CIR, one-factor CIR (Figure 8). But what we can do is take the average of this theoretical value times *i*, and we have the theoretical value of the short rate I at time i, let's say. We have the theoretical value, and we can minimize the sum of squares of errors, let's say, and this is easy. Why is this easy?

Figure 8

Usually $r_i^{(j)} \neq r_i^{(k)}$

Why? Problem $d-dimensional$, one factor model chosen

set $r_i = \frac{1}{d}\sum_{j=1}^{d} r_j^{(i)}$ theoretical value.

r_i and chosen parameters $\rightarrow P(i,i+j)$ theoretical value.

Goal: Minimize

$$\sum_{i=1}^{n}\sum_{j=1}^{d}\left[P(i,i+j)-O(o,o+j)\right]^2.$$

We have an explicit formula for prices on bonds. We can use the same thing using some more complicated interest rate derivatives, but let me show you the result (Table 1). It is for the Mexican market. If we use the original CIR model, which is not good enough, and compare it with the extended CIR, it is not clear enough that it is a big improvement, but I think that for T-bills it is dramatically better, especially for long-term bonds, T-bills (Table 2).

Table 1

CETES prices adjustment performance

	CIR Se^2	**ECIR** Se^2	**CIR** Se^2/**ECIR** Se^2
13 weeks	0.00070426	0.00028588	2.5
26 weeks	0.00043253	0.00040726	1.1
52 weeks	0.00329334	0.00051163	6.4
Total	0.00443073	0.00120477	**3.7**

Table 2

TIBILLS bond prices adjustment performance

	CIR Se^2	**ECIR** Se^2	**CIR** Se^2/**ECIR** Se^2
13 weeks	0.00003637	0.00000463	7.9
26 weeks	0.00003714	0.00001661	2.2
52 weeks	0.00077185	0.00002062	37.4
Total	0.00084535	0.00004186	**20.2**

So, what does forecasting mean for this model? Forecasting means predicting some changes in the parameters (Figure 9). But the original parameterization is much easier. It is what I consider to be the easiest approach to the CIR model.

Figure 9

Forecasting $\equiv$ Forecasting of parameters

Option pricing - easy.

Much more transparent than in original parametrization.

MORE FACTORS.

Usually $r(s) = r_1(s) \oplus r_2(s)$

Pythagoras theorem:

$$r_1(s) \oplus r_2(s) = r(s) \Leftrightarrow r_1^{(s)} = r_2^{(s)},\ b_1^{(s)} = b_2^{(s)}$$

$$d = d_1 + d_2$$

can be extended for $d = d(s)$

$$Y_i(t) = 2\sqrt{Y_i(t)}dW_i(t) + (2b_i(t)Y_i(t) + d_i)dt.$$

$$r_i = s_i(t)Y_i(t)$$

$$Y_i(t) = Y_i^1(t) \oplus Y_i^2(t)$$

$Y_1^1(t)$ and $Y_2^1(t)$ driven by the same Brownian Motion

$W_1(t) \rightarrow$ Totally correlated part

$$Y_1^2 \leftrightarrow W_2(t)$$

$$Y_2^2 \leftrightarrow W_3(t)$$

W_1, W_2, W_3 independent.

One factor usually means that we work with a one-dimensional Brownian motion, more factors means that we have multi-dimensional Brownian motion and many extensions. Mean reversion can even be a stochastic process. There were very beautiful papers written by Deelstra and Delbaen about the long-term behavior of these kinds of models, but what is the Pythagorean theorem here? If you want to consider it, I believe it is the easiest approach to multifactor CIR. It is easiest, but maybe not the best one because we should observe factors. So it is the paper by Duffie and Kan that factors are yields on different maturities of bonds, but if we try to do it like this, the statistics and the valuation formulas are easier.

So, what is the Pythagorean theorem? The Pythagorean theorem means that a square plus a square is a square. The square of the Bessell process plus the square of the Bessell process will give you another Bessell process, a squared Bessell process. So that is why is it sometimes called the Pythagorean theorem. What does it mean, the sum of this process for Bessell or for Bessell with drift, the "plus" in the circle, it means that it is an independent sum of processes.

