Kings College London
Room 4.20 Hodgkin Building
Guy's Campus, London, SE1 1UL

Thursday, 20 March 2014

Meeting to discuss "The (Possible) Effect of Plain Packaging on the Smoking Prevalence of Minors in Australia: A Trend Analysis" working paper[1]

LIST OF ATTENDEES:

Prof. Dr. Ashok Kaul, Saarland University

Prof. Michael Wolf, PhD, Zurich University

Christopher Cox, Review Secretariat

James Collis, Review Secretariat

Lucy Edwards (Notetaker)

[1] http://www.econ.uzh.ch/static/wp/econwp149.pdf

Thursday, 20 March 2014

(11.00 am)

CHRISTOPHER COX: Well, thank you very much indeed for making the journey over to see us, we are really grateful. We should probably start by saying who we are, I know we have informally introduced ourselves but for the purposes of the transcript. So I am Christopher Cox and I am a secondee from the Department of Health and I am seconded to the review for its duration and so part of a very small team supporting Sir Cyril Chantler in his review and I should say sorry that Sir Cyril himself can't make this meeting, he has got other engagements, but we will give him a full briefing in the light of what you tell us and he will have the transcript to see verbatim what was said.

JAMES COLLIS: I am James Collis, I work as a government economist normally and I am also seconded into the review to provide economic support.

PROFESSOR MICHAEL WOLF: I am Michael Wolf, I'm Professor of Statistics and Applied Econometrics in the University of Zurich and I am part of the team working on the statistical analysis of the data to see whether there was any plain packaging effect; and we are going to talk about all of these things in detail.

PROFESSOR DR ASHOK KAUL: I am Ashok Kaul, I am a Professor of Economics at Saarland University, and I work with Michael together on this topic. My research area is mainly Applied Economics, I work as a government consultant for the several German federal and state ministries and also do some consulting for the industry.

CHRISTOPHER COX: Great, thanks, that's very helpful. I thought I should say a bit about what we are hoping to get out of the meeting. I don't know if you are familiar with the review's terms of reference, I can send you a link so you can have a look at them properly, but essentially the review is looking into potential public health effects of standardised packaging and it is in that sense a relatively narrow set of terms of reference, specific to that issue. So we would like to

ask you a number of questions, once you have told us about the work that you are doing, and I should say that doesn't necessarily indicate a particular point of view on our part, we are just trying to understand the information that is being put before us. We will send you a copy of the transcript to have a look at to correct any factual errors after the meeting.

So to start with, we would find it really helpful if you could tell us generally about the work you are doing, how it came about, how you come to be doing that work and what, if any, is the particular aim of that piece of work.

PROFESSOR DR ASHOK KAUL: So as you may know, this work was funded by Philip Morris International, and we started working on this in July 2013, not on this particular piece of work but on this funded research on the statistical effects of plain packaging. The way Phillip Morris approached us was: I was working for a big German consultancy as a technical adviser on a paper on the effects -- on the labour market effects essentially -- of the new TPD, new Tobacco Products Directive, in the EU so they got to know me, that's why they approached me, they asked me and I guess several other people whether we could do this kind of work and --

PROFESSOR MICHAEL WOLF: And he (meaning Ashok Kaul) asked me.

PROFESSOR DR ASHOK KAUL: And we have been working together for many years.

CHRISTOPHER COX: Okay. That's helpful. So obviously we were aware that Philip Morris approached us to arrange this meeting so we knew that that was how the work had arisen but do you want to take us through generally what you have done, what the scope of that work is, and then we will probably want to get quite into some detail around methodology and so on.

PROFESSOR DR ASHOK KAUL: Please feel free to ask any questions. We have prepared a research paper which will appear next week. It is a short paper and the focus of this particular research we will be presenting here is on the effects of plain packaging (or standardised packaging) on the smoking prevalence of minors in Australia. So the focus is on minors in this paper and, well, you could also analyse the full population but as you may know, one of the key

objectives of the Australian Plain Packaging Act was to influence smoking prevalence in particular of minors as in almost every country so that's why we focused on this sub-group of the population.

JAMES COLLIS: Can I just ask; the data that you are using, where does that go up to, so how recent is the latest data?

PROFESSOR DR ASHOK KAUL: So the data covers the years January 2001 to December 2013.

JAMES COLLIS: Okay. So it gives you the full year since plain packaging came in?

PROFESSOR DR ASHOK KAUL: Yes, 13 months to be precise.

PROFESSOR MICHAEL WOLF: 12 months after, well, it was introduced in December 2012 so we have that and then an additional 12 months after the introduction.

JAMES COLLIS: Okay.

CHRISTOPHER COX: Could you tell us a bit more about that data set and how that comes about, how it is collected, or if you know about the methodology of that, so we have a better idea?

PROFESSOR DR ASHOK KAUL: So what is important for you, I think, is that the data set has nothing to do with the people who funded this research, it's a standard data-set, it's called the Roy Morgan Single Source Australia data-set. Roy Morgan is a major Australian market research firm and the Single Source, the specific data-set we use here, has been brought from the so-called "establishment survey" which is a huge survey which covers many areas, one of these areas is smoking behaviour, and these interviews -- it is a survey data-set so interviews are behind this data -- our weekly interviews realised as computer-assisted personal interviews and they are administered door to door and sample about 50,000 Australians per year our data set of minors aged 14 to 17 years has a sample size of about 41,000 the average sample size is about 3,200 per year so we cover roughly a 13-year period. So from an empirical point of view, it is a repeated cross-sectional data-set, so people are not followed over time but a certain stratification procedure starts so that the data-set is representative for Australia as a whole with

regard to gender, age, and territory. And they did it very nicely, we checked that, we didn't simply believe what they did, I mean, it is a very standardised data-set, we simply used it but we checked with regard to the stratification procedure and compared it to the Australian census, for example, and they did very good work.

