

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high-quality educational resources for free. To make a donation or view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at ocw.mit.edu.

PROFESSOR: So what I would like to discuss today is just touch upon what kind of problems the quantitative analysts are solving in the commodity world, the problems that are somewhat different from the other markets. And I'm sure you have the whole year of lectures here from people in different markets. And you will judge for yourself that the models we're building are somewhat, if not completely, different. So this is my goal today is just for you to have some taste of what kind of models we're looking at. So let's start.

Let's start with the following abstract from the Dow Jones publication, the dispatch which announced in 2009 that the Trafigura-- which is the biggest trader, one of the biggest commodity and energy trader-- is potentially on track to post its best results ever in fiscal 2009 on lower oil prices and contango markets.

Remember, 2008, just the year before that, the oil prices shot to the highest possible level, to \$159 per barrel. And a lot of people blamed the traders for the high oil prices. And yet, in 2008, when the prices dropped-- in 2009, I'm sorry, next year-- when the prices dropped to \$30 per barrel, as you remember, to the lowest possible level that we remember, Trafigura is about to make the biggest profit ever. So it sounds like a contradiction.

So they're making money on the low oil prices and contango. And contango, which I have to explain to you what it is-- I assume that everybody here knows what the futures contract is. If not, just let me know, and I will explain. It's very simple concept, so don't be shy and just tell me.

So I would like to show you how Trafigura will make the record profit in the year when the prices are at the record low level, but they are in contango. So this is the

graph of futures contracts, futures prices, oil futures prices on January 15, 2009. Futures contract is simply the contract that allows you to buy today-- let's say, in January-- the barrel of oil for delivery in some future time. That's all.

You always know what the price will be by looking into *Wall Street Journal*, Section C. Open, you will see. If you want to have delivery in August, the price will be around, what? \$55? If you want delivery on February of next year-- so February, 2010-- the price will be \$60. If you want delivery now, the price will be \$35.

So if you want to buy, basically, the spot price, for your knowledge, there is no spot price, as such. And whatever you see on CNBC, for example, on TV, when they give you the spot price of oil, it's the price of the first future contract-- so this one-- of the nearby, the most nearby futures contract.

So as you can see, this is the curve. And recall that the curve is in contango. If the price is monotone, increasing. If they are decreasing, it's called degradation. That's all. So just a useful term for you to know.

So at that time, January 15th, the prices indeed were monotone increasing. They're going from \$35 to February and to \$60 in February. So this is February of 2009. This is February of 2010. And we can see the prices were going from \$35 to \$60.

So let's now see how Trafigura made money. Can you guess, by the way? Yes.

AUDIENCE: They borrowed money [INAUDIBLE] the spot and then sold the futures contract?

PROFESSOR: Exactly. So that's precisely what they've done. What is needed for that, though? There's one little thing that's required.

AUDIENCE: Low interest rates and ability to buy the spot?

PROFESSOR: This price is the ability to buy. But you need a little bit more. You need--

AUDIENCE: [INAUDIBLE]

PROFESSOR: I'm sorry?

AUDIENCE: [INAUDIBLE]

PROFESSOR: No, no, you lock in. You bought at \$35.

AUDIENCE: Storage.

PROFESSOR: Storage-- you need the storage. You need to be able to wait one year, because you already sold at \$60 next year, in February. So you locked in a massive profit of-- so this is your strategy.

Just borrow 35. Buy one barrel. Store it. Then immediately sell, so short means just go and sell the one barrel for next year at \$60.

So you have-- you made \$25-- I'm sorry, you made \$25 just on commodity. You have to-- on February, you will get this-- February next year you will get this \$25. You'll pay the interest, which will be maybe, let's say 10%. So it will be \$3.5. So you made total profit \$21.

If, like Trafigura, you have 50, 60 million barrels of storage, you can easily calculate how much money they have without any risk they made in this particular year. So whenever you hear that traders are benefiting from the high prices, it's not. Actually, if the same stations existed in the high prices, then the interest would be substantially lower. They could drop their profit by 50%, if, let's say, the same where the prices \$125 to \$135 to \$160.

Right? You would have the enormous interest payment. So in reality, you need the low prices, but you would like to have a contango. You really don't care, price is low or high. So to summarize, we need to have-- I mean, just a little thing here-- that the strategy works like a charm. The only thing you need is storage.

OK, let me now invert the question. Let me ask you this question-- let's say we have the same curve. We are on January 1st, 2009, and we are asked to get the storage. You want to have your bosses calling you and saying, I mean, I need storage, oil storage from August to December.

So in August, we have the price is, what, \$55, in December \$58. You go and get

me-- here's your credit card. Just go and get it.

Well, storage is usually-- you get it at the auctions. So a lot of people come to the auction, and they'll bid. So now you're going with the boss's credit card, right? And you have the following dilemma-- if you borrow too little, you will never get it. If you bid too much, you'll have the winner's curse-- namely, you will never be able to recover the money that you paid for that storage.

So you have to have-- before you bid, you have to have a plan. How do you-- or strategy-- how do you recover the money that you pay for the storage through some foolproof, riskless activity-- some strategy like Trafigura. Remember, they have the riskless strategy.

They lock in the profit. They can go to sleep for a year, and then vacation, and then come back and get this profit back. How do you-- in this particular case, you have, whenever you bid something, you also have to have a plan in mind how to recover this money and even more, to get a little bit of profit.

So my question to you-- you need storage from August to December. How much would you bid? These are the prices. If you need more prices, they're all here. So you need the storage from August to December.

AUDIENCE: Do the same thing? Do the same thing they did before?

