Europe Economics

Analysis of the Chipty Report's conclusions regarding packaging changes and smoking prevalence in Australia

30 August 2016

Europe Economics
Chancery House
53-64 Chancery Lane
London WC2A 1QU

Tel: (+44) (0) 20 7831 4717
Fax: (+44) (0) 20 7831 4515

www.europe-economics.com

Contents

1 Introduction

1.1 My full name is Dr Andrew Lilico and I am a Principal and Chairman at Europe Economics, a private sector consultancy, based in London, which specialises in the application of economics and econometrics to problems arising predominantly in the fields of public policy, regulation and competition. Europe Economics' clients include government departments, regulators and competition authorities, companies large and small, professional and trade associations, charities, law firms and public affairs firms. In particular, I am expert in microeconomic analysis and regulatory impact assessment.

1.2 I am the author and project director of this report. However, in preparing this report, I have drawn upon support from the resources of Europe Economics.

1.3 Since December 2012, the sale of tobacco products in Australia has been subject to standardised (or "plain") packaging requirements as a result of the Tobacco Plain Packaging Act 2011 (hereafter referred to as the 2011 Act). Like other Australian legislation, the 2011 Act has been subject to a post-implementation review process. This has involved the two steps I describe below.

1.4 First, the Australian Department of Health engaged an external consultant to conduct a consultation on the impact of the plain packaging measure. The consultation took place from 16 February to 27 March 2015. At the request of JTI (part of the Japan Tobacco Group), I prepared a report in March 2015 entitled "*Review of Current Evidence Regarding the Impacts of Plain Packaging in Australia upon Consumption, Prevalence and Competition/Market Dynamics*", which was submitted to the consultation. In that report, I stated that: "*The reports I have reviewed are consistent in finding that plain packaging has had no statistically significant impact upon Australia's pre-existing decline in tobacco consumption and prevalence. In fact, some of the findings in these reports suggest that consumption may have risen, relative to previous trends (where controlled for), following the introduction of plain packaging*".

1.5 Second, on 26 February 2016, the Australian Department of Health published its post-implementation review report (hereafter referred to as the PIR). This PIR's stated aim is to "*assess the effectiveness and efficiency of the tobacco plain packaging measure to meet its objective in order to determine if it is an appropriate regulatory intervention*". The PIR made no reference to my March 2015 report submitted as part of the consultation which led to the PIR.

1.6 The PIR was accompanied by a report by Dr Tasneem Chipty entitled "*Study of the Impact of the Tobacco Plain Packaging Measure on Smoking Prevalence in Australia*" (hereafter referred to as the Chipty Report)[1] that seeks to model the impact of the introduction of packaging changes on smoking prevalence in Australia, using data from the Roy Morgan Single Source (RMSS) smoking overview data series.

1.7 As I discuss in more detail below, the Chipty Report concludes that there has been "*a statistically significant decline in smoking prevalence of 0.55 percentage points over the post-implementation period, relative to what the prevalence would have been without the packaging changes*". Although others have sought to reference this finding in a far less nuanced manner,[2] Dr Chipty correctly notes that this decline she has identified cannot be attributed to plain packaging. Given that new health warnings

[1] Available at: http://ris.dpmc.gov.au/2016/02/26/tobacco-plain-packaging/. An addendum to the Chipty Report was made available on 19 May 2016, as a result of Dr Chipty having been asked by the Australian Department of Health to convert the estimated reduction in smoking prevalence into an estimated reduction in the number of smokers attributable to the packaging changes.

[2] See, for example: http://www.themandarin.com.au/61097-plain-packaging-cutting-smoking-rates/?pgnc=1.

were introduced in Australia at the same time as plain packaging, her model considers both packaging changes together.

1.8 Dr Chipty has not made publicly available all the RMSS data that she has used in preparing her report, but an aggregated data series has been released in May 2016 by the Australian Department of Health following a freedom of information request (hereafter referred to as "the Data").[3] These Data are the average monthly percentage level of smoking prevalence from January 2001 to September 2015, derived from the full RMSS dataset. They do not include any other indications, in particular demographic data on age, sex, education or various other characteristics of respondents.

