[Slide 1: **Economic Analysis for Newer Agencies Part 1 of 2]**

[Slide 2]

JOHN DAVIES: Welcome to this ICN Training on Demand module. I'm John Davies, Head of Competition Policy at OECD in Paris.

[Slide 3: **Overview]**

JOHN DAVIES: These are the main topics that we intend to cover. You can watch the module all of the way through, or you can jump to specific topics by selecting them on the left of the screen. As well as slides and presenters, we provide a transcript of the module in the window below, which you can also download separately.

We're going to begin by discussing the value of simpler quantitative techniques. Then Simon will take the lead on a discussion of how to obtain data. We believe the less experienced agencies are often too pessimistic about their chances of getting useful data, so we'll make some suggestions about what to look for.

Then I'll talk a little about handling the data, and then we will describe some of the most important and useful techniques: natural experiments, price correlation, price concentration analysis, diversion ratios, and then bidding analysis, before finishing with a description of some empirical techniques for detecting collusion.

We will present specific case studies from competition authorities in Kenya, Hungary, and Sweden as we go along.

We'll finish with some suggestions for further reading on economic analysis in competition, but we'll also provide suggestions for reading on specific topics as we go along.

[Slide 4: Quantitative analysis process]

JOHN DAVIES: To begin with, though, why might we want to use simple techniques of quantitative economic analysis?

Well, one reason is that much of competition law enforcement is concerned with determining economic effects. If we allow this merger to go ahead, will the merged firm have the incentive and ability to raise prices? In an abuse of dominance case, is the behavior of this firm likely to harm competition and thereby harm consumers or the economy as a whole?

The precise questions will depend on the wording of the law in your jurisdiction, and we want this module to be useful to everyone, so we won't go into the specifics of that. But these general questions exist, to some extent, in all jurisdictions.

So, we're deciding economic questions all of the time. And if we're doing our jobs properly, we should make the best use we can of all the evidence we can get, including data, including quantitative evidence.

Our intention in this module is to encourage you and help you to find some sources of useful data for such analysis and then to carry out the analysis to provide evidence to help decide your case. However, it's not a good idea to consider only the quantitative evidence, especially just one piece of such evidence in isolation. All the evidence in a case is just part of a whole. Ideally, all your evidence will be consistent. When it's not, you have to choose. And quantitative evidence is not necessarily better than other qualitative forms of evidence when you have to decide how to weigh out such competing claims. It depends on how good the quantitative evidence is, just as it depends on how good the qualitative evidence is for that matter.

It's not easy to give guidance on how to assess how good a piece of evidence is,

but we'll provide some of our own views as we go through this module, and you'll get better at it by trying the techniques and building your own experience.

[**Slide 5: Getting the data]**

JOHN DAVIES: The author Dostoyevsky pointed out that you can't cook with a hare until you've caught a hare. Similarly, we can't do any data analysis without data. The analytical tools you can use will often be limited by how much data you can get and by how good it is. So, I'm going to hand over to Simon now to lead the section on how to obtain data in a competition case.

[**Slide 6: Getting the right data: the basics]**

SIMON ROBERTS: Thank you, John. I'm very enthusiastic about this topic, having spent six years as Chief Economist of the South African Competition Commission, and also having advised a number of other competition authorities in other African countries in recent years.

As has been emphasized thus far, getting good data is absolutely critical. Without data, we actually don't know what we're talking about, and people tend to fall back on their own experiences and intuitions, which is obviously wrong and can be very problematic. These are subjective experiences and don't reflect necessarily what's actually going on in the market.

In addition, the right data is more important than whatever techniques you use. There's no point having advanced techniques if the data isn't robust and isn't reliable. And reliable and robust data, in and of itself, can point to important trends without necessarily using those very advanced techniques. So, it's very important to be able to understand what's going on at the outset.

I'm going to highlight a few main points before getting into the presentation. And these are points that I'm going to come back to and illustrate as I go through. First of all, typically, you'll be provided with data by the party or the parties in a merger, for example.

And it's important to know what that data is. This sounds blindingly obvious, I know, but I've found that people can talk about price data, for example, that they've been given without actually clarifying whether the prices are ex-works or delivered prices. And this could be quite a big difference. The ex-works price is the price at the factory gate that the firm receives; whereas the delivered priced is obviously affected by transport costs and other factors, and those will differ according to the location of the customer.

Similarly, list prices may have been provided by the merging parties, for example, when in actual fact what you want is to look at the prices after discounts and rebates, the effective prices that the customer actually pays after taking into account all of those things. And this is very important if those discounts have changed over time or if those discounts or rebates vary by different types of customers.

The second point I want to highlight is you always need to get disaggregated data in a number of cases and where relevant. So, just to consider prices again, and prices are very important, if you have average prices, then those prices might hide substantial differences, a meaningful variation, such as between different grades of product if there's different qualities. There could be different specifications; there could be different customer groups which have different prices, such as customers that buy on long-term contracts and customers that buy on spot basis. So, these are very important to

understand, and getting the disaggregated data where these differences are material is going to be important.

Generally, we're going to be interested in data at a baseline which is on price and quantity. And what I've been using so far as examples is price data. And this is really perhaps the most important category of data for market definition particularly. And you need to know obviously how people are responding to those prices in terms of the sales volumes, so the two things go together.

Now, you might say, well, why are we not interested in the characteristics of the products and services? And these are obviously important, but what we actually want to understand is how consumers respond buying those types of products to changes in relative prices. And, so, this is really at the heart of the analysis.

It may seem as if getting this disaggregated data is going to be very invasive, very -- is going to require a lot of hard work. The firms will tell you how much effort is involved in this and will perhaps resist providing the data. But, remember, the firms need to make decisions about things that are material to them, and what product do they get to make -- offer promotions on, where are they going to discount. And they need to consider that by different customers and different grades of product. So, in actual fact is it's likely that the firms will collect data which is disaggregated on this kind of basis.

And, so, it may not be as invasive as you think. They may actually have this data, particularly for larger firms who have got marketing departments and are tracking their performance. And, so, you may find in actual fact that the data you want is data that they have, and they routinely collect and analyze this data.