PROFESSOR: So how much would you bid for the storage? You don't have the storage, so you have to-- when you get it, then you do what we did before, because before what we did was assume that we already have it. Right now, you don't, but if you win it, you will have it. But if you win it, you have to have the strategy how to recover this money. Yes?

AUDIENCE: Depends on how much profit you want to make.

PROFESSOR: Well, it's not, because remember, you are competing against other people. Right? So if you become too greedy, they will outbid you. So let's say bit one penny, right? Assuming that you'll make-- so it should be-- what is the highest you can bid?

AUDIENCE: \$3?

PROFESSOR: Uh-huh. You say \$3. All right, give me the strategy that recovers that \$3.

AUDIENCE: But we do the same process. We borrow money, buy--

PROFESSOR: You don't have to borrow, because it's futures contract. You don't need-- you just go long the August. So buy August, using the futures contract, and sell December using the futures contract.

You don't even need to borrow money. You will need to borrow, maybe, when you get to August, right? Then when you have to pay \$55, then you'll borrow, right?

So you bid \$3, you recovered \$3. No profits. So most likely, you will bid \$2.99, right? That's the highest probably, can go. You have to give yourself at least a penny of a profit, right?

AUDIENCE: The interest.

PROFESSOR: I'm sorry?

AUDIENCE: Also the interest.

PROFESSOR: Let's forget about interest for a moment, just for simplicity. In reality, of course, you never should forget about it, but in order not to make our discussions too complicated, no interest. So you're basically will be at \$2.99, it seems to me, right? You get a penny of profit, and you have a strategy if you get the storage.

You know what to do. You immediately go buy August. Using futures contract, buy August, and sell December, locking in \$3. Pay \$2.99, if you manage to win, if everybody is not as smart as you are, and just bid-- or maybe they want bigger profit. You will win the storage. You will lock. OK, and everybody agrees, right?

This is actually a standard strategy. People use it all the time. Understand, that's what I would call-- that's what the trader would do. That's kind of the trader, so a business guy would do that. This is a common strategy-- was very, very common,

let's say in the '90s.

What the quant will do-- and that's where the added value of the quant is to the organization-- they will do something completely different. What they will do-- they will, on January 1st, they will sell something that is called August-December spread option. You heard the word option, right? Option is characterized by the payout at expiration.

So we have expiration. We have the payout. This is not your typical options, not like IBM option. This is something different.

The payout is determined at the expiration, which is, let's say, July 31st, right before the beginning of August, right? You will look again at the *Wall Street Journal* and look at the December and August prices on July 31st. And if the difference is positive-- this little plus sign means that if it's positive-- then you pay to the owner of the option this difference. If it's negative, you pay zero.

You're more familiar with the options where one of this thing is just a fixed strike. It's called the fixed strike. Here, there's no strike. It simply the difference between December and August contract.

But it's a two-dimensional object, so it's a little bit more complicated to value it. But there is a whole methodology developed, the same way for the regular options. This is for the spread options.

So the quant will sell this spread option on August 1st. Why is it better? Well, first of all, we have to discuss how much will I get for this option?

When I sold it, I give it to you just for you to be confident that I'm not trying to deceive you or anything like that. You will get-- this is the formula that is used to compute the value of that option. And when I substituted all the parameters that are necessary, I get that the value of the option is 447.

So I immediately, on January 1st, got 447. Everybody else is going ready to bid \$3. I have in my pocket 447. I can bid, let's say, 420. Right? Guaranteed if I know that

everybody else is bidding \$30, I will win the storage.

Plus my profit margin is not a penny anymore. My profit margin is \$0.27, right? If I can do even 410, even increase my profit margin if I really want to be greedy in this particular case.

But clearly, I can have a bigger margin. I can have a bigger-- you can guess, by the way, without even looking at the formula, why the value of this option was the payout like this-- was this payout.

Why is it bigger than \$3? It's always bigger than \$3.

AUDIENCE: Discount?

PROFESSOR: No, forget about, again, interest rates. It's not the discount. What is the-- on January 1st, what is the intrinsic value?

If there's no volatility at all, then the intrinsic value of the options-- so at zero volatility, the value of the option is exactly the difference between December and August on January 1st, which is \$3. So \$3 is the intrinsic value of that option. And we all know that the value of the option is greater than the intrinsic value if there is a volatility.

Because there is a volatility, the value is greater than \$3. And actually, it's substantially greater than \$3, because the volatility in [INAUDIBLE] markets is very high. It's much, much higher than what you see in the interest rates, or effects or equities-- indices, I mean. Yes?

AUDIENCE: So you're getting money from taking on more risk, basically?

PROFESSOR: Let's-- we'll get to that. You are asking exactly the correct question, because yes, I am proud. I got 447. I paid 420. I mean, I was bidding for 20. Of course I won.

I got my storage. I bring it back home. Now let's see what kind of risk I brought home. I have \$0.27 in my pocket left. That's my profit.

But now let's assume on July 31st, let's say December price goes to \$80 and the August price goes to \$55. I sold this option. How much do I owe on July 31st-- option is exercised.

How much do I owe to the owner of the option if December, on July 31st, 80 and August is 55? Yes, \$25. I have only \$0.27 in my pocket, right? So what do I do?

That's my risk, as you're telling me. This is a risk that all of a sudden, I owe astronomical amount of money to the person to whom I sold this option. So what do I do?

Do I run to Venezuela or what, I'm interested to know. Where? Can you guess?

Remember, I have storage. And on July 31st, the August price is \$55, and the December is \$80. So you already told me what to do in this situation. That's what traders would do.

They will buy immediately the August at \$55, immediately sell December, because I have the storage. Now I can extract this \$25 using my physical asset. So this is the beauty of the physical or real options-- that I can, by doing certain things, I can extract the payout of the option. So I'm completely protected.