1.9 I have considered, at the request of JTI, the Chipty report, and the Data. Specifically, I consider whether the Data suggest that there is, in fact, a break in the smoking prevalence series at or around the date Dr Chipty claims. I make clear at the outset that, as I do not have access to all of the data used by Dr Chipty in preparing her report (including many of the demographic and other factors that she used in her models), I have not fully replicated Dr Chipty's analysis.

[3] FOI 175-1516, Australian Department of Health. In response to a request to the Australian Department of Health on 17 August 2016, I received the Data on 19 August 2016.

2 Summary of my conclusions

2.1 In this report I explore, using the Data:

- what statistical techniques suggest are the most significant breaks in the smoking prevalence series, finding that at least two breaks rank ahead of the December 2012 break reported by Dr Chipty — namely breaks in December 2004 and May 2012; and
- whether the December 2012 break has any explanatory power once the other two statistically-preferred breaks are added to a time series model and to a time series model with economic factors as extra controls.

2.2 My use of orthodox econometric techniques calls into question the reliability of the conclusion reached in the Chipty Report that there is a statistically significant decline in smoking prevalence following the introduction of plain packaging in Australia in December 2012.

2.3 My analysis of the Data finds two earlier statistically significant breaks which, when introduced into models based on the Data, have the effect of removing the statistically significant impact from December 2012 reported by Dr Chipty. In my time series models, both with and without additional economic factors, there is no such impact.[4] In other words, there becomes no basis for concluding that there was a statistically significant decline in smoking prevalence of 0.55 percentage points in the period since the introduction of plain packaging in Australia.[5]

2.4 Dr Chipty does not report having conducted any analogous tests to my analysis or offer any argument or evidence as to why such tests would be unlikely to produce similar results (i.e. the elimination of the statistical significance of her December 2012 break) in models based on her broader data.

2.5 As I have not replicated Dr Chipty's model in its entirety, since I do not have access to her underlying data, I cannot say definitively what the impact would be in her own models of introducing additional breaks at December 2004 and May 2012.

2.6 I explain in the next section my analysis of the Data, and the modelling I have undertaken in considering these Data. I have also set out in detail in Sections 4 and 5 of this report the results of the modelling I have undertaken in preparing this report.

[4] I observe that nothing in this report should be regarded as endorsing Dr Chipty's econometric methodology or choice of data more broadly or as being in tension with (or support of) the critiques of her models made by other authors. Our purpose has been narrowly focused upon whether there is evidence of a statistically significant break in the Data at or around December 2012 and whether such a break, if it exists, is correlated with any statistically significant change in smoking prevalence.