And, so, at an early stage in the investigation, it's helpful to get a good sense from

the firms themselves of what kind of information they collect to track their performance.

It will probably require getting confidential information from the firms at some point, but it's also important to highlight that simple questions as to how the products are priced and sold, which can be asked also to customers, as well as to the suppliers, can yield a lot of detailed information which is relevant and helps you to understand what's going on in the market.

And one point I want to raise or note at this point -- at this time is that there's a danger if you get embarrassed by your lack of knowledge. The firms know everything about the market, and you generally don't. And we're not meant to know about that as competition experts. We're meant to be -- we have to move from one market to the next. But if you are embarrassed by asking questions which appear dumb or get put off by the firms, then you may not obtain the kind of detailed information that you require, or you may make assumptions about what that information is. And, so, asking the dumb questions at an early stage is often very, very important.

And, so, just to summarize, I'd say be careful -- and this is highlighted on the slide -- be careful about jumping to answers without understanding the data. So, for example, if you've got national data, don't assume there's a national market. There may be an important local market that you may need to get data on prices and quantities and market developments at a regional level or at a local level.

We talked about the importance of looking at effective prices after rebates and discounts, and this is another very, very important area that in my experience can be -- can be not understood fully and relates to the information that's being requested.

[Slide 7: The relevant data: some examples]

SIMON ROBERTS: In this slide, I give two examples just to illustrate the points that I've made already. And these are examples which are drawn from actual cases that I've worked on in South Africa and that I've been involved in analyzing. On the left-hand side, there's prices for different grades of coal. And these are -- these are -- these charts are hypothetical charts because they're based on actual experiences.

And you can see that there's different -- three different lines here. There's a line for two -- two lines for Grade A, the same grade of coal, a higher quality coal; and there's a line for Grade C, which is a lower quality coal. This coal is sold to different buyers, including power generation. It's also for industry -- industries that want to run furnaces and use it for various types of industrial processes.

And, first of all, you can see very, very big price differences between the different grades. This may not mean that there are different markets, but it's very important to understand, and if one looks at an average price over time, which would also depend on the weighting that one would use, you'd get a very, very different picture from if you understand it in terms of these different grades.

Secondly, in terms of looking at the difference between the Grade A and the Grade C, you can see there are different dynamics that are driving this or likely to be driving it. The relative prices don't stay stable over time. Those are likely to be broad factors that are impacting on it such as international commodity prices. These are traded commodities.

There's differences that you need to understand. And, likely, problems or issues are a substitution between the lower quality coal, which in South Africa is used largely

for electricity generation, and higher quality coal, which is used by a particular industry, such as the steel industry, et cetera. So, you need to understand this, whereas, you know, if you don't get the disaggregated data, you can see there would be important differences.

The second point to highlight is for the Grade A coal, it's exported, as well as being used locally. And the FOB refers to free on board. That's the price at the port. In South Africa, there's Richards Bay Port, which is where exported coal is priced. And you've got the same grade of coal, which is priced free on truck, FOT, so at the mine, loaded onto the truck. So, these -- this an ex-works price.

And you can see that interestingly the price of the coal locally sold, the free on truck price, is actually quite a lot lower than the export price. And this is really important in terms of -- well, it's really important in terms of different types of cases that we had because there was a big question as to whether exported coal could be switched to supply the local market and whether local coal -- locally sold coal could be exported instead.

And the explanation which came out in the case in fact was that the export terminal had a capacity constraint. So, the exported price was at a particular level, and dollar terms changed over time, whereas the locally sold coal could not be switched into the export market easily because the export terminal had this capacity constraint. And, so, there was local dynamic rivalry which went on over local customers which drove the local price below the export price. And this could and potentially did make quite an important difference in terms of looking at market shares and understanding how those markets worked in practice.

I just wanted to highlight the fact that this disaggregated -- this disaggregated pricing data revealed very important potential differences in markets and market

definition, which then fed through to the analysis in question.

On the right-hand side, I've got pricing for a consumer product. Again, these are made-up or hypothetical price series, but for -- based on an actual -- an actual case. And I won't say specifically which case it is.

And you can see that there is an average price for different brands of the product and that this average price, which was a list price, didn't change very, very much. There's some differentiation between brands, as you can see, but the trend over time is relatively stable. And this might suggest that you look at an overall market for -- for these products.

However, when one obtained the price net of discounts and rebates, it was very interesting to see that at a particular point in time, a particular month, the prices of some of the brands fell by around about 30 percent. And this pointed to a particular competitive dynamic where rivalry -- increased rivalry in what was perceived as a more mass-market or lower quality brand had led to a pricing response. But the pricing response hadn't been to reduce -- by the large firm hadn't been to reduce the price of all the brands together. They had only reduced the prices of a particular category of brands, and they hadn't done it through the list prices. They had done it through the discounts and rebates.

So, they had responded to potential and actually actual customer switching by reducing prices of a group of brands, which actually indicated a potentially different segment of the market -- a potentially different market altogether, so for one type of brand and another type of brand. And without getting the prices net of rebates and discounts, none of this dynamic would have become -- would have become evident.

[**Slide 8: The right data, not just the data readily available]**

SIMON ROBERTS: So, in terms of getting the right data, much as the data was readily available, there's a number of things that we can -- we can do. As I've highlighted already, it's very important to interview industry people to understand how the market works in practice. And, in fact, I'd emphasize that you want to go to their offices and their factories. And ideally, you know, out of -- out of kind of the very legal setting and to be able to understand from them, if you like, on the shop floor, how they go about selling their products.

If you can see around the factory, that's also really good because you can often get a better feel for the types of products. You can pick up how the products are packaged and whether that's relevant and whether there are different -- very different kind of pack sizes and ways in which the products are supplied, bulk loads versus smaller -- smaller loads, which may or may not be important, but it gives you an insight and understanding into how -- how these things work. And people actually like -- generally like talking about what they do and explaining it to an interested outsider, which in this case is obviously yourself.

It's also important to understand how the firm is organized. So, getting an organogram of the company is important at an early stage as well because you can then direct the questions to the relevant people in the company. You can understand how the marketing department works, for example, in particular, because the marketing information and information on sales and prices, as we've seen, is going to be something that's very, very important.