So it seems to me that my storage and that spread option is the same thing. Because I'm fully hedged with my storage. Right, is that clear? Do you understand how I managed to escape terrible predicament? Yes?

AUDIENCE: What happens when the value of what you have in storage falls greater than what you received from the auction?

PROFESSOR: But it's an auction, right? So I received 447. Everybody bids \$3 or \$2.99 and then for 20. There's no value of the storage, except for the value what people are bidding for. So if my bid is the highest, that's it. I received it. So is that your question, or are you asking me something else?

AUDIENCE: I want to visualize the payoff. It's when the--

PROFESSOR: Oh, you're talking about what happens if vice versa, December not goes to \$80, but to \$20, right? And this one still remains at \$55. I do nothing, because the payout of the option is equal to zero, because the difference is negative.

So I owe nothing to the owner of the option, and I do nothing with my storage. It's a fully hedged proposition. So the storage and this option is one and the same. OK? Any questions? So conceptually, we're on the same page, right?

OK, so that was of course of the caricature of the real situation. So reality is somewhat more complicated, as always. In reality, let's say I go and bid for two years. I need the storage for two years.

There are many spread options I can sell, right? In the this, my example-- again, in the caricature example, I sold August-December. I could have sold August-November, August-October, right? Or I can sell-- if I have for two years, I can say May, November, June, September, and so on. If I have two years, it's 24.

You can understand how many options I can sell. I mean, just options where-- options to put the oil into the tank today, and extract it six months later with some profit, or maybe three months later. And so there are a lot of these options.

So first of all I have to determine-- so I can sell a lot of this option against the storage. I have to determine a, what is the most valuable portfolio of options I can sell against the storage? This is-- our strategy starts shaping up. So we want to optimize something.

And what are we optimizing? We're optimizing the value of the portfolio that we can sell against the storage. Value is-- remember this cash that I get from selling the option?

I want to get not just 447. I want to get 447,000-- just whatever. just I want to maximize, and bring it to the highest possible level based on the information-- price information, volatility information-- that exists at this particular moment. I want to determine what kind of portfolio options.

Now, this is optimization. What are the constraints? And now the constraints are requiring that you have some technical, contractual, legal, and environmental understanding of what's going on.

The simplest constraints I can give you is that you cannot put the gas in the ground or oil in the tank as quickly as you want. There are certain constraints on the injection rate. There are certain constraints on the withdrawal rate.

You cannot instantaneously extract gas from the ground, or oil from the tank. You can never have a situation-- remember that whenever the option expires, you have to do something to extract. Remember, to extract the value, whenever you owe somebody-- the option holder-- \$15 or \$25, you have to put the oil in the tank, wait six months, extract it.

Under no circumstances-- so your option portfolio should be sold in such a way against the storage that under no circumstances, you are in the situation when the option is in the money. So option is positive. You have to do something to extract the value.

So you have to inject, and there's no space, because you've done something before, because some other options expired before, and you put already oil in the tank. And the tank is full.

And now a new option expires, and tells you put more into the tank, and there's no space. In the opposite situation, when you need to sell from the tank and the tank is empty, because some other options from that portfolio managed to completely deplete the oil from the tank. Questions?

So let's try to make an attempt to write this optimization problem. Just for fun, so that you have an idea what's going on. So first of all, let's see what we're trying to optimize.

Let's start from the end. This contract simply tells me that I wanted some particular-- F is the price of the futures contract with expiration, let's say, the month of June. So this particular term of the sum tells me that how much I will have to pay in June if I

buy-- this is the volume.

Let's say right now is January. If I buy in June this amount of oil using my future's contract, how much I have to pay? This is kind of cash from my pocket. Right? I mean, it's clear.

The same thing here. There is option here. This is just straightforward buying and selling, buying into the storage and selling from the storage. Agreed?

A little bit more complicated is this list. This is the option as I told you before-- the option that injects oil into the tank at one month, and extract it in month I, and extract it in the month J, which is later. So you extract in June, extract in November.

So this is the value of the option. This is the volume associated with this option. I sold this option. Therefore, I have the positive cash flow. So this is 447 in my pocket.

This is another option that I sold. This is an option which is kind of opposite, which is first you extract, then you inject, which is typical situation when you receive your storage with already oil in the tank. The curve is not contango, but degradation.

So it's not profitable for you to sell it now because the prices are high, and then replace it later. So it's kind of symmetric situation to this one, to the options that we already discussed. It's just it works when the curve is in the opposite direction.

So you try to make money no matter how. Just the prices look. You're trying to make this money. So you try to optimize this portfolio. OK?

Now, let's talk about constraints. To determine the constraints, I introduce right now the Boolean variable, which is 1 or 0, which basically tells me if, when I come to the exercise of this option-- like in our case on July 31st, in our first example-- and I see the option is in the money or not. So do I have to do something or not? Do I have to inject into the storage?

And so is it \$15 or it's minus \$15? If it's minus \$15, the difference between the

prices, I don't do anything. So then it's 0. If it's plus \$15, I have to do something, because the owner of the option asks me \$15, right? And so I have to put oil into the tank, and extract it sometime later. So that's one. The same thing with the option which is opposite-- first withdrawal, then injection.

So then injection constraints will be quite simple. So at the expiration, I look. If this one is 1, then it means that I have to do something. And what I have to do-- I have to inject at the months I , I have to inject this volume X_i , which was withdrawn later. So that is injection.

If some option before that was in the minus-- so I have to inject before and withdraw now, that will cancel my injection. So if simultaneously, I have injection, because one option is in the money, and withdrawal because the option that was exercised before is in the money, they kind of cancel each other. So that's why this sign is plus and minus.