So, the sum of these kinds of processes will be the same process, but only if we have this β,—this drift, and β the same in both cases. If not, we will have to work with a multidimensional problem. It's going to be extended for delta or even the stochastic process. I want to put some correlation between factors. Usually in the two-factor CIR model, in some applications, one factor is the real interest rate. Another factor is inflation. They are uncorrelated. At least in Third World countries we can't assume that inflation is uncorrelated with real interest rate. It is simply not true. Maybe in the United States it is true. I don't know. I didn't do any statistics on developed countries, but in Third World countries, it's obviously not true.

So, we can put one part. We can consider a three-dimensional Brownian motion (Figure 10). One part will be independent parts and the other will be totally correlated parts, driven by the same Brownian motion. There are some technical things, but that is not the main problem. In a two-factor model we can write like this, and this correlation, because I can consider this function, sigma, completely arbitrary using this method, I can consider using any kind of correlation between factors. So, this is the first part. I think that using this kind of parameterization, this approach, keeps everything easy. It's even easy to price any derivative products. Interest rate derivative products, options on future bonds, etc. There are many, many pricing formulas in the Ph.D. thesis by Boris LeBlanc, but there are big errors in some calculations. Now I'll switch to the original CIR model and talk about risk premium.

Figure 10

$P(0,t)$ is

$$E\left[e^{-\int_0^T (s_1(s)Y_1^1(s)+s_2(s)Y_2^1(s))ds}\right]\cdot E\left[e^{-\int_0^T s_1(s)Y_1^2(s)ds}\right]\cdot E\left[e^{-\int_0^T s_2(s)Y_2^2(s)ds}\right]$$

And previous method applies!

What is too difficult? Well, let's start from the well-known Ito formula, dW squared of *t* (Figure 11). It is Ito calculus. We teach Ito calculus in Mexico at my university. So, let's say we have W squared of TW—of course W of T is the classical standard Brownian motion, one-dimensional. But we need some kind of autonomous equation, because we would like to have R, which is W squared, we have to put the β of T.

Figure 11

2) What is too difficult

Ito formula

$$\mathrm{d}W^2(t) = 2W(t)dW(t) + dt$$

$$\text{Set } W^2(t) = r(t) \quad = 2\sqrt{W^2(t)}\ db(t) + dt$$

$$dr(t) = 2\sqrt{r(t)}db(t) + dt.$$

$$b(t) = \int_0^t \operatorname{sgn} W(s)dW(s) \qquad W(t) \text{ and } b_t \text{ dependent}$$

complications arise.

Typical errors:

1) $\sqrt{x^2} = x$

2) $\int_0^{T_y} W^2(s)ds,\ W_0 = x \overset{\text{Law}}{=} \int_0^T W^2(s)ds, W_0 = x - y$

dr(t) becomes exactly a one-dimensional squared Bessell process. What is β of T? β of T is called a Levy's transformation of Brownian motion. It is a Brownian motion itself, but it is not the original Brownian motion. This is the major problem in our analysis. For technical reasons, the sign of x is minus one if x is smaller or equal to zero and is one if x is greater than zero. So, it is easy because this is the absolute value. Absolute value times sign will give the value itself, the original value W. I think that it is easy, but now there are plenty of errors.

Because of this, sometimes it's a little bit complicated and sometime it's very complicated. If I work with this process, I don't care about the nature of this Brownian motion. This is simply Brownian motion β. But if I want to work with W(*t)* and *r*(*t)*, so it is W(*t)* and β *t*, they're dependent and complicated. One typical error is the square root of x-squared is x. It more or less appears in the Longstaff paper about the double square root model where the square root also appears in the drift, and I'll talk about this more in a moment. This is a very bad error.

This is an example (bottom of Figure 11). Where is the error here? That we can translate the square. If, of course, we can translate the square of a function like this, because if we translate, we have to put something here and it will be a linear term as well. So, this is completely and totally wrong. We can apply this kind of translation for Brownian motion itself, but we can't apply this kind of translation for any power of Brownian motion. This is shown in a paper by the French gentlemen

Leblanc, Renault, and Scaillet in finance and stochastics. In 2000 they published a correction to the paper, but this error remains.

They used to price options, barrier options using this kind of model or, let's say, the Orstein-Uhlenbeck process, which is the Vasicek model more or less. If we want to calculate options, this error is another warning. If you want to calculate the hitting time of the linear barrier of the Orstein-Uhlenbeck process or the Vasicek process, it is very tricky and difficult if the linear barrier is different from zero.