It's also going to be important to get marketing and strategy documents. These

will be confidential documents, typically, but these will explain to you how the firm sees the market themselves and how they track information. And we have an example of kind of the dashboard, which is often used by companies.

And all of this will help you to avoid the common pitfalls which I already talked about and understand the kind of questions that you're going to need to look at in more detail as part of the analysis**.**

[Slide 9: 'Consumer Co' Quarterly marketing dashboard]

SIMON ROBERTS: So, this is an example which shows a typical dashboard type of report that companies would prepare. This is given on a quarterly basis, but this may be monthly or quarterly or annually. And it shows that they are tracking -- in this case, tracking sales in what they call a premium target market. So, that would be something that would be important to understand. And to track that, they must monitor it. So, they would have data on it.

They've got the premium market or premium target, as they've called it here, and they seem to be contrasting that to a mass market. And, so, there's a number of questions which you need to go into, but certainly obtaining information about the prices and quantities sold in these two segments would be -- would be relevant to understand the market dynamics.

And it appears as if there are -- there are some particular branding strategies, as we can see, something called brand preference here. And they're tracking SKUs, or stock keeping units, so there's particular units that they look at and sensitive to.

So, it's obvious that in actual fact they obtain and record and collate and analyze a lot of data to produce these types of reports. And obtaining that data is essentially just

obtaining stuff that they have readily to hand and, you know, we should go ahead and get this kind of information.

[**Slide 10: Data exist: you just have to ask]**

SIMON ROBERTS: So, as we can see from the title of this slide, the point we're making is that data do exist. Again, there may be very small firms which don't collect this kind of information, but if firms are going to make decisions, then they have to -- they have to collect this. And firms are making decisions typically in the context of some uncertainty and variability -- costs are changing, exchange rates will be changing, demand will be changing, the rivals will be behaving in different ways. And, so, they need data which enables them to make decisions about responding to these kinds of issues.

Of course, at the outset, I want to highlight there is data from national statistics agencies. Generally, this is too aggregated for the kind of analysis of competition markets, although as I'll come on to later, in actual fact in South Africa, we found the producer price data was quite useful for initial screening of cartels. And there were signs of a number -- a number of industries and markets of potential collusive behavior, which came just from this level of data.

And in some cases, in fact, six agencies that I'm aware of have been able to provide data on a sub- national basis, which can point to regional variations and possible local market dynamics. Trade data for economies which are very open can also be helpful in understanding markets that may be wider than national, which may be of imports from the country just next door, and it may become increasingly important as regional integration arrangements advance. And trade data can provide data on prices

and quantities, depending on how the data is made available.

But putting those things to one side, the data that exists that you want to pursue in information requests is information typically from -- from the firms. It may be that this information is obtained also from third-party data gatherers. So, firms provide the data to a third party. So, for example, the most well-known is probably A.C. Nielsen for data on consumer products. Or the data is obtained from supermarkets, for example, on the different types of products.

There are also companies in natural resources and agricultural commodities which specialize on collecting this kind of information. And you may find that although those companies charge -- could charge for the data, the companies you're investigating buy this data and have this data and use it. And it may be possible to obtain this data quite readily from the participants.

And, so, there's a number of avenues to get quite important detail data, which is not just on the individual company but on the market overall. Companies who are members of trade associations and industry associations may be providing that information to those associations, and those associations may be tracking the markets and providing data back to the companies. This may be facilitating cartel conduct in actual fact, as has been found in a number of countries across different continents. I'm aware, for example, that Chile has found trade associations highly involved in cartels, as well as in South Africa.

But while companies are not very aware of the likelihood of being caught for collusion or even that what they're doing is coordination, they may be providing quite a lot of detailed data and asking for the data from the trade and industry association may

assist quite a lot in the investigation. Just to give some examples, we've found that the Chamber of Milling, the Chamber of Baking, the Animal Feed Manufacturers Association, the South African Petroleum Industry Association, the Fertilizer Society, and the Cement and Concrete Institute all collected quite detailed data on sales volumes, and in some cases also on prices.

And this data readily provided a very good and full picture of the market, which didn't require going to each of the market participants individually. So, that's something that I think is worth bearing in mind.

And lest we don't forget, I think, to ask for data from other market participants, if you've got a merger, it's important to ask for information from the non-merging parties about-- about their products and prices, because clearly you're going to be interested about the rivalry that they -- that they will continue to provide to the merging parties.

And customers, especially larger customers such as supermarkets, which may have sophisticated data gathering and reporting, are important to go to, as well.

[Slide 11: Using powers to obtain information]

So, competition authorities generally have very far- reaching powers to obtain information, for the very reason that detailed product and firm-specific information is essential for competition analysis. And not using these powers to require the provision of data implies in actual fact that the authority could be open to some kind of challenge for not fulfilling its mandate and responsibilities.

At the same time, the parties clearly do not want to provide detailed information which may open themselves up to greater scrutiny. And they will also potentially and likely challenge the information request of authorities on the grounds of relevance or the

onerous nature of providing the information, and they'll complain bitterly about having to provide truckloads of documents and how much time and effort and cost will be involved in this.

So, a couple of points flow from this. First of all, in terms of relevance, clearly, the authority needs to be prepared to explain why data is essential for this analysis, and data on price and quantity and disaggregated data. And to economists, this is kind of fairly obvious, but in my experience, there are important benefits for just patiently explaining why these types of facts are necessary for the economic analysis of market dynamics, whether it's in a merger or another type of competition matter. So, explaining why these -- this data is necessary if a misleading picture is not to be given, I think is an important part of the process.

Secondly, to the point we've made already, which is that firms may complain bitterly about the quantity of information, but in actual fact, there may be a lot of information that they collect themselves and actually is readily available in a form which is useable because they're using it for their own marketing purposes to monitor the performance of sales staff, et cetera.

So, in actual fact, there may be much more easy ways of getting the data which exist already, which is disaggregated and is material. And they need to -- firms need to explain what information they collect. It goes back to the importance of getting marketing and strategy documents at a relatively early stage. And as I've explained, this is likely to be quite disaggregated, where there are important dynamics that are at play.