The same thing with others options, and the same thing with pure futures contract. And that should be less than the injection rate. The same thing with withdrawal rate, right? I mean, it's a similar consideration.

Now, maximum capacity constraints, it tells me that if I start with this inventory in my tank at time 0, and this is basically how much I will get at any time I , for any months I , that's how much I will have in my tank by that time I . You have to believe me, because it's not a trivial thing. But the same thing with minimum constraints.

What you really have to understand, that unknowns are x , y , z , v , and unfortunately, ω is also unknown. ω is just-- it's basically these are the volumes. This is the kind of control variable-- the price.

And the price can be-- the spread can be positive, can be negative, can be 0. And this is becoming a very ugly, non-linear problem, very quickly-- very big. So you have a lot of variables.

For two years, you have, what, 12 times 11, which is 132. Let me just-- no, 24. I'm sorry. It's 24 times 23, so it's 12 times-- so it's a big number of variables, plus an

equally big number of constraints, and the constraints are non-linear. So the problem is pretty hard.

I leave it up to you to decide how to solve it. That's why they take optimization courses here. I can suggest that there are several approaches people would take.

There's approximation, where this problem is approximated, let's say by linear programming, or quadratic programming, or whatever you want. You can do it through Monte-Carlo simulations. Or there is an interesting approach through the stochastic control. I recommend you a paper by Carmona and Ludkovski, exactly how to do this-- how to make a decision of injecting oil or withdrawing oil based on the stochastic control, stochastic optimization methodology. But it is quite an interesting paper.

So either of these approaches can be used. And they are used, and give you sometimes different results, but you can [INAUDIBLE]. Any questions so far?

AUDIENCE: Is the stochastic control solution an optimal solution or exact solution? Is it giving the solution that the Monte-Carlo simulation's approximating?

PROFESSOR: Let me put it this way. None of these solutions is better, or pretty, or whatever, because of the parameters that go into the problem. And you can have the most precise methodology, but if your parameters are only 50% accurate, in reality, just all those parameters necessarily for Monte-Carlo simulations with stochastic control - we really don't know what they are. We can only guess through some implied market brand with volatility, and so on. And if you're wrong, you're wrong. Even your method can be absolutely precise.

I like it, because it's really very nice mathematics there. We personally don't do that. I'm not going to tell you what we're using, but we do something different. And our methodology is chosen to be robust.

We chose it because we want the methodology to be extremely robust. So we don't want small changes-- we don't want to have the situation where small changes, which usually comes when you over parametrize. I mean, people like to go and

later-- in the second half-- I will discuss some of the models people use.

For example, they see the richness of the behavior of the prices, and they want to introduce a lot of parameters to capture this richness. By that, they sacrifice the stability and robustness. So little changes in the parameters and the value can change substantially. So we prefer a different approach, where we maybe a sacrifice some of the value, but we will gain the robustness, stability. And this is the most important thing.

And most importantly, everything that we use in the model can be verified by the outside regulators and controllers, which in this day and age is extremely important. Because these days, to have a model that is calibrated-- even the word calibrated causes a lot of antennas just going up. So we try not to calibrate anything. Everything that we do can be verified-- every little brick that goes into the model can be traded in the market, and can be verified. This is very, very important in this day and age. Yes?

AUDIENCE: [INAUDIBLE]. What do you mean by calibration [INAUDIBLE]?

PROFESSOR: We'll get to that, but very quick answer is this-- just you have a lot of parameters in the model. I'll show some of the models people use. These parameters are not-- you don't see them in the market. It's the result of your model.

It's because you chose this particular model. Black-Scholes uses one parameter. If you use Black-Scholes with jump diffusion term, and some other, you introduce another ten parameters-- these parameters are not observable. So you have to somehow calibrate it to some market data.

And the calibration becoming a pretty interesting, because calibration usually is done through some kind of least squares approach. So you try to look at the observable data, and then adjust your parameters using some least squares approach to match that data. Least squares is notoriously bad problem, so you have many solutions.

The solution set can be very flat, so you can stop very early. So there are problems with least squares. It's non-linear optimization, so by itself, it's very difficult. So calibration can be very, very unstable. Any more? OK.

Now I have to tell you the secret-- there is no Santa Claus. Just remember, I was telling you that I go and I sold the option to get 447. I've got the option. In reality there is no option market. Nobody will buy this option from you. The market for this option-- non-existent.

Now what do we do in this particular situation? That's where the whole beauty of Black-Scholes comes into play. Black-Scholes didn't get the Nobel Prize for integrating some very simple payoff function that you can do in first year of school.

They got the Nobel Prize-- or they got kind of appreciation for what they've done, because they showed that, with this value, they don't also have the strategy that allows you by doing something every day or every half a day, you can replicate this [INAUDIBLE]. So if you paid for the option \$5, then you can show what to do every day. They tell you what to do every day in order to get back this \$5 at the end.

So that's what their main achievement. Or if I sold an option, and received \$5, then I will do something that-- I will use this \$5 to replicate any payout I owe at the end. So I sold it to you, for example. So at the end, I owe you \$50. But they showed how, through the dynamic hedging, and using this \$5 that I receive up front, I can be able to meet my obligations to you. So they showed how to replicate the payout of the option.

So in reality, I don't have to have the market. If I sold somebody an option for 47, I don't have to have anybody who buys it. What I will do, I will simply use this dynamic hedging strategy that Black-Scholes advised me, adopt it to the spread options, which at the end will produce for me 447. That's it.

But here again, this is another task that quants will have to do. Not only they will find the value. They will produce every day the strategy that allows you to extract this 447.

So we already showed what to do with the storage. But now you have to do on the next level of complexity. You have to tell to the trader what to do every day in order to, in reality, get this 447 from selling this option.