CHRISTOPHER COX: Do you know, is that data-set used for any other purposes or -- it would just be helpful to have some insight --

PROFESSOR DR ASHOK KAUL: Yes, in consumer research, it is used in consumer research and also it has been used, for example, by Wakefield, a very renowned Australian tobacco control advocate, in her research on tobacco control in 2008 and 2009 in published papers, in peer-reviewed published papers.

CHRISTOPHER COX: So in your view it is a well-regarded data-set?

PROFESSOR DR ASHOK KAUL: Yes, indeed. We have reviewed several data-sets because it was not the case that Philip Morris asked us to do the research with the Roy Morgan data-set but they gave us the option to have look at all the data-sets that were available and this was the best data-set we could find.

CHRISTOPHER COX: In terms of what you have done with that data, it would be helpful to know is, for example, the method that you have used a sort of standard, well-recognised method, is it a peer-reviewed method, and I guess the other question, just to be clear on, is this sort of analysis your particular speciality in this area?

PROFESSOR DR ASHOK KAUL: So maybe we can walk you through the presentation so that you can see the methods.

CHRISTOPHER COX: That would be helpful, please.

PROFESSOR DR ASHOK KAUL: We can also explain to you how standard our methods are and whether we are experts in these methods and what alternative methods would have been. We have focused on this in quite some detail in the paper. What we suggest is -- we all have the

paper but I also will open the presentation here (on the computer) so that you can have a look at the graphs, so you can read at the same time and see the graphs, a couple of graphs in the appendix. So Michael, do you want to start with the section 3.2?

PROFESSOR MICHAEL WOLF: Section 3, yes, so we fitted a very standard, very simple time series model that would explain how -- first we computed observed prevalence in a given month because we always have a sample of minors that are either smokers or not so there's a zero long variable: smoker no or yes.

PROFESSOR DR ASHOK KAUL: Let me start with the research question, so it is about smoking prevalence, you could analyse different things, smoking intensity, for instance, so we analysed smoking prevalence, the most important variable, and the question we asked is whether there is empirical evidence showing that the implementation of plain packaging in December 2012 changed the smoking prevalence of Australian minors. That's the research question. To do this, we did a trend analysis and Michael is going to go into the details.

PROFESSOR MICHAEL WOLF: First the computed observed prevalence in a given month as the average of the zero-one variable (smoker) so just as the proportion of minors in a month that were smokers and then we did a trend, a time series analysis, and we looked at how that developed over time and maybe it's useful to look at figure 2, that shows the time series development of this variable.

PROFESSOR DR ASHOK KAUL: From January 2001 to December 2013.

PROFESSOR MICHAEL WOLF: So one can see a general downward trend with quite a bit of variability around the trend and the variability can be explained by the fact that every month there is a new sample, so it is not the same people followed over time, then it would be much less ragged, so it is clear there will be some depending on, you know, randomness, sometimes above and sometimes below the trend; and then we fitted a very simple linear time trend by modelling prevalence as a linear function of time and then we fitted that by standard minimising

least squares methodology.

JAMES COLLIS: Can I just ask how you are using the data, you said it is a different sample every month, are you treating it as a sort of quasi panel data-set so --

PROFESSOR DR ASHOK KAUL: It is a repeated cross-section. That is important. We could build a “pseudo panel” if you want but the thing we do here is we exploit the full data-set as a repeated cross-section as it is and the method we have picked, because it is a repeated cross-section and sample size varies over time.

PROFESSOR MICHAEL WOLF: After sort of condensing all of these people in a given month into one number, the average, your prevalence then becomes a univariate time series so there is no longer any cross-sectional dimension, it's just one number for a month, and so then it is just a standard univariate time series analysis. So we fitted that line by least squares and then maybe if we go up to figure 1, so one can see that the sample sizes are not constant over time. There is a general slight downward trend, but even, you know, it is not the same number in a given month and so to take that into account and get the most efficient estimate we use a very standard technique called "weighted least squares”, so in a given month we used the number of observations as a weight so those months that had a higher number, we give more weight, are given more weight, and those months with a lower number are given a bit less weight, so it extracts the most information possible from the data.

CHRISTOPHER COX: Do you have any particular reason as to why the sample size varies so much over time?

PROFESSOR DR ASHOK KAUL: No, we don't know.

PROFESSOR MICHAEL WOLF: That is up to the people collecting the data, I have no idea.

CHRISTOPHER COX: I am just wondering if there is any --

PROFESSOR DR ASHOK KAUL: There is no sample attrition because it is not a panel but --

CHRISTOPHER COX: No, but also whether you would have any way of knowing or correcting for

any sort of -- I hesitate to use the word "bias" but sort of systematic change in the sample related to changes in size --

PROFESSOR DR ASHOK KAUL: No, that is taken care of because in the stratification procedure they make sure -- that is the important point, of course, in a panel data-set it could happen because of sample attrition, that there is a systematic bias evolving over time, that can't happen here because of the stratification procedure, that's very important.

CHRISTOPHER COX: So you mean it still remains a representative sample even though it is smaller?

PROFESSOR DR ASHOK KAUL: Exactly.

JAMES COLLIS: Would it be to do with response rates at all, are they aiming to have the same sample each time but sometimes they get a lower response rate?

PROFESSOR DR ASHOK KAUL: Well, it's a computer-assisted interview where they walk door to door so they can make sure that the sample size stays at the rate they want, so we don't know why, maybe it's a deliberate decision to sample less people but we don't know why they did it, we take the data as they are.