So, I think it is important to push back against the threat that these data requests are an overreach, that these data requests may be far too voluminous, may be very

expensive, because in actual fact, it may be that this is being -- these dangers are being -- or these costs are being overstated.

So, don't back down in the face of these threats. I think if you do back down you may well find that too little is collected, and later on in the analysis, the analysis can be attacked for not being robust. And, so, you really need to pursue these. And when there's time constraints in mergers, for example, it's important to pursue these as early as possible.

If you move more quickly in something, which is time-bound, such as a merger, it also means that you can adopt this iterative approach. You can obtain data, understand what's going on, and then go back to drill down into more detail in particular areas where that's necessary. In mergers, of course, the merging parties want the merger typically to be heard and decided, so they have an interest, an incentive, in providing data, as long as it can be clearly explained why it's relevant and necessary for the analysis.

In enforcement matters, in my experience, the stance is much more adversarial and the contestation of the information requests will be much more vigorous because ultimately the parties are typically quite happy if the investigation drags on and on and on and is not resolved, no matter how they may claim to want the case to be heard and decided and this is not necessarily their -- their incentive.

So, you've got to craft information requests and just to kind of put a pointer in the ground, typically my experience, and this is not just from my experience but talking to people in other authorities, what you're going to want is monthly data for at least three to five years on prices, quantities sold, et cetera. And, so, that's where you want to put that -- you want to put that out and sensitize to legal advisors to companies or the companies

themselves. So, this is the kind of information that's going to be necessary if a proper market analysis is required.

Of course, this is unlikely in complex mergers where there are problems such as horizontal mergers, and so it's not going to be in all the cases, of course. It's going to be in cases where you're required to do that economic analysis in order to be able to come to the right and reasoned decision.

[Slide 12: Example: getting the right data matters!]

SIMON ROBERTS: While this is not a legal presentation by any stretch of the imagination -- I'm, as is obvious, an economist -- if companies are providing confidential information, of course, then the checks and balances to protect that information need to be in place, and one needs to be careful about -- about sharing that information and how that information is summarized in an analysis and maintaining a confidential report and separately for non-confidential analysis. So, that kind of goes without saying, but it's very important just to highlight it.

Just to give an example of where getting the right data matters and to, I guess, highlight the learning by doing, which we're all going through in terms of competition analysis, from a particular merger that happened in the early years in South Africa. There was a merger between two distributors of alcoholic spirits. These folks, of course were Farmer's Winery and Distillers or Distell, and there's a link to the case decision of the tribunal, which is provided on the slide.

And in this merger, the Competition Commission had followed the market definition of European -- of cases taken in Europe involving alcoholic spirits. And, so they defined the market in terms of whiskey, brandy, vodka, gin, as separate product

markets. These are separate products, they've got different characteristics, and as in other countries, this is how the market should be defined.

And this is -- this is a tempting thing to do, and -- and I think an obvious thing to do, but in this particular case, it's interesting because it highlights why there may be a danger in doing it. When it came to the competition tribunal hearing, which in South Africa is a separate stage and much of it is in public, the competition tribunal questioned this real market definition because in the documents which were discovered for the hearing, what came to light is that there had been a study done called the brandy study in a particular province of the country. And in the study, the companies had been looking at what happened when the price of brandy had changed and how responsive consumers were to brandy.

And instead of the products being quite separate, what one found or was revealed in the document was that when the price of brandy changed relative to other spirits, people switched between the spirits, and in quite large quantities. In other words, if you did a SSNIP test, you'd have found that they -- it's likely that these products were in the same product. In other words, that there wasn't a separate market for brandy and whiskey and vodka and gin because they're separate product types, but in actual fact, there was a product for alcoholic spirits and because people actually switched as a matter of fact.

And this shows why time series data on prices and quantities in different geographic areas, because this was -- as it only happened in a particular province where there'd be particular pricing behavior that had been put in place and, you know, it emphasizes why this is -- this is necessary. And it made a difference to the market

definition and the measures of concentration in the merger in fact.

In the tribunal hearing, it wasn't possible to get the time series data, but from the marketing documents of the merging parties and witness testimony, the tribunal ended up with a market definition which looked at a lower priced value market, which put all of these spirits into one market. So, there's a value market for spirits, not a separate market for brandy, whiskey, vodka, gin, et cetera.

When one moved up to where there was -- there were proprietary brands, well-known brands, typically international brands, then the tribunal decided there was a separate market, because there people were more brand-conscious. And there's a very large price difference between the value brands, the lower-end brands, and these proprietary brands, these brands which were advertised and which were -- which were known. And that had different dynamics.

And then there was a top-end market, really high premium market, of course, where you have very, very expensive whiskeys and vodkas, et cetera. And those were potentially in a separate market.

So, it illustrates, I think quite -- in a quite nice way how getting this time series data by different geographic areas is relevant for understanding these competitive interactions and why getting the right data really matters.

[Slide 13: Seeing things from the customer's perspective]

SIMON ROBERTS: Just a word before handing over to John, who's going to go into -- into surveys, as well as techniques for analyzing -- analyzing data, I'm highlighting why it's important to get out there and ask, so get out there and ask -- and ask customers particularly as well as the companies who are supplying the products.

In this response, you can do mystery shopping exercises, as well, where you can -- you can go out and pretend to be a shopper, be a customer, and obtain information, and build up at an early stage a firsthand knowledge of how prices are set and how markets work. And there's a number of examples that you can find about mystery shopping exercises.

And, thirdly, which is really at the further end, is surveys. As I said, John's going to talk about this, but this means time is required, of course, but you build up a more comprehensive data set by surveying, for example, retail outlets to understand how buyers -- buyers behave or different market participants behave depending on the case that you're looking at. And this is what John's now going to go into in more detail.

JOHN DAVIES: Thank you Simon.

[Slide 14: Surveys]

JOHN DAVIES: Most data sources on customer behavior are so-called revealed preference techniques. We're looking at data that show actual customer behavior. If we can get reliable revealed preference data, that's great. But sometimes we can't, so we need an alternative.

Furthermore, revealed preference data can only tell us about what customers actually did in circumstances they have really faced. But, often, we're interested in how customers would behave in a different situation. We can draw inferences about that from revealed preference data, but also we can use an alternative: stated preference data, from a survey.