So you told this trader, I'm selling this portfolio for \$100,000 for the storage. OK, so, but there's nobody to sell it. And you'll say, don't worry. I will do it through dynamic hedging. At the end, you will get your \$100,000.

So he will believe you. And then he will do it. He will get \$100,000.

Then, on the next level you have to work with the guy who operates the storage. And you will tell the person when to inject oil, and when to withdraw, and to get all of this magic which we've discussed before. So I think it'll be a very logical point for me right now to stop. And in the next half, we'll discuss how to model the power plant.

So we're going to the next topic. So far, we have covered how to model a physical asset named storage. The same approach can be used to model practically everything that we're interested in.

For example, we can use the same thing to model the tankers, power plants, these storage refineries, anything, power lines, pipelines, everything can be modeled using this methodology. Of course, the nature of the beast will be different. And you have to understand the nuances and the-- as I said, the constraints.

This modeling, for example, tanker requires understanding of all the routes, and the constraints in the ports, and all these things. But the conceptual philosophical approach is the same. You have to find some optionality that adds some additional value.

So I obviously don't have time to go over the whole thing. But let's, again, for you to have a taste of this, let me just-- let's decide how to model the power plant. Assume that you are the manager of the merchant power plant. Merchant means that you decide when to run it, or not run it.

You run it to maximize the profit of this power plant. So how do you decide? Let's

say you decide once a day, in the morning, whether to turn it on or turn it off, or not to turn it on.

So how do you decide? Well, sounds complicated. In reality, it's very simple. You wake up in the morning. You look, first of all, at the newspaper and find out what's the price of electricity today. And price of electricity you know for each hour-- sometimes for each 15 minutes. It's determined.

So you know for each hour, or maybe the daily price, enough for the whole day. So that's the price. If you sell it, that's how much money you will get.

On the other hand, to produce one megawatt-hour of power, you have to do something. You have to turn on your turbines. So you have to bring some fuel.

You have to put some fuel into the turbines, just to make them move, and produce electricity. Fuel can be anything. Let's say it's natural gas, but some fuel.

Now you have to determine-- so you know how much money you'll get for one megawatt-hour of power, right? But how much it will cost you to produce it? So how much fuel you need to determine how much-- first of all, how much it costs you in terms of fuel to produce this one megawatt-hour, you have to know the efficiency of the plant. Because efficiency of the plant is telling you how much units of fuel-- let's say how much MmBTUs, so millions British thermal units of natural gas-- that's how it's measured-- how much MmBTUs of natural gas you have to burn to produce one megawatt-hour.

Well, this measure of efficiency for the power plant is something called heat rate. So heat rate is exactly this coefficient by which you multiply the-- so if I say that heat rate is 7, it means that I need 7 MmBTUs of natural gas to produce one megawatt-hour of power-- that's all. So in our case, there's nothing to be concerned about. It's just simply some constant that is given.

It's some constant between 7 to 20. 20 is a very inefficient plant, very rarely run. Seven is right now more kind of a standard constant. That's the constant corresponding to the natural gas power plant, which is-- right now the majority of the

new plants is the natural gas plants.

And there are some other costs associated with producing one megawatt-hour-- just air-conditioning, labor costs, and so on and so forth. Typically, they are not the biggest component. So if, let's say, this is seven, what is the price of natural gas right now? Do you remember?

AUDIENCE: \$3.20.

PROFESSOR: That's-- I wish, but it's not. It's around \$4 right now, let's say, per MmBTU. So you need \$28. This is probably around \$3, so you need \$31 to produce one megawatt-hour.

AUDIENCE: I was thinking of gasoline.

PROFESSOR: Oh, no, not that gas-- natural gas. Sorry about that, but just even gasoline is right now around \$5.

AUDIENCE: Really?

PROFESSOR: Yes. Or maybe you are filling your tank someplace which you'll share with us. [LAUGHS] So is that clear?

That's basically-- so you wake up. You look at the price of the power. Let's say it's \$50. You know your plant is, let's say, heat rate 8. You look at the price of natural gas. It's \$4, variable cost \$3. Will we run the plant?

What will be your profit? So \$4, 8, 3, so it's \$35 it will cost you to produce one megawatt-hour. This is \$50. You made \$15 profit. You turn it on. It runs.

If, on the other hand, the price of power will be \$30-- you look in the newspaper-- you go back to sleep. And nothing happens. You get zero for that day. So this is the maximum between this price and zero. Agreed?

Well now, before I go to the next slide, I have to-- together, we'll have to determine what is-- if I want to buy a plant-- if I want to buy a power plant, how much I'm willing

to pay for that. Well what I'm willing to pay-- I know that every day I'm getting this thing, right?

But I don't know. Each of these things is a random number. So I have to kind of buy- - I have to construct the distribution of this price, of this price, maybe correlation between them. It's a two-dimensional distribution, so correlation, find some expected value. And that's how much I will want to pay for the power plant.

So now we come to an interesting question. So now, real work starts. So I know that this is the value that I have to integrate, but I don't know with respect to which distribution.

So I have to now construct the model of the power prices, fuel prices. And then take this two-dimensional distribution and try to find the expectation. And that's how much I'm willing to pay for this power plant-- agreed?

Well, so let's start with kind of building these components. That's what, again, quants will do using already the experience from other markets, and so on. You already heard the other lectures.

So what we'll do, we'll first look at the distribution of power prices to see. Because I don't know how to model the evolution of power prices. I know if I use simply Brownian motion-- let's say it's my first idea.

Well, let's see. Brownian motion gives us the distribution which is, as you remember, at any point of time is kind of normally distributed. If I look at the terminal distribution of price, which is driven by Brownian motion-- if I assume that the price is driven by Brownian motion-- then at the end, I'll get the normal distribution.