PROFESSOR MICHAEL WOLF: If we go back to figure 2. Then you can see this dotted line in whatever colour that is, this magenta, you see it in numbers and equation 3.2 so that is the time trend that we fitted and it fits the data really well, so we did all kinds of diagnostic checks whether we should need a higher polynomial but that very simple line does a very good job in modelling. Of course, it cannot capture these up-and-down movements, that is not possible, it's due to randomness; but it captures well the general down movement and in particular there is about a 0.44 decline in percentage points on average per year; so the smoking prevalence has been down-trending since 2001 at an average rate of roughly 0.44 percentage points.

CHRISTOPHER COX: That is amongst 14-17?

PROFESSOR MICHAEL WOLF: Right, 14 to 17. So after having done that and what is of key

interest if we look at the very end, how does observed prevalence deviate from that trend line starting from December 2012 and then to really obtain it, we have to focus, we have to kind of key in on that part to see -- get a clearer picture. And then we computed deviations of the observed prevalence from the trend line and that is done in figure 3. So these are the deviations.

PROFESSOR DR ASHOK KAUL: So here the X-axis is now timed from December 2012 to the end of December.

PROFESSOR MICHAEL WOLF: So a negative number means that the observed prevalence was below the fitted line and a positive number means that it was above the fitted line, in particular we see that there are 13 numbers; seven are negative and six are positive and if we compute an average of these 13 numbers then it is minus 0.41. So on average some would be low, the deviations on average are somewhat negative, meaning that on average the observed prevalence lies somewhat below the fitted line and then a naive interpretation would be that plain packaging has led to a reduction on average by these about 0.41 percentage points but then the statistician steps in and says well, it is an estimate so really we have to go beyond it; we cannot equate an estimated number with a true population effect, so estimation is not equal to truth and we have to go beyond that and add more information to the pure estimation. Estimations are a first step but we can't stop here. So we did two things, first just an informal check, if we compare before and after, so we have December 2012, we have one year after that, so after the policy, but we can also look at one year before. And if you look at it just as an informal analysis, just by looking at it, there doesn't seem to be, we are going to go beyond it, but at this stage right now there doesn't appear to seem any difference, before and after it's almost like a mirror image.

CHRISTOPHER COX: So just to clarify my understanding: at this point you are saying you have looked at the deviation from that general downward trend line and the first chart you put up

shows, you know, there is quite wide and random variability.

PROFESSOR DR ASHOK KAUL: Right.

CHRISTOPHER COX: And on this chart, it is not dissimilar in terms of how much deviation from that line there is before and after.

PROFESSOR DR ASHOK KAUL: Exactly. If you look at figure 4 and remove the red line which is the intervention date, that looks very similar 12 months before and 12 months after.

PROFESSOR MICHAEL WOLF: It would be difficult to guess from the plot if there weren't the red line to guess when did the intervention actually take place.

CHRISTOPHER COX: And the difficulty remaining of how much of that is essentially due to randomness and so on.

PROFESSOR DR ASHOK KAUL: That is the statistics, that is still looking at --

PROFESSOR MICHAEL WOLF: Staying with an informal analysis comparison, if you look at -- the average after the introduction was minus 0.40 (typo corrected) but the average before the introduction over the same number of months, 12, is minus 0.60 (typo corrected) so it is actually a bit lower-- if one were naive and equated, you know, estimates with population numbers then one would have to say that the quote on quote effect of plain packaging was more in the year before it actually took place compared to the year after.

PROFESSOR DR ASHOK KAUL: Which, of course, doesn't make sense as a statement.

JAMES COLLIS: I suppose the other challenge here being how you separate that out from other effects.

PROFESSOR DR ASHOK KAUL: Sure, right.

PROFESSOR MICHAEL WOLF: Right. Only -- here we did very similar, we only have the general time trends so we did not include any potential further variables so -- but this is the first step and also with the samples as we have, it would be very difficult to include a lot of further --

CHRISTOPHER COX: I don't know if maybe you want to continue and we will come back to the

question but so far you have looked at the data over time which has the relationship in a sense to plain packaging insofar as you have a line saying this is when this intervention occurred rather than anything arising from this data-set per se, to suggest that there is a link with standardised packaging. But I think it might be better if you carry on with your presentation and maybe we will come back to some of those questions.

PROFESSOR MICHAEL WOLF: Now, we go back to figure 3, so in a sense we can look at these numbers as estimates of treatment effects, so there's a negative number meaning prevalence was below the fitted trend lines, or one could take this as an estimate of plain packaging in a given month. So according to the estimation, seven were negative and six were positive, but then these are only estimates and we want to add more information and do a statistical inference and we can do that by adding confidence intervals to these estimates; and maybe if we can look at those first and then we can talk about how those were computed and what else could have been done. So what we did here --

PROFESSOR DR ASHOK KAUL: So that is figure 5, the same graph we saw before, plus the confidence intervals we added --

PROFESSOR MICHAEL WOLF: So what we have done in the given month, we have added a confidence interval at 90 percent confidence level and so as inference, if the entire interval were to lie below the horizontal line, if the entire interval was negative, then that would be a statistical significance indicating indeed an effect in the desired direction. If on the other hand, the entire interval lied above the horizontal line, that would be a statistical significance that the thing (meaning the introduction of plain packaging) backfired and did the opposite of what was intended; however, what we see in every single month is that the interval is both below and above, meaning that the lower end point is negative and the upper end point is positive, so the zero is always in the confidence interval, meaning there is no statistical significance for any effect, be it positive or negative. So I think it is important here to add just in case that we -- it is

statistically impossible to prove that there is a zero effect, we cannot prove a zero, so what we say here, there is no evidence for an effect, which is not the same as evidence for no effect. But evidence for no effect statistically cannot be done. By the laws of statistics, you cannot prove that it's zero. You could prove that it's not zero, you can prove it is negative if the entire interval is below the line --

PROFESSOR DR ASHOK KAUL: At a certain confidence level --

PROFESSOR MICHAEL WOLF: Right, or you could prove it is positive if the entire interval is above the line. But you cannot prove zero. So to interpret a given confidence interval, any of these numbers in the confidence interval is a plausible number at a chosen level, which is 90 percent, in particular zero is one of them but zero is one of many, right, so we cannot prove it is zero but we don't have evidence that there is a negative effect in any given month since zero is always in the confidence interval.