There is some explanation in this beautiful paper about Bessell processes I mentioned already by Going and Yor: on Mark Yor's website you can find it. Mark Yor is, I believe, the best mathematician in the theory of Brownian motion.

There is another error. It is that linear risk premiums are impossible. There are many things about linear risk premiums. A little bit later, I will talk about linear risk premiums that are impossible in my formulation, which I believe is the correct formulation of the problem.

There are two papers published about the solution of our outstanding problem. Pricing asset options with the CIR model as a short rate, but which is independent of the asset. This is simply trivial. If we can assume that these things are independent, if asset prices are independent of the interest rate model, what can I do? Just take the conditional expectation, and everything is trivial in the sense of not for the calculations because we can't—even in this case we have to invert the LaPlace transform and inversion of the LaPlace transform is not trivial.

This is pathetic, that if we want to try a more general solution, we can do it solving this equation for any trajectory of Brownian motion (Figure 12). Let's try to do it. We know this is Brownian motion, but it should be like this, right? We, of course, know that the trajectories of Brownian motion are highly irregular. So, trying to solve this equation for any trajectory of Brownian motion is senseless.

Figure 12

More general case solution of:

$$\frac{P''(s)}{P(s)} = W(s, w) \text{ for any } w$$

¡This is pathetic!

Let me talk about linear risk premium. What is observed in the financial markets? If you have a financial market, we have real world for assets and risk-neutral world for interest rates, because we have prices of interest rate derivatives. We don't observe the short rate. It is something we can't observe. You only observe interest rate derivatives.

In some papers it's written that in one month, we can take the yield of the one-month interest rate, the short rate, but it is not the case because we simply observe interest rate derivatives. To talk about risk premium, we need assets. It was correctly done in the Cox, Ingersoll, Ross paper, but the explanation was so difficult I think that now nobody can understand exactly what they did, but it is explained very well in this paper by Chris Rogers about arbitrage and equilibrium pricing. So, if there is an equivalent martingale measure for discounted prices of financial assets, there are economies supporting these prices as equilibrium prices. This theorem was proved by Chris Rogers in the paper I was talking about. I was talking about which model of interest rates one should use.

So, let's talk about the one-dimensional market (Figure 13). The one-dimensional market is one-dimensional linear Brownian motion. Figure 12 is the classical model for asset prices, geometric Brownian motion. And *r* of *t* is the CIR model. Let us discount.

Figure 13

$$dS(t) = S(t)\left[s\,dW(t) + m\,dt\right]$$

$$r(t) - \text{CIR}$$

$$dr(t) = 2\tilde{s}\sqrt{r(t)}dW\ \ (t) + \left(d - 2br(t)\right)dt,\ \tilde{s}, d, b > 0$$

$$dZ(t) = Z(t)\left[s\,dW(t) + \left(m - r(t)dt\right)\right]dt$$

$$dZ(t) = s\,Z(t)dW^{*}(t)$$

If we can drop the drift, there is no arbitrage.

This is FTAP

Impossible if $r(t)$ is CIR!

This is once again my CIR model—classical, not extended. To get zeta we discount the present value, discount the prices, and we immediately have this formula. So, a risk-neutral measure means that we can drop this drift away. Simply, if we can write things like this, it is a local martingale. The solution of this equation is well known. It's simply geometric Brownian motion. If we can drop the drift, there is no arbitrage, and there is economy supporting price. This price is the equilibrium price, but this is the fundamental theorem of asset pricing.

The general version was proved by Delbaen and Schachermayer. But this is impossible if *r* of *t* is the CIR model. We can't drop this drift away. It means that we can't use the Girsanov theorem. It means that this exponential local martingale is not a true martingale, so we can drop the drift.

It is a little bit of a more modern explanation of what was written in the paper by Cox, Ingersoll and Ross. This means there is arbitrage, but in doing all this, we want to work with the CIR model to price interest rate derivatives. So, what we really want is the CIR model in a risk-neutral world (Figure 14). I can prove the following theorem. Do we care about risk, about interest rate models, in the real world? I think not, because we want to use interest rate models to price, not for some abstract, real world for interest rates that we can't observe. So, this is a very unsymmetric situation that I define this model. Why not? I can't use this model for pricing because it's too difficult. But I don't care! Simply, this is the theoretical model in the real world, which can't be observed. Then we do the same. We discount. We use the same model as before for asset prices. With the same model, we discount, and we can drop the drift away, and in this case, what is the risk-neutral world? We change one Brownian motion to another Brownian motion, W to W*. It is the main message from the Girsanov theorem.