So if I look at the, say equity-- this is S&P 500-- indeed it's-- we all know that S&P has fat tails, right? But as we can see in a moment, this is not a fat tail. This is very close to the normal distribution.

For the equity guys, of course, for the guys who trade the stocks, it's an enormously fat tail. But from the commodity point of view, this is just perfect, normal distribution.

This is the distribution that we deal with.

This is-- the tails here are just-- so the normal distribution just out of the window immediately. So Brownian motion out of the window-- we cannot construct this distribution of the terminal prices of power using that. That's one thing.

So we have very fat tails. And this is the kind of the parameters that specify the distribution. If I look at the, let's say the equity index, the kurtosis is very close to the kurtosis of the normal distribution, which is 3 exactly.

But if I look at the Nord Pool, which is the power prices in Scandinavian countries, kurtosis is 26. If I look at the Whanau price, it's 76. So it is, as you can see, as far from normal as possible. And we've seen it in the picture. And these are the numbers corresponding to that.

Moreover, look at the behavior of the prices. This is the price in Texas, for example-- power prices in Texas. What immediately jumps-- I mean, look at the prices, and say, wow, how is it different from what you see, let's say, in the equity world, the stock market, for example?

What is it that immediately jumps at you? What is different here? Go ahead. There's no right or wrong answer.

AUDIENCE: Spikes.

PROFESSOR: Spikes-- that's the key word. Like no other market, we have spikes here. That's a major, major issue for us from the modeling point of view.

You take any standard, say Brownian motion-- it will never exhibit the spikes. Not only that, the volatility, of course, in the prices, as we already expect, is huge. It's hundreds.

S&P volatility right now is 10% percent. This is 100, 200, sometimes 1,000, right? All the intuition you have about the behavior of the prices, the behavior of the random variables, is mostly evolved, [INAUDIBLE], under this 10, 15, 20% volatility assumption.

When the variables behave in a completely different way, when they have volatility of 100% or 200% or 300%, so that's another thing that will be challenging us-- spikes and high volatility. It's just the same thing. It's everywhere, so I just wanted to bring that it's a different region of the country, it will be exactly the same thing. So it's common thing.

So if you look, and start to summarize what we're trying to capture is mean reversion and spikes are more or less the same thing-- so we know if we go far away, it will go back-- high kurtosis, regime switching-- I will talk about that in a moment-- and non-stationarity. That goes without saying. That's true for most of the markets.

And now the problem that we have to capture-- or not the problem but another phenomenon that we have to capture-- that the power and the fuel-- natural gas-- exhibit a very particular structure of the correlation. So correlation is not a number. It's depends the heat rate in the market. Remember heat rate is the efficiency of the power plant that they're running in the market.

Depending on the efficiency, you may have very high correlation or low correlation. We'll discuss that at the end of this talk, which is why this here. Right now, I'm giving you something-- we observe, and our models, preferably, should capture that as well.

So requirements for our model are pretty intense, and pretty difficult. I don't think that you ever assume the requirement of the correlation should have some particular structure, just depending on some parameters, and so on. OK, any questions?

So the models typically-- the first thing people do, they take the models that they know, and they try to apply. Let's just very quickly go through the models. I mean, you've seen these models already in the previous lectures, right? Let's start with the straightforward, geometric, Brownian motion, which out of the window right away, right? I mean you agree, because no spikes, no height, no correlation structure

nothing.

So it's clear that in order to-- we want to kind of start with Brownian motion, geometric Brownian motion, try to maybe modify it a little bit. So what is spike? Spike means that just things go up, and then they're pulled back.

So there should be some mean reversion, right? So mean reversion is good, right? But unfortunately, mean reversion, if you-- first of all, it's pretty strong, so it goes, remember, if the price goes from \$30 to \$1,000, and go back to \$30, mean reversion should be extremely strong, right?

Second of all, what pushes it to the \$1,000 level? You need to have a jump. So that's why people introduce the jumps-- so have mean reversion-- jumps, jumps and mean reversion. Still it is becoming pretty clear that this will not work for the following reason-- it goes from \$30 to \$1,000, and then it goes back, because it's a spike. It goes back within three or four days. It comes back.

So the mean reversion should be extremely strong. The force of pulling back should be extremely strong. But if it's so strong that it moves you back from \$1,000 to \$30, imagine what it does when you're on the level of \$30.

Because it will be completely flat-- just nothing-- you could not move there, right? So people observed that, and introduced, well, they figured out maybe the mean reversion will be different at the level of \$1,000 than when you are at the normal level of \$30, so the difference. So you introduce right now so-called regime switching. So it means that all the parameters change, so you introduce the kind of the high price level, low price level, parameters change.

And now we're getting to the example that we discussed in the first half. You end up with the model with 10 parameters, 12 parameters, which is absolutely impossible to manage. I mean, I can give you another, probably, hour discussing why this approach in general is not good. And not just in commodities, but everywhere else.

I mean, I'm not a big fan of making the models extremely complicated, because introduction of one complexity leads to introduction of another complexity. And you

just cannot manage this thing. It's impossible to manage. Theoretically, it looks fantastic-- unmanageable.

So we have to figure out something else. So all these methodologies-- and I put them there-- the methodologies people actually use-- the stochastic volatility, the regime switching, multiple jumps, and so on. This-- it's all used, but it's not what I want to suggest.

I want to suggest something completely different. These are the methodologies that typically are coming from fixed-income world or foreign exchange or equities and so on. I want to introduce something completely different. Any questions so far?

So I want to introduce the methodology which is more suitable and more understandable from the commodity point of view. Because actually, what's the price of commodity? Price of commodity is driven by only two things. Can you guess which?