JAMES COLLIS: To what extent do you think that reflects the large volatility in the data? So when you showed us your graph there is a lot of noise of the data --

PROFESSOR MICHAEL WOLF: For sure, exactly, that noise in the data, like a typical magnitude of the deviation from the line, is very much reflected in the width of the confidence interval. If the deviations had been smaller, the confidence interval would be shorter, so of course that is the key factor going into -- the key factor in determining, in addition to the estimation uncertainty about that line -- because it is the estimated line, not the true line, but that's a second-order effect -- so the first order effect determining the width of the confidence interval is the typical magnitude of the deviations from the line. But there's nothing we can do about that because the data are by their nature quite noisy and we have to take it into account and that is reflected in the confidence interval.

PROFESSOR DR ASHOK KAUL: Then on the other hand, the data-set is not small so we have some power.

PROFESSOR MICHAEL WOLF: Then we can work through how we did this and what variations could have been done. But I can say from upfront the methodology that we have employed is the one that gives the most leeway to finding an effect, if there had been any, in the sense that the methodology we employed gives the shortest intervals. You know, of any reasonable methodology anybody would ever employ, it gives the shortest confidence intervals, meaning it gives the highest possibility of getting an interval that is entirely below the line or entirely above of the line. And even if you are giving the most leeway in that sense, still for every single month the zero is contained. So, again, we are not claiming there was no effect, we are saying there is no evidence for any effect; I just want to make it very clear.

CHRISTOPHER COX: I would just be interested on -- with calculation of the confidence intervals, I don't claim to be a statistician, we do have some people assisting the team who are, but just for my own information; are there any particular assumptions involved in calculating those confidence intervals that are relevant in a sort of - assumptions about the real world, if you like, and standardised packaging?

PROFESSOR DR ASHOK KAUL: You will find the details in the paper in section 3.2.3 and Michael will walk you through the details of the assumptions behind the calculation of the confidence intervals in figure 5 and in particular, how the confidence intervals would change if we make changes in the assumptions that underlie our analysis.

CHRISTOPHER COX: That would be really helpful.

PROFESSOR MICHAEL WOLF: First, I mean, there are some changes that can be made and there are some changes that do not depend on any assumptions and others do, so maybe let's walk through these bullet points on page 5. The first one, and there is no assumption here whatsoever, so we have chosen the confidence level to be 90 percent and that's the lowest confidence level anybody would ever apply in empirical work. More typical is 95 percent and 95 percent would make the intervals wider. So we have chosen the smallest, reasonable

confidence level 90 percent, nobody would do 80 or 60 or 50, so 90 percent is the smallest one anybody would ever use, so in that sense we have given the most leeway by making the intervals as short as possible in terms of the confidence level. Next, what we have done, we have just used the standard text book methodology, anything that is programmed into a standard statistics package, and that makes a key assumption, namely that the deviations from the line have a certain parametric distribution, namely they have a normal distribution and that's a key assumption going into determining the width of the interval. If the distribution is not normal, then that could make the intervals shorter or wider, in particular if the distribution has heavy tails, that would give a wider interval, if that were to be taken into account. If the distribution has short tails compared to the normal, it would make the interval smaller.

CHRISTOPHER COX: Could you just maybe put it more in lay man's terms about what that would mean in practice, the tails?

PROFESSOR MICHAEL WOLF: The tails meaning the probability of observing very small or very large numbers, so the normal distribution -- obviously you have to think about a distribution that is standardised, that is standard deviation of one. You can take any distribution and scale it so it has a standard deviation of one and means zero and then also probability of observing a number larger than three or larger than four and so that heavy tail distribution means compared to the normal distribution, there is a larger probability of getting a very extreme number.

CHRISTOPHER COX: I see your colleague is drawing me a Bell curve!

PROFESSOR DR ASHOK KAUL: That's a Bell curve and this is a different distribution.

PROFESSOR MICHAEL WOLF: Which is flat in the middle and has heavy the tails. And just to give an example, if you take financial returns, on a daily level, financial returns have very heavy tails, they are not normally distributed, so it can happen that you get a very small or a very large number with a considerable probability that is not consistent with normal distribution.

CHRISTOPHER COX: Just to try and crystallise it for my understanding. In practice what would that signify? The different shape or distribution in terms of smoking behaviour amongst youth; what would that actually mean?

PROFESSOR DR ASHOK KAUL: No, we have to be more precise here, it is not about the smoking behaviour, it is about the errors as a deviation from the trend so that's something very technical but it's not an assumption we make. In figure 5 we make this assumption, but then we can have a look at the data, whether we are in the normal case or whether we have fat tails or light tails. So we can check that.

PROFESSOR MICHAEL WOLF: So the key here is this is an assumption, it doesn't have to be taken for granted or, you know, in blind faith. It is an assumption that can be checked by looking at the data, and this is what we did. Just, you know, before we go through that check, I just --

PROFESSOR DR ASHOK KAUL: Back again?