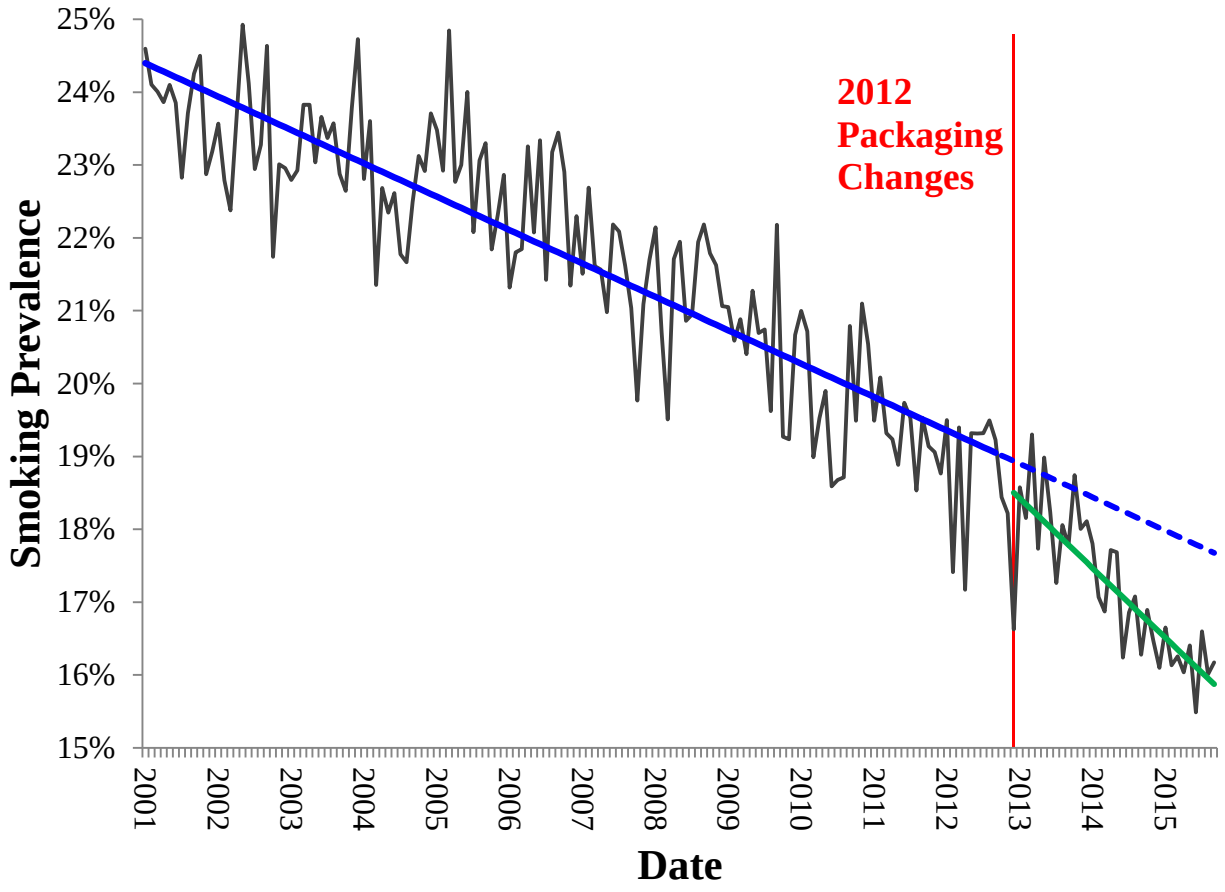
[5] As noted in paragraph 1.5, the Chipty Report does not go as far as saying, as Dr Chipty cannot even on her own analysis, that plain packaging has caused such a reduction.

3 Modelling of the Data

3.1 As noted above, the Data underlying one of Dr Chipty's charts from her report, namely Figure 1: Overall Smoking Prevalence, was released by the Australian Department of Health following a freedom of information in May 2016. These Data were the monthly percentage level of smoking prevalence from January 2001 to September 2015, from the RMSS. They do not include many other RMSS data, in particular demographic data on age, sex, education or various other characteristics of respondents.

3.2 Dr Chipty presents the first figure from her report thus.

Figure 3.1: Overall Smoking Prevalence



Source: Figure tab of Document 13 from FOI 175-1516, reproducing Chipty Report Figure 1.

3.3 We can see in this Chart that Dr Chipty invites the reader to accept that the data after the 2012 packaging changes, with a best-fit line in green, belong to a different series from the data before those changes, with a best-fit line in solid blue. The checked blue line shows how, according to this thought, one might have expected smoking prevalence to evolve absent the packaging changes, the difference between them being a measure of the impact of the packaging changes upon prevalence.

3.4 Much has been written about the Chipty Report.[6] As noted above, I have focused upon whether the Data suggest that there is, in fact, a break in the smoking prevalence series at or around the date Dr Chipty claims.