AUDIENCE: Supply and demand.

PROFESSOR: Exactly, supply and demand-- that's all it is. That's what we'll try to do. Maybe we can-- we have a hard time modeling the commodity prices using just these standard methodologies from different markets, which rely only on the prices themselves.

Maybe if I introduce some fundamental modeling as well, I can do it without losing the most important part of this model, is it matches the market data. So I have to a, I want to model supply and demand. But I also would like to match every market data that I have.

So maybe I introduce a completely different complexity, maybe not. So let's try to see if I can succeed here or not. Before we're going to go into the depths of this, let's discuss how the power prices are formed.

Power prices are formed in the following way-- let's say that the market consists of two power plants. And there's Generator One and Generator Two. Every day, they get prices of power formed at the auction.

So every time the generators will submit the bids, which says the following-- I can generate 50 megawatt for a given hour. I will generate 50 megawatt. If you ask me only to generate 50, I will generate them at \$20.

Why? Because I will run my most efficient power plant, which is the cheapest one to run. I will generate 50 megawatts. If you want more, if you want me to-- I'm bidding-- if you want me to do 100, then you'll have to pay me \$25, because I will have to introduce less efficient power plants, and so on. 200 to \$30, and if you want me to me generate 600, you will have to pay me \$50.

So that's my bid-- so-called bid stack. That's what I'm sending to the auction. The other guy is sending similar things.

The auctioneer-- there is an auctioneer, organization-- called independent system operator, ISO. They collect all these bids. They know what demand will be tomorrow.

They collect, and then knowing this demand, they will-- first of all, they will sort out all the bids in the most optimal way. So they'll put it all together by sorting. So this is the auctioneer combined all these two bids, created this graph-- so basically based on how much they need to generate. So based on the demand, this will be the price. That will be the clearing price of the market.

And then, because the auctioneer knows all the bids, the auctioneer will send the demand for this generator and that generator. So you will generate 60, you will generate 600, or whatever. The final price is the highest, basically highest price that is necessary to meet the demand. So if the demand is 600, so this will be the price. If it's 800, this will be the price.

So the price is clearly the function of demand for any given day. Well, even if this is the case, then if I look at the-- if I plot, do the scatter-plot demand versus price, I'll have to see something similar. Well, let's see. When we take a particular market, and do the scattered graph, so demand versus the price, you see this is the graph that we expect to see.

It's a little bit fat. Why is it fat? Why is it not a straight line? Why is it not the line as before, but it's fat? What is random here?

Remember I told you that each generator will bid approximately how much it will cost them to generate at the particular day. That cost depends on the fuel price, and that is a random number. So what you can see here is a lot of these curves, but they are kind of randomly moving because of the fuel price affects the cost of running.

But conceptually, philosophically, this is exactly what we discussed-- that these guys will bid this curve. And depending on the demand, the price will be simply the value on that curve. Are you with me? So far so good?

So this is-- that's what we're trying to model. So if we can model that, if I can model for every day the bid curve that the auctioneer, the independent system operator sees, I know the price at that particular day.

Moreover, I don't want to get too complicated, but in reality, I don't even have to know precisely the curve. I kind of have to know precisely the distribution, because remember, for the value of the option, you don't need to know where the price will be at expiration. You just need to know how the price will be distributed.

If you catch the distribution correctly through, again, dynamic hedging and so on-- that's what Black-Scholes tells you-- you can actually value the option correctly. So you can value the power plant correctly. But that's beside the point.

The point is right now I'm trying to model, somehow, the randomness of this bid curve. So now, to summarize, the power price is the function of what? Of demand, clearly-- we already know it is that.

But also the fuel prices, because the fuel prices determine the cost of generation and therefore, how much each generator will bid into the market. That's dependent on the cost. If the natural gas goes through the roof, the price of generating one megawatt will be very high. So the person will be-- the generator will be bidding very

high prices. And that's what we see here.

And also outages-- we have to model outages, because the market is only finite kind of fleet of the power plants. And if a couple of them will go down, the price of power can be affected dramatically. The 1997, the Indiana, the price went from \$40 to \$7,000 because tornado hit the nuclear plant. You take a big chunk of the generation from the stack, and all of a sudden, you have to run absolutely anything, including some very expensive diesels, and so on. So the price of power, obviously, goes very high.

OK, so these are three things that we try to model. So how will we model that? Well, before even I get to the modeling of them, let me just again outline our modeling approach.

If I know the fuel-- let's say there are no outages-- if I know the fuel price, I know then how much each of the generators-- how much it will cost for each of the generators in the market. Because I know everybody in the market. I know how much for each of them it will cost to generate.

So fuel, if I know fuel price, I can generate so-called generation stack. It means the cost of generation for each of the generators, for each of the participants in the market. The outages will simply allow me to take out some of these participants out of the picture, basically. So I need to know that, but that's-- at any given day, if I know the outages, I know the fuel prices, I can construct the cost of generating for everybody.

So it's close to the bid stack. It's something that they will bid-- their cost, maybe plus some profit. I don't know what their profit margins are. But what I know, I know the market prices. I know the options, how they trade, and so on.

So the bid stack will, in general, follow the generation stack. It will be, more or less, the same thing, maybe some added profit requirements. But I can back these profit requirements from the market prices. That's where my supply demand approach and the market will get together.

I will adjust the generation stack in such a way, maybe moving it up and down in such a way, that I matched the prices, and I matched the option prices, so measured the volatility of the market. So now the circle is completely closed, if I can succeed in doing that. Questions? This is the key.

Well, let's see if I can do that. So first I want to-- now if I manage to do that, then if I know how to model the evolution of fuel prices, if I know how the outages are modeled, and I model the demand, then I can determine how the power prices will be moving in time. So power price is the function of the bid stack and demand, as you remember. So if I know the evolution of the bid stack, the evolution of the demand, I can determine the evolution of the power price-- completely different approach.