PROFESSOR MICHAEL WOLF: Yes, go on. As I say, if that assumption of a normal distribution of the error terms you know, doesn't hold, that could make the intervals wider, it could make them shorter, or it actually could leave them more or less unchanged. In particular if you have some, you know, very heavy-tailed distributions, like big deviations are more likely compared to normal, then that would give us a wider interval and the methodology to compute such intervals, taking into account the distribution, is there. It is called the "bootstrap method", it is computationally very involved and, you know, we could do it, and if you ask us to, we are going to do it. But we wanted to keep it simple here so that other people can replicate our work; but we checked, you know, if we did it what change that would make. Would it make the intervals shorter or wider and that's what we did and we did an analysis of the residuals, that is, the deviations from the fitted line.

PROFESSOR DR ASHOK KAUL: So this is figure 6 which plots the so-called normal QQ plot of the residuals of the fitted model.

PROFESSOR MICHAEL WOLF: So meaning if all these blue dots were on the line more or less -- it's never going to be a hundred per cent perfect -- but if they are on the line, that would indicate a normal distribution; and deviations from the line, systematic deviations, indicate a non-normal distribution. And what we see is that they are above here on the left and more so on the right, there are above the line, and that indicates two things: first, the tails are heavier compared to the normal distribution so the heavy tails would give us, if you take it into account, would make the intervals somewhat wider and then there is also in addition some skewness, so it's not quite symmetric, it's slightly skewed to the right, also in the computed summary statistics. Both of these effects are not dramatic but it's somewhat heavy-tailed, it is somewhat skewed, and if you take it into account, and we can, if you ask we will, but the effect that it will have, it will make the intervals wider, not dramatically wider, somewhat wider. So, again, in that way we didn't feel the need to do it, since we are doing something that is somewhat incorrect but incorrect in making the intervals too short and so, again, in that way we are giving leeway to finding a possible effect by having the intervals arguably a bit too short. So the confidence level isn't even 90 percent, it is somewhat below and it’s not important whether it is 87 or 86, it's somewhat below 90 percent.

JAMES COLLIS: Just on -- when you say it's slightly skewed just to bring that back to what that means in terms of the actual tobacco data, is the tail skewed towards lower than average prevalence or higher?

PROFESSOR MICHAEL WOLF: Here it is slightly skewed to the right which means, you know, large numbers above the line are a bit more likely than large numbers below the line so to go back to the financial data, they are actually typically skewed to the left meaning a large loss is a bit more likely than a large gain, sometimes you have a bad day in the market, we have very good days but the very bad days occur a bit more frequently so they are skewed to the left and here it's not very strong so the large deviations above the line occur a bit more frequently than large deviations below the line, but it's not a dramatic effect.

PROFESSOR DR ASHOK KAUL: It's very small.

PROFESSOR MICHAEL WOLF: So it's not quite symmetric.

JAMES COLLIS: Does that reflect at all -- obviously in terms of what the prevalence is you have got a minimum that can be at zero whereas the maximum is a hundred, which is much, much further up.

PROFESSOR DR ASHOK KAUL: Which doesn't happen very often either.

PROFESSOR MICHAEL WOLF: So there is sort of more room to be above than below and already this trend line, I mean, we are already at the very end, we are at six per cent so it is getting pretty close to the lower, so it can't go below zero but there can be a larger, positive shock with a bit more probability.

CHRISTOPHER COX: Can I ask, you may have already answered this, but do you think the changes over the sample size over time would have any bearing on the discussion we have just had? Would you have more confidence in the results if sample size had been maintained as it was earlier in this series?

PROFESSOR MICHAEL WOLF: I think it would have been -- the ideal situation is that the sample size is constant and we don't have to use the weighted least squares, we can use the ordinary (least squares), that would have been somewhat preferred. But I think the differences are not gigantic and then they (the sample sizes, not the differences) have been also pretty stable from 2008 on so it is not that -- just at the very end they would have been much lower than throughout the period. So it is not -- to me it is not a cause of particular great concern.

CHRISTOPHER COX: Do you want to carry on?

PROFESSOR DR ASHOK KAUL: Yes. No, we have a couple of more points where we discuss the widths of the confidence intervals. So we stopped at the second bullet point. Yes.

PROFESSOR MICHAEL WOLF: So the third bullet point again, well, we want to be very, you know, safe and thorough, well, the data has been collected over time and what we have when we

computed the confidence intervals, also the assumption that we made is that the deviations from the line are independent; so even though the data is collected over time, it is plausible that the deviations from the line are serially not dependent. And that assumption goes into determining the width of the confidence interval, although it would be a second-order effect. If there was anything and we took it into account, it would be a very small change. Still just to be very, very safe we then did a time-series analysis of those residuals and we can go to figure 7. So we looked at the so-called auto-correlation function and the partial auto-correlation function. But the key thing here is that there are these little bars and there are -- no, that's too small, leave it as it is -- maybe just go to the first one. All that you sort of have to understand or take into account at this point, there are bars and there are these horizontal lines around the bars; and if there were bars sticking out above or beyond the horizontal lines, that would be an indication that there is a time series nature of those deviations from the line. But in both plots, in the ACF, the auto-correlation function, and the PACF, the partial auto-correlation function, all of these bars are inside the lines and so there is no indication of any time series nature to be taken into account.

JAMES COLLIS: Is that true even though the first half of these series, the bars are always above the line generally and the second half they are always below the line?

PROFESSOR MICHAEL WOLF: Yes, but that's sort of meaningless random noise; the key is whether you find one sticking out, above/below, I mean, no, that doesn't come into play. If you had some -- anything that, you know, sticks out of those horizontal lines, then you start interpreting that. But anything within the lines and to the statistician sort of -- I mean, that's sort of the end of the story.

PROFESSOR DR ASHOK KAUL: So that was the third bullet point showing us that it was not necessary to correct the assumption of independent IID error terms.