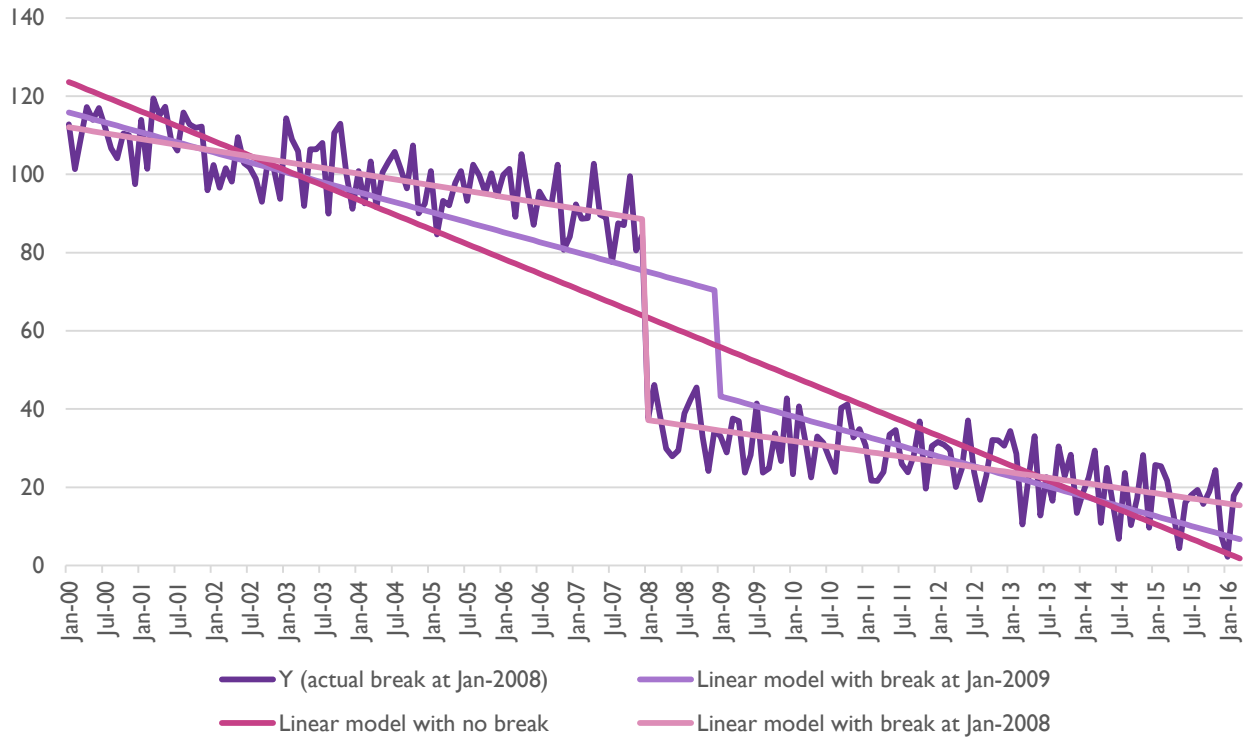
3.5 This is an important question for Dr Chipty's broader analysis, because although she does consider the possibility that there might be statistically preferred breaks in the months close to the break she claims — specifically in October or November 2012 — she does not report having investigated whether there might be any other breaks in the smoking prevalence series.

3.6 There are two ways other breaks in the series could potentially be relevant to Dr Chipty's findings.

- They could mean that there is actually no break in the series where Dr Chipty contends, but instead that the break she identifies is proxying for the true break or breaks.
- They could mean that even if there is a break at or close to the month , the break she identifies either has no statistically significant impact on smoking prevalence or has a lower or higher impact than her models suggest.