In some cases, parties might put forward existing customer surveys, but I'm going to focus here on how competition authorities can design and use surveys themselves.

That word "surveys" covers a multitude of possible techniques, though. Some surveys are very structured, trying to reach a statistically representative sample of customers, whether through face-to-face interview or telephone surveys, or even self-completion internet surveys.

Typically, you would need a market research company to help you carry out a really large survey. At the other extreme, you and your colleagues might be calling up customers to ask them about how they use the products. In that sense, almost all competition authorities use surveys. However, our focus here is on the larger scale surveys in which you're trying to generate a data set that is large enough to produce quantitative results.

This is an area where some people are excessively skeptical in my opinion, rejecting simple numerical measures from surveys but accepting purely qualitative data. It seems odd to reject, for example, the findings of a survey of a few hundred people as being statistically insignificant, but nonetheless, to pay attention to the views of four or five customers who came in to talk to the competition authority or who were questioned at hearings.

However, you can obviously place more reliance on the numerical results of a survey if your sample is large enough and if the surveyed customer sample and the responses are likely to be unbiased. And all of this will depend upon how well you design the survey. I've suggested here on this slide some rules for good survey design.

Avoid questions that are too hypothetical. As a survey seems to give you the freedom to generate data on anything you like, such as responses to hypothetical monopolists of sets of products, it's very tempting to ask the questions we really want to

know the answer to. So, you might be tempted to try asking, "[i]f, hypothetically, the prices of all groceries in supermarkets within a ten-kilometers radius rose by 5 to 10 percent, would you go and shop in more distant supermarkets, or would you start using convenience stores?"

But no customer has ever given any thought to such a question, and the answers are likely to be meaningless. The most reliable survey questions are likely to be those that ask about what customers actually do rather than what they might do.

Try to avoid biasing the questions. The wording of questions and the topics raised in earlier questions can produce very different answers. It's always a good idea to make questions as well-defined and concrete as possible. Phrasing should not leave things open to interpretation. So, if you want to know which products customers will switch to, give them a list of options; don't just ask an open question.

And if you're asking about a 10 percent price increase, try asking about it in real prices. Say suppose the price of a ticket went up from 4 Euros to 4,40 Euros. Try to design the survey questions so that they follow the customer's own experiences.

Good market research companies understand these matters well and can advise you. And with experience, you'll get better at them over time, too.

Finally, always share and discuss the draft survey with the parties to the case and their advisors before surveying customers. One reason to do this is tactical. It makes it harder for them to challenge the methodology later if they don't like the results, but it can also be very useful. It's hard for outsiders like us properly to write questions that really capture the reality of the market.

And, also, if you have an opportunity to test some of the survey answers against

real data on how people actually behave, you should do that. Maybe you could even build a question in to which you already know the answer to test the reliability of your survey. And if it's a question to ordinary consumers, try it out on a colleague not involved in the case.

The U.K. Competition Commission, where I used to work, made a lot of use of surveys, particularly in its Phase 2 investigations of mergers. The economists and statisticians there became quite good at both the theoretical and practical aspects of running effective surveys, and there's an excellent article that two of them wrote encapsulating this experience, which I recommend very strongly. The link's there on the slide, and also in the transcript in the window below.

We're also going to see an example of the use of a customer survey later on when Arvid Fredenberg from Sweden presents an analysis of a merger in which the authority surveyed customers to estimate diversion ratios between different bookstores.

[Slide 15: Cleaning the data]

JOHN DAVIES: Before we get into analytical techniques, we need to talk about data handling and data cleaning. Obviously, if you can't be confident that your underlying data are accurate, then that's a reason to put little weight on a piece of quantitative analysis. You always need to check your data.

[Slide 16: Errors in the data?]

JOHN DAVIES: Take a look at this data set. This shows sales volumes and revenues by week. If you look closely at the weeks in June, you'll see that revenues are hugely greater than in other months, but sales are not so much increased. So, what happened in June? Did prices suddenly go up by a factor of 100? Maybe they did.

Maybe you found a cartel. But it's surely worth checking whether someone entered cents instead of dollars for those months.

Always graph your data. You might be able to see immediately if something is wrong. You certainly would if you graphed this data set.

Always look at any outliers in your data, but don't automatically exclude them either.

[Slide 17: Cleaning the data]

JOHN DAVIES: And you need to be very careful, indeed, when combining data from different sources. Do they really all mean the same thing when they use a word like "sales"? Here's a data set from another company. Do they count all of the same things in revenue as the first company did? One of them might net off taxes or other costs; another might not. You have to know exactly where these numbers come from and what they mean before you start using the data.

And then you might need to correct them, to clean the data. In this case, for example, we're going to need to convert dollars to euros to be consistent with the first data set, and we're going to need to change some other things, too.

This is not intended to be a course in data handling, but data handling is actually more important than any of the techniques we'll be covering later. If you're analyzing rubbish, your results will be rubbish.

I do want to make one point, and that's to encourage you to keep a perfect audit trail of what you've done, all the way from taking raw data to your final analytical results. Using Excel or other spreadsheets can be quite dangerous for this because you can do so much so quickly that you don't always have a good record.

Think about the data sets we just looked at. An analyst who's good at Excel will just want to dive right in and fix those problems. So, let's divide those June figures in the first table by 100. If the second table is net of taxes, let's add a column to make it consistent. Oh, and the figures in the second set are in dollars, so we'll put in the exchange rate and convert them all to euros. And let's move the columns around, as well, so they're all in the same place in both tables. You can do all that very quickly in Excel.

[Slide 18: Keep a complete log of all data cleaning and analysis]

JOHN DAVIES: But now, you have a data set that's different from the raw data and if you made an error in making it, it's going to be very difficult, indeed, to find it or to rewind your work back to where you made the error and start afresh from there. Someone trying to reproduce your results, including economists working for parties who want to show you're wrong, maybe in court, could reach a different answer from you, and it will be a long and embarrassing process to reconcile the two.

You need to keep a log of what you've done. One way to do that is to avoid spreadsheets completely, instead using a statistical package like Stata. In these packages, you can record every single thing you do -- your data manipulation, your commands, and your output in a log -- so all of your commands from data importing and cleaning through to the analysis are recorded. You can also use that log to create a file that would list and repeat all of your commands. In Stata, it's called a do- file.