So let's start with the evolution of the fuel, with building the fuel model. Well, I told you that there is natural gas. In reality, we have to model all that stuff. Because each of them has a curve because of volatility.

It looks a lot, but in reality it's not, because we have a pretty good idea. Unlike the power prices with the spikes and all this crazy behavior, we have a pretty good idea of the behavior of these things, because they're all storable. By the way, power prices are jumping because of non-storability of power prices, because they're not storable.

Because they're non-storable, you cannot smooth out the changes in the demand. Just so you have-- your action's immediate. These are all storable commodities, so the behavior is much more regular. They're much closer, let's say, to the equity prices. And so we can use some standard models.

So from the modeling point of view, they are not particularly difficult to model. So particularly just natural gas, heating oil and fuel oil, just the coal-- just we have a pretty good understanding of how these things are modeled. Outage-- well, we take the standard model from-- and reliability theory provides us with a very well-developed mechanism and apparatus how to model these things.

I mean, usually we do it through some kind of Poisson, or version of Poisson process. It is very well understood. There's a lot of literature on that. We can model that very easily.

Where do we have the parameters for these Poissons? The government provides us with these. There's government data. Sometimes we get it directly from the power plants. Everybody keeps track of the frequency of the outages and so on and so forth.

Demand-- how will we model demand? Well, we'll model demand through the temperature, typically. Why temperature? Because for temperature, we have a lot of data. It's statistically a very stable thing. So there are many different approaches to modeling temperature.

It's up to you. There's a lot of literature, so I'm not going to go into detail. So we choose something. We now model temperature. From that, we model the demand. And it works pretty well.

So now we have-- we modeled the evolution of temperature, evolution of demand, evolution of fuel, evolution of the outages. Now we can construct the generation stack. So remember, the generation was this curve, this curve that was a function of demand, but also the fraction of the outages, variable costs, and so on, and the fuel, the vector fuels.

And then there's these parameters α and these α parameters we choose to match the futures, the forward curve for the power prices and all other market parameters as we need. So that's how we get. So we're matching the market, and we're matching the kind of supply and demand formation.

Now, it is very clear why with this approach, I can capture spikes without effort, without any effort. And the reason is very simple. This is my-- remember the stack, my bid stack or generation stack, whatever. So for high generation volumes, it's becoming more and more expensive to generate.

And after a certain time, you've exhausted all your cheap plants. You have to go to

very, very expensive plants, like diesel plants and so on, which runs maybe once in a lifetime, maybe once a year, twice a year. And very expensive. But they determine the price. They are the ones that determined the clearing price at the market.

Now, let's see. If we are somewhere here-- this is demand. Demand is driven by temperature. Temperature is typically a normal thing with maybe mean reversion. If it goes up, it typically reverses some mean.

So this is the distribution of demand around here. Well, if demand moves left or right, up, down, the prices don't change much. So that's your normal regime.

But let's assume demand is somewhere here. If you are to the left, the prices are very small. But the moment you move a little bit to the right, your temperature moves a little bit to the high region. You immediately spike into the \$1,000 territory.

But remember, temperatures mean reverting, so within five, six days, it comes back. You go back. Your price goes from high levels back to the normal level. That's your spikes. You get it completely for free. And I'll show you right now a couple of graphs. One is the actual graph. The other the simulated using this approach for the same market.

So as you can see, this is the actual price for that particular period. These are the prices that we generate-- just plain Monte-Carlo simulation. As you can see, this can be easily this. Just from the point of view-- distribution of the spikes and so on, it's exactly the same thing. This is what we were after.

Moreover, I'll tell you even more. If I knew the past of the fuel, then obviously my approach-- there are two graphs here. You cannot distinguish between them, because I substituted exactly the right price of the fuel-- just actual price, which was historically. And therefore my price and the actual price were the same.

So, but the reality, of course, I don't know the past. But as I explained to you, I don't need to know. I just need to know the distribution of fuel prices is correctly captured. Once I have the correctly captured distribution of the fuel prices, I have, according to this, correctly captured distribution of the power prices.

Moreover, so as you can see, we have a very nice behavior. Now from the parameter's point of view, this is the model. Look at the kurtosis. This is summers. Kurtosis of the model and kurtosis of the empirical is very, very close. So the distributions are very close. I mean, the skewness and so on-- it's a very, very good capturing of the distribution.

So this approach works very nicely. It's completely different from what you get used to, but it's really the one that works. Moreover, final benefit-- bonus point-- this is simulated correlation structure, remember? And this is the actual one-- very close.

And the beauty of it is you don't have-- it's not an input. It's an output. I don't ever use the correlation as an input in my model. I got it. I got the distribution of power, got distribution of fuel, natural gas, put it all together, compute correlation and that's what I've got. And this means that this is really a correct approach. And I don't need to put it-- I don't need to look for the distribution that has this property. This property comes for free, just as a result of the completely alternative way of modeling this thing.

Now, what is the negative side of this? The negative side is it's extremely difficult to do it and maintain it, because you have to maintain the information of every power plant and was built, was retired, and will be built, and so on. Because you have to look at the power prices 10 years, 20 years from now. You have to know what's going to be there, what kind of stack you have to forecast.

That's a lot of information to keep. I mean, you have to have a big organization to work on that, to maintain it, to build the model. Because each region in this country has a different market. So you have to-- I mean, it's a massive undertaking.

It takes years. So it's not like you can get it from the can. It's expensive.

So I think that will be a good point for me to stop. If you have questions, please let me know. Questions? I think it's music to my ears. Thank you.