3.7 Consider Figure 3.1.

Figure 3.2: A hypothetical series with an actual break at January 2008 and a model with a break imposed at January 2009



3.8 Here we imagine a series[7] in which there is a break at January 2008. A good model for that series would introduce a break in January 2008, producing something along the lines of the light pink line

[6] For example, I am aware that Dr Chipty was instructed by the Australian Government in proceedings brought against its legislation introducing plain packaging at the WTO, and that others involved in those proceedings have commented on her analysis in that context, see: http://mic.gob.do/media/22058/20160323%20-%20DOM%20Integrated%20Executive%20Summary%20(EN).pdf (paragraphs 75 – 82). In addition, Professor Sinclair Davidson has also commented on the Chipty Report: http://catallaxyfiles.com/2016/03/17/australias-plain-packaging-experiment/.

[7] The numbers are generated randomly, with a constructed trend and break. The values themselves are purely illustrative.

in Figure 2. But suppose that, instead, a modeller did not inquire of the series what the best statistical break was, but, instead, asked whether introducing a break in January 2009 (as per the purple line) would produce a statistically significant improvement to a model that simply had a single best-fit line for the whole series (as per the dark pink line). A break imposed at January 2009 will, in this case, be found to be statistically significant, even though there is no such break in the underlying series, because it is proxying for the actual break in January 2008, despite being some time later.

3.9 To avoid this problem it is thus not sufficient to test whether any imposed break is statistically significant. Rather, we need to investigate the data points in the time series to find what the most statistically significant breaks are.

3.10 Specifically, we can use the Quandt likelihood ratio (QLR) test, which tests for a structural break with an unknown date.

Intuition of the QLR test for the best structural break in a series

3.11 Section 4 below explains the mathematics of the QLR test in detail. Here it suffices to set out the key intuition behind the test.

3.12 In essence, a QLR works as follows. In order to identify a statistical break (if any such break exists), we need sufficient data and a sufficient proportion of the total data to be before and after the potential break point. Therefore, the first step is to trim the series so that we look only at the internal portion where a break could be found.

3.13 Next, every point (every date, here) in that internal portion of the series is tested to see whether the best statistical model for the series as a whole is different from the best statistical model for the series after that date. Whenever there is such a difference, that date becomes a candidate break point and there is a statistical measure taken of how significant that potential break is.

3.14 Once all the candidate breaks have been found (if indeed any are found — one option is that the test rejects the hypothesis that there is a break at all), the QLR test then considers which of those breaks is most statistically significant — i.e. which is most likely be an actual break in the underlying data-generating process in the world. That is then identified as the "best" break.

Iterative implementation of QLR test for multiple breaks

3.15 The QLR test identifies the best break in a given series. But suppose there is more than one break in a series. What then?

3.16 The QLR test tells us that the best model of the data after the break point is different from the best model for the series as a whole. That leaves open two further questions.

- Are there any breaks in the series before the QLR break result?
- Are there any breaks in the series after the QLR break?

3.17 To investigate these questions, we can perform a second (or third, or subsequent) QLR test for the set of data before or after the overall-preferred QLR break. Let us refer to the first QLR break (the break identified from the whole series) as QLR Break 1. Then, let us suppose that we are not interested in breaks before QLR Break 1 but are interested in whether there might be additional breaks later. We can perform a second QLR test on the restricted series of data after QLR Break 1. If a statistically significant break is found for that subset we can refer to that as QLR Break 2. If we are interested in whether there might be further breaks after QLR Break 2, we can perform a third QLR test on the series after QLR Break 2, to identify a QLR Break 3 (if such a break is