To run some different analysis, you can just rerun that code to the point at which you got the clean data set and then run your new analysis after that in a new do-file. So, everything is recorded and everything is reproducible.

So, if you can get any kind of statistical package, I recommend that you use it, not because you're going to need to use a lot of technical statistical analysis but because of the log that it creates. If you really need to use Excel, try to create an audit trail by building up from one sheet to another, cleaning and analyzing using functions, not by directly changing data. Clearly annotate your changes so that you or anyone else can check and reproduce what you did to get from raw data to analytical results. Never manually mess about with the data using cut-and-paste.

Finally, get someone else to audit your work, especially any data cleaning. There are simple spot checks that can be done. If all you've done is to rearrange data, for example, then the total should add up to the same as before. And at the very least, get someone else to check in this way. For a piece of work, though, that's really critical to a decision, it's best to get someone else to reproduce the data manipulation right from the beginning, so they have to independently make the decisions you faced, rather than simply judging your decisions.

This might all seem like a lot of work, and you'll be impatient to get on to the more interesting analysis start finding answers to the competition questions. But you can waste an immense amount of time and severely damage the reputation of your authority and your own reputation if you make data errors. And it's essential to recognize that everyone makes such errors. Some people just have better systems for picking up their errors than others do.

[Slide 19: Uses and abuses – is the analysis appropriate?]

JOHN DAVIES: So, the initial data could be wrong, or you might make errors when importing and cleaning it to create data set that's wrong. Most often, though, when

there is a problem with quantitative analysis it's not that something is being measured wrongly, but rather what is being measured is not really the same as the question we're trying to answer.

Quantitative analysis usually depends on making some simplified assumptions about how the world works. For example, we might assume that the effect of a price rise on demand or supply happens rather quickly, within a year, say. But reality will be messier than that. There will be people who might not change their buying behavior that quickly in response to a price rise. If there are only a few such people, our assumption might not matter much. But if there are many, then the results of our analysis could be misleading.

One of the most common mistakes you can make is to use a technique that depends on certain assumptions, even though those assumptions did not actually fit the facts of the case. We'll discuss how to interpret and present results in some of the examples as we go along. If there is a single rule, though, it is to understand and to state the assumptions behind your analysis. You need to help nontechnical decision-makers understand how much reliance they can place upon the analysis.

There's a myth that senior decision-makers who are not economists do not like quantitative analysis. In my experience, that's not true. They love quantitative analysis because it seems to give them a simple, clear way of deciding a case. And that's fine, but if the numbers actually lead away from reality, they will result in the wrong decision, and one that the parties can challenge in court.

Your analysis of data produces evidence, and it sits alongside other forms of evidence to form the whole of the case. It's essential to allow others to see and criticize

your analysis, so make sure it is clear and reproducible. Try to get a colleague to act as a devil's advocate, querying and criticizing your analysis just as you did with the data handling.

If possible, make it fully available to the parties to the case and to their advisors. Sometimes, you'll be using confidential information that cannot be made available to everyone, but try to find ways of exposing as much of your analysis as possible without breaching your duty of preserving confidential information. For example, can the data be shared with some trusted third party, or can you anonymize the data to share it.

[Slide 20: Some Techniques]

JOHN DAVIES: Now we come to some techniques for data analysis. We're going to look at several different techniques that can be used for different purposes in a competition case. In each one, we will explain the purpose and data requirements for the techniques. Then we'll explain how to carry them out and give practical examples.

We'll try to give you enough detail so that you can have a go at using these techniques, but if you are going to use them in a real case, do go to some of the further reading that we provide.

[Slide 21: Overview]

JOHN DAVIES: We'll be covering six techniques, listed on this slide in red. With the exception of the last one on detecting cartels, they're all quite similar in their intent. They are all about determining the competitive interactions between firms that sell substitute products between firms in competition with each other.

You'll notice we don't have a section labeled "market definition," nor one labeled "competitive effects." That's because most of these techniques are useful for several

purposes. If you can find out about substitution patterns between a set of suppliers, you could, in principle, use that for market definition or for estimating the effects of a merger between two of those firms. Or, you could use that information to help you decide whether one single firm has market power, whether it's dominant.

We're going to begin by looking at what you can do if you just have prices and no other data.

[Slide 22: Overview]

JOHN DAVIES: One of the most common sets of techniques in competition analysis, particularly for market definition, looks at the relationship between prices. There is an obvious intuitive appeal to the idea that prices of products in the same market should be similar and should move in similar ways. Usually, this sort of analysis involves looking at the time series of prices for two or more products. Although, occasionally, you might look at cross-sectional data such as the price of two products in different geographical markets.

But let's concentrate on two time series of prices. It's very straightforward to calculate a correlation coefficient between two time series of prices. That coefficient will be close to one if they're closely correlated; close to zero if they're uncorrelated; and it will be negative if they're negatively correlated, as might be the case for the prices of two complementary products.

[Slide 23: Price correlation for product market definition]

JOHN DAVIES: Here's an example from an EU case, a merger between two producers of carbonless paper. Carbonless paper is sold in reels, large rolls, and in sheets cut from those reels. In this case, there was, therefore, a question of product market

definition: Are reels and sheets in the same or separate markets?

There was also a question of geographic market definition, were there separate national markets or wider, possibly Europe-wide, markets? The European Commission used price correlation analysis to provide evidence on those questions.

[Slide 24: Price correlation for product market definition cont'd]

JOHN DAVIES: Here, you can see monthly pricing data for the two types of product -- sheets on the left, reels on the right. The series look rather different, and if you calculate the correlation coefficients, indeed, they are different. The Commission reported correlations ranging from negative correlation up to positive 0.5. That raises a question: How high does a correlation have to be before we can say that products are in the same market? Unfortunately, there's no good answer to that question. Even 0.5 is quite low, and obviously correlations toward zero and less show that prices are not moving together in the manner we expect if the products are competing.