Figure 14

But what we want is CIR on RNW!

Theorem: if in RW

$$dr(t) = 2\tilde{\sigma}\sqrt{r(t)}dW(t) + \left[\delta + 2\frac{\mu}{\sigma}\tilde{\sigma}\sqrt{r(t)} - \left(2\beta r(t) + \frac{2\tilde{\sigma}}{\sigma} r^{\frac{3}{2}}(t)\right)\right]dt$$

Then RNW exists and in RNW we have CIR.
Similar result if

$$dS(t) = S(t)\left[(\lambda+1)r(t) + \mu dt + \sigma dW(t)\right], \text{ for any } \lambda < 0.$$

So, this is about linear risk premiums, but I don't start with the Cox, Ingersoll, Ross model. What is the difference? Because this term means that I pull down the process faster than for the Cox, Ingersoll, Ross model itself. So this exponential will be the local martingale. The problem is when the local martingale is not a true martingale. The local martingale is bad. The true martingale is good because it can produce change of measure. When the process takes too large a value, it's too high a probability. For this minus term, the power of 3/2 pulls down the process faster, a very high value of interest rate is forbidden here because it is unlikely. A similar result is if we have two things like this, we can do the same, but this is exactly about linear risk premiums. I believe that in the theorem, which appears all over is the idea that linear risk premiums for the Cox, Ingersoll, Ross model are impossible is not exactly true.

Let's talk about the Longstaff model for interest rates. It was a very beautiful paper, but, unfortunately, the solution provided by Longstaff was wrong. The solution appears in many textbooks. So, it is another warning that you have to be very careful working with this kind of process. Square root processes are difficult .

This is the Longstaff model (Figure 15). It is very easy to check that *r* of *t* is the square of *y* of *t*, but *y* is very complicated because the sign is once again plus or minus one. It depends on the value. If the sign of X is plus one, if X is larger than zero if it is minus one, it is zero or less. So, Beaglehole and Tenney discovered the error in the Longstaff model.

Figure 15

2) Longstaff model for IR

$$dr(t) = 2\sqrt{r(t)}dW(t) + \left(1 - k\sqrt{r(t)} - 2\lambda r(t)\right)dt, \; k, \lambda > 0$$

$$r(t) = y^2(t), \text{ where}$$

$$dy(t) = dW^*(t) - \left(\lambda y(t) + \frac{k}{2}\operatorname{sgn} \; y(t)\right)dt.$$

So this is the simplified version (Figure 16) by Beckel and Tenney. We can recognize it to be a kind of Vasicek model. Why is it interesting? This is exactly an exponential martingale. –the expected value of this martingale is zero. So, if it is a real martingale, the expectation of these must be one zero, zero of the exponent. X to the power of zero is one. So, you have to make some very easy calculations. It's an undergraduate calculation. So to price both in this simplified version we don't need to use a Green function, which is a little bit more complicated than published by Beaglehole and Tenney, but we need to solve this kind of equation. This is a completely undergraduate differential equation, so it is easy, but I was trying to solve the Longstaff original problem, and it is impossible. Okay, it is possible with some constant, but if we make some calculations, this thing appears, and it is local time at 0 of the process. It is complicated.

Figure 16

Simplified version:

$$r_1(t) = y_1^2(t)$$

$$\text{and}\quad dy_1(t) = dW(t) - \left(?y(t) + \frac{k}{2} \right) dt$$

Somebody told me that there is a book by Carters and Shreves in which they make some kind of calculation like this (Figure 17). No, because we have local time that appears with a factor plus because minus k over 2 minus L will be plus local time.