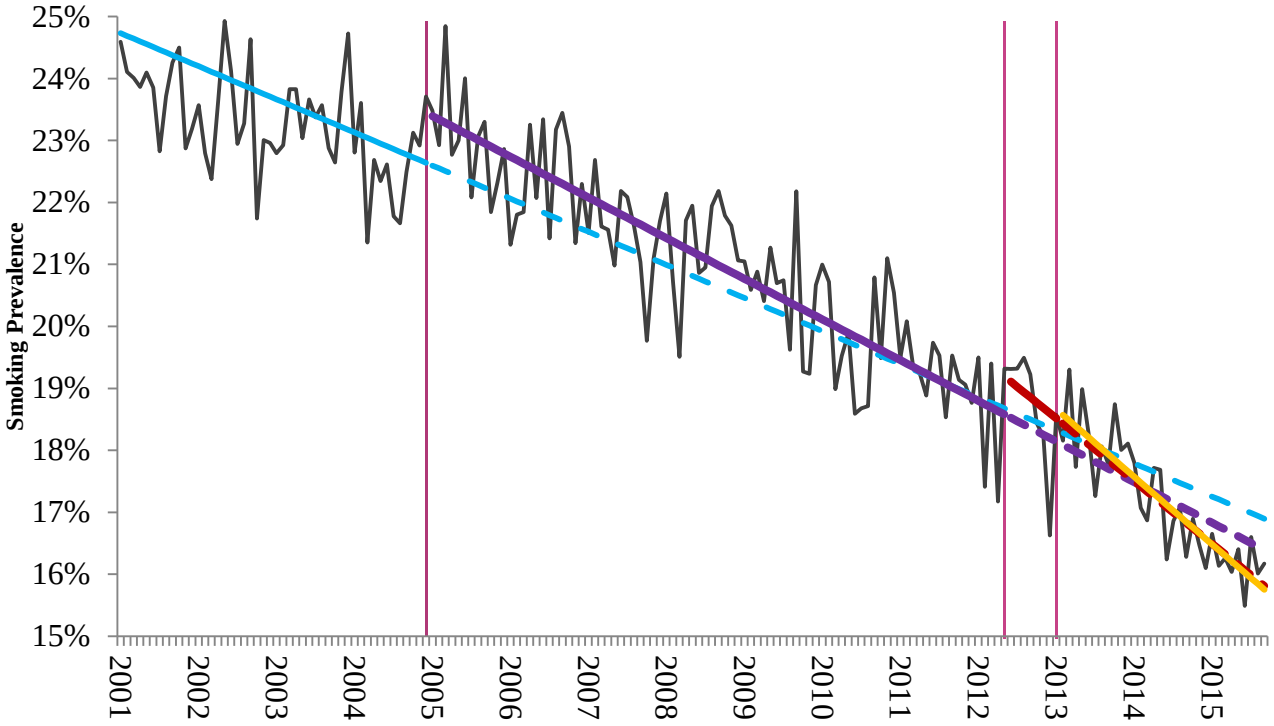
statistically significant). And so on until we either produce no more breaks or reach a date we are interested in testing. For example, if we are exploring whether a pre-imposed break is statistically significant, we might iterate through other breaks until we either come to a break at or around the point of the pre-imposed break or demonstrate that there is no statistically significant break in the series at or around that point.

Results of the QLR test for the Data

3.18 We applied the QLR test iteratively to the Data. The first QLR break we identified was in December 2004. We performed a second QLR test on the RMSS data for January 2005 to September 2015. We found a second break at May 2012. We performed a third QLR test on the RMSS data for June 2012 to September 2015. We found a third break at January 2013, close to the Chipty assumed break date of December 2012.

3.19 Thus, the Chipty break is only the (at least) third most significant break in the series. In Figure 3.3 we illustrate how the breaks segment the RMSS series.

Figure 3.3: The breaks in the RMSS smoking prevalence series



Implications of the simple iterative QLR test

3.20 We have so far seen that the Data have (at least[8]) two series breaks that, in a statistical sense, rank ahead of the Chipty break of December 2012 or January 2013. That suggests that, in Dr Chipty's models, she should have included dummy variables or other controls for December 2004 and May 2012 instead of or in addition to her December 2012 control.

[8] We emphasise that although we describe these other breaks using terms such as "Break 2" and "Break 3", this should not be interpreted as these being necessarily the second and third strongest breaks — although they can rank no higher, they could rank lower than that. Hence we qualify with "at least".