The Commission did something better to provide a benchmark, though, by also estimating correlation coefficients between products that were agreed by all parties to be in the same market -- different types of paper. Those correlations were much higher. So, as the price correlation between reels and sheets was so much less than the price correlation between different types of paper, which were definitely in the same market, this evidence suggests that sheets and reels are not in the same product market.

[Slide 25: Price correlation for geographic market definition]

JOHN DAVIES: Similarly, with geographic market definition, the Commission calculated correlation coefficients between national price trends to the two different product groups. I've just shown reels because it's easier to see. Again, we have a

problem with what benchmark to use. The Commission comfortably concluded that the correlation coefficients were high enough to say that prices in Poland, Italy, and France were correlated, and to say that U.K. prices moved very differently from those countries. But Germany was in the middle, requiring some additional evidence, perhaps.

So, we've already seen one problem with correlation analysis -- there's no good guide to how much correlation is enough. There are some fundamental problems. You can get high correlations between prices of products that are not in the same market; and you can get low correlations between prices of products that are in the same market.

[Slide 26: Price correlation: pitfalls]

JOHN DAVIES: If you have two price series, with a trend, both rising, for example, then they will appear to be highly correlated, perhaps very highly correlated, even if they are unrelated. There are lots of things that increase on average from one year to the next, and all of those series will be highly correlated.

Technically, you should only calculate a correlation coefficient between two stationary series, two that do not have a trend. So, if you see that the prices are trending, and in most countries there is at least some inflation, you can estimate that trend and remove it. Then you calculate the correlation coefficient.

I won't go into the technical details, but if you have a statistical package, what you really want to do is either remove the trend and then test that each series is stationary, then calculate a correlation coefficient; or you can test for the stationarity of the difference between two non-stationary series.

One very common reason for prices to be correlated for reasons other than competition between the products is shared costs. For example, many products use oil --

plastics and air travel, for example. They're not in the same market, but their prices will move together because their prices are both affected by common movements in the oil price.

In the example on the right, I have made up prices for two hypothetical products, which have a common cost element but which are otherwise completely uncorrelated. And you can see how similar the two series look.

So, these are all reasons for worrying about spurious correlations. We can also have situations in which there is a spurious lack of correlation between two prices that are actually competitively related. But one product might be increasing for quality -- in quality, for example, and then its price might be diverging from the rest over time.

Or we might have demand shifting between the products as the prices change. Now, that's actually quite good news, because if we can get the data, we'll be able to work out price elasticities, which is much better than computing correlations, as we'll see in a moment. But it does mean that we will see low correlation between the price series.

None of this is an absolute reason not to perform correlation analysis. It has advantages. You don't need much data, and it is a quick and simple technique. However, as always, you need to be careful how you do it and how you interpret the results. Always graph the data. Graph each series, and graph the difference between. You'll probably be able to see if there are trends.

Try to understand whether there might be common cost or demand shocks. And, as always, don't just calculate a number and assume that it's told you the answer. It tells you something, and it will help you get to the answer if done right, but it's just part of your evidence. You need to understand it and how it fits with all your other evidence as

part of your overall understanding of how the market works.

[Slide 27: Overview]

[Slide 28]

JOHN DAVIES: We can only get so far just looking at prices. What we really want to know is what's happening to demand when prices change. If prices rise, whether for a single firm or a group of firms, or even looking at the relative prices between two firms. How much will demand change? That's obviously important in understanding any firm's ability and incentive to raise the price.

Before we go into elasticities, though, I want to mention another equally important factor to take into account when thinking about those market interactions. And that is the margin that the firm makes on any additional sale, or the margin that it loses when its sales fall.

[Slide 29: Critical loss analysis: consider a price increase]

JOHN DAVIES: We talk about a critical loss -- how much would a firm's sales have to fall for a price rise to be unprofitable? When a price rises, there are two effects: sales fall, but the margin on each of those sales rises. If the initial margin on a firm's sales is very high, then only a small loss of volume, a small loss of sales, will be unprofitable. Let's do some simple arithmetic.

Suppose we're making a big margin. Suppose we're making a 50 percent margin on every sale. Then if the price goes up by 5 percent, our margin increases by 10 percent. So, if we lose 10 percent of sales as a result of that price rise, that price rise will be unprofitable.

Suppose in contrast that our margin is much lower, only 10 percent, say. Then a 5

percent price rise will cause our margin to increase by half. We could lose up to a third of our sales as a result of that price rise, and we would still be more profitable at the higher price.

And if margins are really low, if they're down near zero, then a firm might be prepared to accept almost any loss of sales if it results in a price that gives it some kind of margin. So, different elasticities can have different effects on profitability, depending on how large the margins were to begin with.

Of course, the arithmetic will not always be as simple as this. Elasticities and costs will both be different in the long run than the short. But the point in principle is an important one. A high elasticity does not automatically imply that a firm is facing a tough competitive constraint. Indeed, firms facing a high elasticity will generally have very low margins. So, you might not always be able to measure margins accurately. In fact, profitability analysis is very hard, but do at least bear in mind, when you're looking at elasticities, that margins matter.

We are not, in the end, interested in how many sales competitors can take from one another. We're only interested in that figure as an intermediate step when we're looking at how their competition with one another affects one another's profits.

So, back to determining elasticities. We want to look at how prices and quantities sold have varied over time. That tells us something about elasticities. However, there are two problems with just looking at some data series to answer that question. Firstly, there might not be very much variation in the data. You often find that prices, or sales, or both are rather stable over long periods of time. You can't get evidence on how demand and prices relate to one another unless they have been moving.

Secondly, there is the problem of endogeneity. In estimating elasticities, we're trying to find the slope of the demand curve. But the only price and quantity combinations we're ever going to see are at the intersection of a supply and a demand curve. Either of those can move. Supply curves will shift, for example, because of cost changes. Demand curves could shift because of changes in tastes or overall GDP and so on.

What we really want is to be able to identify price changes caused by shifts in the supply curve because those points would trace out a demand curve. But prices are endogenous. They are determined by both supply and demand. There are economic techniques for sorting this out involving the use of instrumental variables. We won't go into that here, but we will instead look at something simpler, natural experiments. But this approach only works if you're lucky, if you've got a good natural experiment in your data.