Figure 17

Calculations Easy

$$E\left[\exp\left(\int_o^t f(s)W(s)+g(s)dW(s)-\frac{1}{2}\int_0^t\left(f(s)W(s)+g(s)\right)^2 ds\right)\right]=1$$

$$\alpha\ E_x\left[\exp\left(-\left(\frac{\lambda}{2}+1\right)\int_0^t W^2(s)ds-\frac{\lambda k}{2}\int_0^t W(s)ds-\frac{\lambda}{2}W^2(t)-\frac{k}{2}W(t)\right)\right]$$

$$\alpha\ E_x\left[\exp\left(\int_0^t\left(f(s)W(s)+g(s)\right)dW(s)-\frac{1}{2}\int_0^t\left(f(s)W(s)+g(s)\right)^2 ds\right)\right]$$

if and only if in $(0,t)$

$$f'(s)+f^2(s)=\lambda^2+2$$

$$g(s)f(s)+g'(s)=\frac{\lambda k}{2}, \text{ and}$$

$$f(t)=-\lambda$$

$$g(t)=\frac{-k}{2}$$

If it's plus local time, we can't use this calculation, (Figure 18) and this is very, very complicated. Very, very complicated means that I can calculate the LaPlace transform of the price of the bond, but it is still very, very complicated using very highly non-trivial stochastic analysis, excursion theory or something like that. There is no way to invert this LaPlace transform. So, it is out of the question completely.

Figure 18

$$P(0,t)\,\alpha\ E_x\left[\exp\left(-\left(\frac{\lambda^2}{2}+1\right)\int_0^t W^2(s)ds-\frac{\lambda k}{2}\int_0^t |W(s)|ds-\frac{\lambda}{2}W^2(t)-\frac{k}{2}\left(|W(t)|-L_t^0\right)\right)\right]$$

We can only calculate $P(0,T)$, T exponential time independent of the process.

Very, Very Complicated...

Why the Longstaff model (Figure 19)? I should talk about CIR, but pricing bonds using the Longstaff model is equivalent to pricing defaultable bonds using Merton's structural approach. There are two approaches to pricing defaultable bonds. One is the approach by Elliott, Jeanblanc, and Yor and one is the structural approach. The Elliott, Jeanblanc, Yor approach is the subject of many new books published or coming soon about this new approach by Merton. Structural approach means that the default occurs when the value of the firm goes below some level. It is an older approach, but it is nice.

Figure 19

3) Pricing Asset options=Pricing defaultable bonds in the Merton´s structural approach

$$\text{Interest rates} = r_1(t) \oplus r_2(t)$$

$$\text{Asset driven by } W(t)$$

$$r_1 - \text{one dimension CIR driven by } W(t)$$

$$r_2 - \text{CIR}$$

Let me try to do it. If we have interest rates dependent on the value of the firm… for example, we had a very big financial problem in Mexico in '94. There was a very strong correlation between assets and interest rates. We can't consider it independent, at least if you consider emerging markets, so it is exactly the same. If I want to price asset options with defaultable bonds, which is the same as Merton's structural approach, pricing bonds with Merton's structural approach is more or less the same as pricing options, asset options, with this interest rate. So we can once again decompose the interest rate into a totally correlated part, an independent part, and this is equivalent to pricing bonds in the Longstaff model. So, it is too complicated. This is why I called it too difficult to apply. We can't really use the CIR model as a short-rate model pricing asset options if there is some kind of correlation between interest rates and assets.

Thank you very much.

MR. TILLEY: I think I speak for all of us in the room when I say, "Wow." What questions do we have for our guest speaker?

MS. VALENTINA ISAKINA: If we can't apply the CIR model, what model would you suggest?

DR. SZATZSCHNEIDER: I suggest any other, which is not a square root. You can consider the Vasicek model. You can consider some power of the Vasicek model. If you have 100 students working on a numerical solution, perhaps they can invert this LaPlace transform, but I don't believe so.

MS. ISAKINA: So, by stepping away from the square you're taking care of independence? Is that what you're saying? You mentioned that we cannot apply CIR because you cannot assume that the interest rate models and the…

DR. SZATZSCHNEIDER: Maybe you can assume in the United States, I didn't make any studies about the U.S. market. But probably you can apply this. I don't know. It is like some kind of warning—don't believe some things written in textbooks because there are plenty of errors, and I was trying to focus the second part of my talk on these kinds of errors.

MS. ISAKINA: Do you know of any papers that address the application of this to the U.S. markets?

DR. SZATZSCHNEIDER: There are many statistical studies, and usually it is that CIR is a bad model, but it is a bad model if it is two-factor model. The two factors are not enough. Maybe the three-factor model could be good because why two factors? One factor is inflation. Another is real interest rate. But if you consider more factors or a time-dependent model, you can get almost a perfect fit. There are many statistical studies, but usually with negative results for the U.S. market talking about original CIR, not the extended CIR, because it is normal Hull and White parameterization, it is not good enough to get something explicit, even for piece-wise constant functions.