PROFESSOR MICHAEL WOLF: As I say, that only would effect that uncertainty about the fitted

trend line which already is only a second-order effect; and then we tune around with the second-order effect to some tiny degree. Overall, the difference would be meaningless so we did not even go through that, but if you ask us to do it, we will, I can say right now with a hundred per cent certainty you wouldn't be able to see -- I mean, it's a meaningless difference that would come out from that but now the last bullet point, that's an important conceptual difference. If you go back to plot number 5. Below -- let's make it maybe a bit bigger. Above, above. Yes, number 5. Good, good, good.

PROFESSOR DR ASHOK KAUL: So we are back to the deviations of observed prevalence from the linear a trend, plus the confidence intervals added from our simple methodology.

PROFESSOR MICHAEL WOLF: Right. So now the important thing -- that really is an important distinction -- these confidence intervals are what the statistician calls "pointwise", so we have pointwise, one month at a time, we have the 90 percent confidence about the true effect being within the interval. But it's one month at a time and it's really -- that is the methodology that should be applied if one only is interested in the potential effect in a given, pre-specified month before looking at the data. So, for example, I mean, it would be natural to me to be interested in December of 2012, as a month of natural interest, so if the policy maker says, "I am interested in the potential effect of plain packaging in December 2012".

PROFESSOR DR ASHOK KAUL: Because it was a lot in the media so the question with the natural question: is there a one-time effect at least?

PROFESSOR MICHAEL WOLF: Then the right thing to do is to look at this pointwise, just focusing in on December 2012 with a confidence level 90 percent; and so what we have done is one month at a time. On the other hand, and I am sure some people would take a different stance and say, "Is there any effect at all? Allow me to look over those 13 months and, you know, search for an effect", and the statisticians call it "data mining" or "cherry picking". You know, I don't have a particular month in mind upfront, going to look at the data and look where do I see

the most effect; and actually it wouldn't be December 2012, it would be October or September 2013, it's August actually that has the smallest estimated effect, if you will, at minus 2 per cent or a bit below. And then if you go looking, if you go looking for something significant then the statistician would tell you that you should not compute the pointwise confidence intervals, you have to make it uniform. So not one month at a time have that 90 percent confidence but together globally; not locally in a given month, but globally over the entire period it should be 90 percent and that methodology is out there; it would make the intervals much, much wider, so not just a tiny bit, it would be really a big difference. So if you did it in a uniform fashion -- which from a statistical point of view, you know, if you want to convince me there was an effect after looking at the data, I would say you have to do it uniform not point wise -- if you did that, again it would make the intervals much wider. So even by only doing it one point at a time, and even then we still sort of do it incorrectly, the cherry picking, data mining, looking for an effect, and even then nothing shows up. So that's the last variation that we listed. Again, also, if you want us to do it, we will do it, but what would be the point? The point is that all these things get much wider and what is of interest, whether zero is in the confidence interval or not. And it was in there before all the time, it will be in there after all the time as well; so there is no point --

CHRISTOPHER COX: I don't think it is for us to commission any research, thank you all the same. I think that's for your clients.

PROFESSOR MICHAEL WOLF: I think I have nothing further to add at this point, but any questions --

CHRISTOPHER COX: Unless you have any burning questions, James, I just want to go back to a fairly basic question that we may have skirted over slightly but remind me what exactly the data is purporting to show? We have been talking about prevalence which could be a slightly ambiguous term and I recall you saying it was on a sort of binary basis that the data is put into

that model, but what does prevalence actually mean in this case? Is this ever smoked, smoked in the last week?

PROFESSOR DR ASHOK KAUL: I would have to look up the exact definition, there are several different definitions and as far as I remember, they stick as closely as possible to the official government definitions. I think it should be something like: "Do you smoke? Yes or no", and you count as a smoker if you smoke either on a daily basis or more than once a week, something like that, we could look it up in the questionnaire.

CHRISTOPHER COX: That sounds very familiar in terms of I think the definition of a regular smoker is someone who smokes at least once a week so it would be helpful if you could just confirm that actually but --

PROFESSOR DR ASHOK KAUL: I will write it down.

CHRISTOPHER COX: That's helpful to know. The root of my question was whether there is an effect within prevalence that wouldn't show up on this because it is a sort of binary approach --

PROFESSOR DR ASHOK KAUL: We will check that. Yes.

CHRISTOPHER COX: It might be for instance someone was smoking ten a week and changed to five a week, let's say or, indeed, vice versa, you wouldn't get at that from this data?

PROFESSOR MICHAEL WOLF: No, you don't get it from that. You probably would have to look at smoking intensity and that is a different exercise, so we don't address that.

PROFESSOR DR ASHOK KAUL: For a reason we don’t know in the tobacco control literature 90 per cent of the focus, I would say, is on smoking prevalence. As an economist, of course, I would also care about smoking intensity, that is conditional on smoking, how much do you smoke, and this number has also changed in all the OECD countries, as far as I know, over time. Even heavy smokers smoke less than they used to do 20 years ago. We know that.

PROFESSOR MICHAEL WOLF: From a statistical point of view smoking intensity could go up. If

those quit who are smokers but smoke very little, then among the group of smokers, intensity would go up. I have no idea, we haven't analysed the data; it could go up, it could go down or it could remain unaffected.

PROFESSOR DR ASHOK KAUL: For a reason we don't know, all these anti-smoking measures tend to focus on smoking prevalence, that's why we focused on smoking prevalence as well, but the data don't -- well, the data are richer, you could do more.

CHRISTOPHER COX: Did I understand correctly that the sample changes literally every month?

PROFESSOR MICHAEL WOLF: Yes.