3.21 As noted above, we do not have access to all of the RMSS data to explore whether the breaks we have identified would persist once controls for age, sex, education and the other rich set of variables available in that dataset were added as controls. It is possible, in principle, that if there were material discontinuities in these other series, the results above could change (there might be no statistically significant break in December 2004, for example), though Dr Chipty has not offered any evidence of considering these other breaks or of providing any statistical reason to reject them in favour of her preferred break.

3.22 Although we cannot control directly for the impact on the results above if variables for these various demographic and economic factors were introduced, we can explore how likely it is that if such controls were introduced that would change our results. We do that next, by using different types of time series analyses.

Time series models

3.23 When data evolve through time, it is common to model and forecast them using a class of what are referred to as "autoregressive–integrated–moving–average" (ARIMA) models. Such time series models attempt to describe the behaviour of variables by exploiting any systematic relationship between a variable's current value and its past values.

3.24 Two key components in ARIMA models are an "autoregressive" (AR) term and a moving average (MA) term.[9] The "autoregressive" term describes how the present value of the variable depends on its previous value at some point in the past (say the previous month, or three months ago, or twelve months ago). The moving average term describes how the noisy fluctuations around the current values of the variable depend on the noisy fluctuations observed in the past.[10]

3.25 ARIMA models are particularly useful because they can provide an accurate description of a time-series variable by using only the information contained in the variables itself, i.e. without the need of additional control variables. For example, it is plausible to think that the evolution of smoking prevalence depends also on the evolution of a number of additional factors such as, the price of cigarettes or the proportion of the population that is male as opposed to female. The use of ARIMA models allows describing how smoking prevalence has evolved even if data for these additional factors is lacking — as, for example, we lack certain RMSS demographic data here.

3.26 Among the available range of ARIMA models, the "best" ARIMA model (i.e. the one with the correct *orders* for the AR and MA terms) can be selected based on the outcome of some standard statistical tests.[11]

3.27 In considering the Data, we found that monthly changes in smoking prevalence are best described by an ARIMA model with no autoregressive component and a moving average term of order 1 (i.e. a simple MA(1) process).[12] We can see how this MA(1) model compares to the smoking prevalence data in the figure below.

---

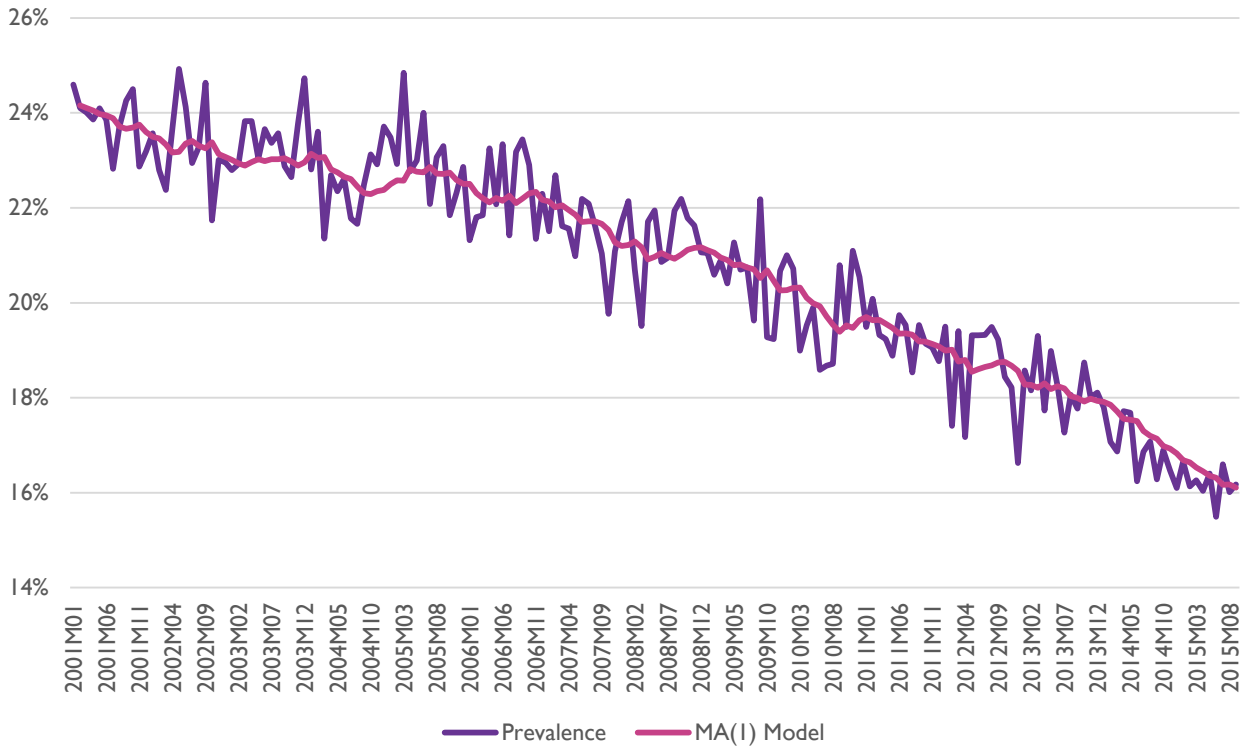
[9] The other element, the "I" in ARIMA, which stands for "Integrated", in this context means the model is calculated in first differences (i.e. in changes in values, rather than in levels).