[Slide 30: Overview]

[Slide 31: Natural experiment "shock analysis"]

JOHN DAVIES: My example is from a U.K. competition case from 2005, when the CC was very lucky. The case was a merger of two producers of steel pilings, the structural barriers you can see in the slide. As the name indicates, these are made of steel, and their purpose is to act as retaining walls, holding back earth. There are other ways of holding back earth using structures, for example, made of concrete or stone. And those other products were not made by the merging parties.

So, market definition was very important in this case. If there was a narrow

market for steel pilings, then the merging parties' market shares would be very high after the merger, more than 70 percent, as there were very few firms selling these things in the U.K. But if the market were wider, including concrete products serving the same purpose as the steel pilings, then the market shares would be much lower, less than 20 percent if the two firms combined.

So, the merging parties were arguing that there was continuous and easy substitution between steel and concrete products, and therefore that their merged firm would have a small market share of a wide market for both types of retaining structures.

Now, this illustrates a difficulty we often face from the hypothetical nature of the hypothetical monopolist test. What's the question to which we really want to know the answer for market definition in a case like this? It's as follows. Suppose that all suppliers of steel pilings increased prices by 5 to 10 percent for a sustained period of time but suppliers of retaining structures made from other materials held prices constant. Would sufficient customers switch away from steel products to other materials to make that price rise unprofitable?

It is a very complex and artificial question because we will rarely be lucky enough to witness just the price rise that affects the products we want to increase in price, in this case steel products, without affecting the products whose prices we want to hold constant, in this case concrete and other forms of retaining structure. And, of course, we also want to be sure that the price rise we see is not the result of demand changes in our market.

[Slide 32: Natural experiment: a SSNIP test]

JOHN DAVIES: Well, the U.K. Competition Commission was very lucky. In the year before the merger, there had been a sharp increase in the global price of steel,

which is the main component of the cost of making steel pilings. Here's the chart. Industry opinion was firmly that this global price rise reflected a large increase in demand in China. So, as far as the market for retaining structures in the U.K. was concerned, it was an exogenous supply shock, an increase in the cost of one of the steel products, for no reason that has anything to do with U.K. demand. And we were told that this increase in cost raised prices somewhere between 5 and 10 percent, while the price of concrete products was unaffected.

So, what happened? If the merging parties' arguments were correct, we should have seen a significant shift in demand from steel pilings to concrete, but we did not. We saw essentially no movement at all in the demand for steel pilings, suggesting to us that engineers choose steel pilings over concrete and others for technical reasons largely unrelated to price.

So, the CC concluded that the market as narrowly defined as steel sheet pilings. And the market shares of the two merging parties were therefore very high.

That's an example of a natural experiment. Look for purely exogenous shocks to prices or other competitive conditions. Often the purpose is to conduct a critical loss test. In this example, if there had been some measurable shift to concrete but not much, we would have had to look at whether the loss of sales would have been enough to make a 10 percent price rise with costs unchanged unprofitable. So, we'd have had to do some analysis there. But there was no shift to concrete, so we did not need to.

Quite commonly, researchers find natural experiments by looking at the effect of entry or exit of large competitors on demand. If you can find clearly exogenous changes in price, that can be even better. You might get these as the result of an exchange rate

movement, for example, particularly if you're considering whether imports and domestic production are in the same market. So, have a look for any sudden shifts in the exchange rate of your currency.

Taxes can also provide a nice natural experiment in which a price will suddenly change for reasons unrelated to market conditions. Francis Kariuki, Head of the Kenya Competition Authority, will now present a case study of how his authority used a change in value- added tax to understand substitution patterns in a merger case.

[Slide 33: A natural experiment in the dairy sector in Kenya]

FRANCIS KARIUKI: In this session, as you can see from the title of the slides, is that we are going to discuss on how we can use a natural experiment to determine whether two products offer competitive constraint to each other and, therefore, we find whether they are [inaudible] in the same relevant product market.

Specifically, we are going to discuss a merger in Kenya which was in the milk processing sector. And this was in September 2013.

[Slide 34: Processed and unprocessed milk?]

FRANCIS KARIUKI: As you can see from the slides, is that 2.1 billion liters of milk are marketed in Kenya. And the marketing channels are two. One, the formal channel and the informal channel.

And out of the 2.1 billion liters, specifically 500 million liters are marketed through the formal channels, and when they talk about the formal channels, it's where we have milk processed, packaged, and put into the supermarkets and the retail stores. And the other remaining 1.5, which is 76 percent of this milk, is marketed through the informal channels, that is through milk dispensers, milk buyers, and other outlets like that

delivery to people's residences through specific milk producers.

The big question in this merger was, do these two different products offer competitive pressure to one another? And also, as a recap, is that this merger was between the largest and the fifth-largest milk processors in the country, and both of them combined for 46 percent of the milk which is marketed formally. The specific companies were Brookside Limited and Buzeki Limited.

[Slide 35: VAT Imposed on Milk]

FRANCIS KARIUKI: In September 2013, when we were under the merger, is that the government imposed a 16 percent value-added tax on the milk which is marketed through the formal channels. And, obviously, this offered a price shock to the milk which is marketed through the formal channels. And, admittedly, the prices of the milk marketed through the formal channels increased by 16 percent. This increase in price and this price shock offered what I may call a natural experiment.

What happened after that is that we saw in terms of the sales of the milk marketed through the formal channels, the fresh milk sales declined by 17.3 percent, and the long-life milk, which is also processed, sales declined by 11 percent.

[Slide 36: VAT Imposed on Milk cont'd]

FRANCIS KARIUKI: And this means that the people who are consuming the processed milk obviously shifted consuming the formal -- the informal milk, that is, the milk -- the milk marketed through the informal channels.

And based on this natural experiment, the authority concluded that, one, the milk sold through the informal channels and the milk sold through the formal channels had the same relevant market and they offer competitive pressure, and the informal -- and the

milk sold through the informal marketing channels offered is a substitute to the milk sold through the formal channels.

This led to the approval of the merger because after -- the resultant firm who only controlled 4.3 percent of the market when we take the 2.1 billion liters sold annually. Thank you very much.

[Slide 37: John Davies]

JOHN DAVIES: Thank you Francis.