CHRISTOPHER COX: So it would be just helpful to know a little bit more about how the sample is identified, because presumably it is to be balanced across that age group, you would need equal numbers of 14 year-olds and 17 year-olds each month, otherwise you would end up with a skew, wouldn't you, if, say, smoking uptake was more intense amongst 17 year-olds than 14 year-olds.

PROFESSOR MICHAEL WOLF: From a statistical point, it would not be desirable that in one month they are all 14 and next month they are all 17. I don't know. I was given the data to analyse. How the data was collected, I cannot make any statement here.

PROFESSOR DR ASHOK KAUL: The stratification procedure, the way they put it and the way we checked it, is as good as you can do it, I mean, there is no reason to have any suspicion on the stratification procedure, it is very well done and very well documented.

CHRISTOPHER COX: But you mean by that that you do end up with a sample that follows whatever pattern of split between different age groups.

PROFESSOR DR ASHOK KAUL: As closely as they targeted it. We will have to look up again in the questionnaires how closely they targeted it but from what I remember, and we check that in quite some detail, they tried to have a very, very detailed stratification procedure.

CHRISTOPHER COX: That is helpful.

JAMES COLLIS: Just a couple of questions on the stuff you were showing us. So your trend line of looking at changes against, is that purely like a time series trend, you are not trying to control for any other explanatory variables?

PROFESSOR MICHAEL WOLF: No. It is a very simple, purely time-series analysis, not controlling for anything else. That could maybe be done at a later step, you know; we have to take it one step at a time. It is a natural starting point and then one can get more fancy going on from here but, you know, one step at a time.

JAMES COLLIS: You say one step at a time, would your view be it would be too soon to try to do that sort of thing, given the amount of data you have so far?

PROFESSOR DR ASHOK KAUL: No, no, but it takes some time and when we started with this research, this is the first output we produced and obviously it is easier to produce a time series paper based on a univariate time series than doing a full-fledged micro-data analysis. It takes some time.

PROFESSOR MICHAEL WOLF: I was given the data at the end of January, you know, so it takes time.

PROFESSOR DR ASHOK KAUL: We wanted the December data to have a full year so we started this paper and it was finished three weeks ago so the next obvious step, which has not been done yet, is to analyse the full-fledged data-set.

JAMES COLLIS: One or two of the ideas that spring to mind is you could look at just putting in dummy variables for the change --

PROFESSOR DR ASHOK KAUL: Of course.

PROFESSOR MICHAEL WOLF: In a full-fledged micro-analysis, you want to throw in things: income, occupation, education, etc.; but all these things make much more sense if you look at the full sample, not just minors. You want to throw in age but for the minors it is all the same, and the income is all probably zero, and the education is -- so there's little variation in those

interesting other variables so a full-fledged micro-analysis would call for the entire population, not just the minors.

CHRISTOPHER COX: Did you have any other questions James?

JAMES COLLIS: No, I don't think so at this stage other than -- so you were showing us the sort of confidence intervals when you don't see any statistically significant change. Is it possible within the framework you have to sort of test explicitly for a structural break around December or is what you showed us basically equivalent to doing that?

PROFESSOR MICHAEL WOLF: To do again, with the univariate time series data, that would be equivalent to doing that. To really test for a structural break it would from a statistical point of view make much more sense to analyse the micro-level individual data and then you have a much larger sample size as you do not condense all the people into one single number; you keep the individual information and then testing for a structural break and also, you know, looking at the form of a possible structural change, what kind of shape did it have if any, is really much more useful to do the micro-level analysis. But then again, it probably would require to look at the full population and not just the minors.

PROFESSOR DR ASHOK KAUL: And it would take some more time.

PROFESSOR MICHAEL WOLF: That would take more time, yes, surely, that's much more involved.

CHRISTOPHER COX: Where does your research go from here then? That is your next step, is it, to do that?

PROFESSOR DR ASHOK KAUL: That would be the next natural step.

CHRISTOPHER COX: I just wondered if there was an issue about, over time, would you expect to come to a more secure or different conclusion, you know, are there additional elements that you can build in that would --

PROFESSOR MICHAEL WOLF: It would be good to have ten years more of data, but I think the people want answers now or at least they want the first answers now.

CHRISTOPHER COX: Presumably in ten years you would be looking then at a sort of longer term effect whereas what you have been able to do so far is really -- it is almost looking, well, it is not a negative, you are looking for a distinct sort of kink in a curve at a particular point in time rather than --

PROFESSOR MICHAEL WOLF: Right. So if we are interested in, you know, has there been any effect within the year after the introduction, then we have all the data that we are going to get.

CHRISTOPHER COX: But you might be able to look over a decade; is there any discernible effect.

PROFESSOR DR ASHOK KAUL: Sure, but obviously it is impossible now.

PROFESSOR MICHAEL WOLF: Well, if somebody says, well, even, you know, it was introduced now but it will take five years to take an effect, that -- we cannot answer that question based on the data. So the question we can address, you know: has there been any effect within that year after the introduction? And that we can do.

CHRISTOPHER COX: I think I was more, wondering what would be possible in the future rather than -- I mean obviously you can only do what you can do now because you have got the data that exists now, but what might exist in the future that would enable you come to a fuller answer, if you like, and I guess the really knotty question is: how do you differentiate any effect, if there was one, that you attribute to standardised packaging as opposed to, say, increases in taxation?

PROFESSOR MICHAEL WOLF: That's a very good point because we don't take these other variables into account in this simple time-series analysis. But if we have the -- we can analyse the micro-level data and we have that information, the other variables that we can throw in in addition, such as gender and age and income, occupation, the region where people live, we have the information of the taxes there were perhaps other policy interventions, banning smoking outdoors -- or indoors. All these things can be also thrown into the model and then we can control for the other variables and compare to the plain packaging effect, if any.

JAMES COLLIS: Given it's obviously not that long since the policy came in, we only have about

a year or so of data and like we were saying, there is a lot of noise within that data, you would think in order to find a statistically significant impact, it would have to be a fairly big, immediate impact. Do you have any feel from the work you have done for how much the prevalence would have to change for it to be picked up as statistically significant, given how much the data varies randomly?

PROFESSOR MICHAEL WOLF: I could not give a specific number right here. Obviously, a very, very, very small effect will never be picked up. You can always argue that if you say there is no evidence but maybe there was a tiny, tiny effect. And that is always true, if there was a very tiny effect and unless your sample size in gigantic, you will never be able to detect a very, very tiny effect.

PROFESSOR DR ASHOK KAUL: Yes, but then on the other hand what is a little bit surprising, I guess, is that from a purely descriptive point of view, you don't see anything. Even if you are not talking about statistical significance, I mean, sort of hoping for an effect, even a small one, you would at least expect something, a little, tiny effect, 12 months after in comparison to 12 months before but we don't find that, we find the contrary.

JAMES COLLIS: I suppose the difficulty being here, you were saying that the previous 12 months wasn't necessarily a typical 12 months either, so given the tax changes that came in.

PROFESSOR DR ASHOK KAUL: Yes, but we had a tax change there as well, I think, in December 2012, so if at all that plays -- if at all that would kind of attribute something to plain packaging which was in fact a tax change so that would make things even worse. So suppose plain packaging has no effect but the tax increase has the obvious effect, namely it induces a price increase then we would hope to observe a drop in prevalence. We don't observe it either. That's the surprising thing actually.

CHRISTOPHER COX: I think we have a separate list of tax changes we can look at.

PROFESSOR MICHAEL WOLF: I mean, all this is only for the minors, you know. If we look at the

full population, the story may well be different, it may not. So we are not going to make any statements about the full population at this point.

JAMES COLLIS: Is that something that is on your next steps to do?

PROFESSOR MICHAEL WOLF: Yes, sure. Yes. For a statistician it is much more interesting, individual data, you can run some more complicated -- I mean, it's sound methodology but it is very, very simple so (meaning the time-series analysis we did in the report on minors) -- and the other stuff can get quite a bit more complex and intellectually a bit more challenging.

CHRISTOPHER COX: So to come back to your conclusion, and I think you put it very clearly earlier, about, if I understood it correctly, it is impossible to prove that there was no effect.

PROFESSOR MICHAEL WOLF: Right, that cannot be done.

CHRISTOPHER COX: But this appears to show that -- rather it doesn't show that there was an effect. Is that the correct formulation?

PROFESSOR MICHAEL WOLF: In a court case, if there is no evidence for guilt, that doesn't mean we have proven innocence, but just there's no indication for guilt and then, you know, the verdict -- it is a stronger thing to prove innocence which statistically is not possible, we can say: do we find evidence for an effect, for guilt?

PROFESSOR DR ASHOK KAUL: We did the best to find an effect but we couldn't.

CHRISTOPHER COX: Other than on what you have described as a very naive view that would be --

PROFESSOR DR ASHOK KAUL: Of course.

CHRISTOPHER COX: So I am just thinking of the first chart you put up which shows the downward trend in prevalence and do you think that that is a true reflection of a continuing trend?

PROFESSOR MICHAEL WOLF: Obviously it can not go on for ever because at some point if you extrapolate that line, at some point that line will turn into a negative. At some point the downward trend has to slow down as it reaches zero.

CHRISTOPHER COX: No, sorry, that wasn't the point I was making, it was that over the period you

have looked, you think there is a continuous or fairly continuous -- the line of best fit shows a fairly continuous downward trend but what it doesn't show is any particular acceleration or deceleration that would be contemporary with --

PROFESSOR DR ASHOK KAUL: To put it differently, we see this line in all the OECD countries, it looks very similar across all countries and in some countries we had heavy anti-smoking measures, in other countries we didn't, but we see essentially the same line in all countries. Even if you look at the quantitative data, you find a minus 0.4 percentage point effect per year. That is also very similar, had the same effect in Germany, I think in the UK must be very similar as well. I don't remember the exact numbers but I had a look at all the OECD countries and that looks very similar across the countries over the last, let's say, one and a half decades.

CHRISTOPHER COX: Okay, that's great. Is there anything else you wanted to tell us about?

PROFESSOR DR ASHOK KAUL: No, we are through with the presentation and the research so far.

CHRISTOPHER COX: I just want to ask in terms of -- you have given us this paper which hopefully we can keep.

PROFESSOR MICHAEL WOLF: Of course.

PROFESSOR DR ASHOK KAUL: And you can also download it from next week on.

CHRISTOPHER COX: That was my question, just to clarify that. This is going to appear --

PROFESSOR MICHAEL WOLF: The working paper series of the University of Zurich and then automatically there are some global website like SSRN and RePEc so everything that goes into the working paper series will then automatically will be entered into those systems, so it would be out there in the open for everybody.

CHRISTOPHER COX: Is there a plan for it to be published in a journal?

PROFESSOR DR ASHOK KAUL: Yes, we are thinking about it, we finished that a couple of weeks ago so the usual process is extended and we wait for feedback. We have published all our papers in peer-reviewed journals so there's no reason to stop here.

CHRISTOPHER COX: So it is just a bit early as far as this work has got before you get to that point. Okay. Well, that is incredibly helpful and thoroughly interesting, particularly for those such as myself without a statistics background. So any other questions?

JAMES COLLIS: No, nothing from me.

CHRISTOPHER COX: Thank you again for making the trip over to see us.

PROFESSOR MICHAEL WOLF: It's always nice to be in London.

(12.02 pm)