[10] Within the broad class of ARIMA models, a specific model is characterised by an order for the autoregressive components (p), and an order for the moving average component (q). The order simply indicates the lag of the relationship linking current values to past values, so, for example, an autoregressive term of order two AR(2) indicates that the current value of a variable depends on the variables' value observed two periods earlier.

[11] For example, the Akaike Information Criterion (AIC) or the Schwarz Bayesian Information Criterion (BIC). For the purpose of selecting the best ARIMA model to describe the behaviour of monthly smoking prevalence we used the BIC statistic, applied iteratively across possible ARIMA models so as to identify those that perform best.

[12] The ARIMA models have been applied to the first-difference of smoking prevalence in order to avoid non-stationarity problems.

Figure 3.4: Chart of MA(1) model



3.28 As set out in Section 5, there are structural breaks in this MA(1) model at December 2004, May 2012, and January 2013. We introduced dummy variables for each of these breaks.[13]

3.29 Next we tested for the significance of the three breaks sequentially, through the use of dummy variables. More specifically, we added sequentially dummy variables representing break B1, B2 and B3 to the MA(1) model of Table 5.1 to test whether these are statistically significant. The results of these models are reported in Section 5 and indicate that:

- B1 is statistically significant when alone;
- B1 and B2 are simultaneously significant when both are present;
- when break B3 is also added to the model, the only break to remain statistically significant is B1.

3.30 Thus, in a pure time-series model smoking prevalence has breaks before the introduction of plain packaging in Australia in December 2012 (one in December 2004, and one in May 2012) and these earlier breaks eliminate the significance of the later break.

Implications of the pure time series tests

3.31 The time series data suggest that, not only do two other breaks rank statistically ahead of Dr Chipty's imposed break of December 2012 or January 2013, but once they are introduced into a model, it is likely that Dr Chipty's break would cease to have a statistically significant impact (as indeed it ceases to have a statistically significant impact in the best time series models).

3.32 Thus far we have used only pure time series models, to which breaks have been added. The relevance of such models in this case is based on the presumption that the underlying drivers of

[13] A "dummy variable" takes a value of 0 before the break and a value of 1 thereafter.

smoking prevalence in the RMSS series would be able to be characterised well by a pure ARIMA time series. One potential challenge to this might be that relevant economic determinants of smoking prevalence might not only be demographic characteristics such as age or sex and education, but also economic variables such as price or income. These latter variables might be subject to their own significant discontinuities, which might undermine the susceptibility of the RMSS data to pure time series analysis.

3.33 To explore this point further, we next introduce measures of price and income as controls.

Testing for multiple breaks in a model with additional economic explanatory variables

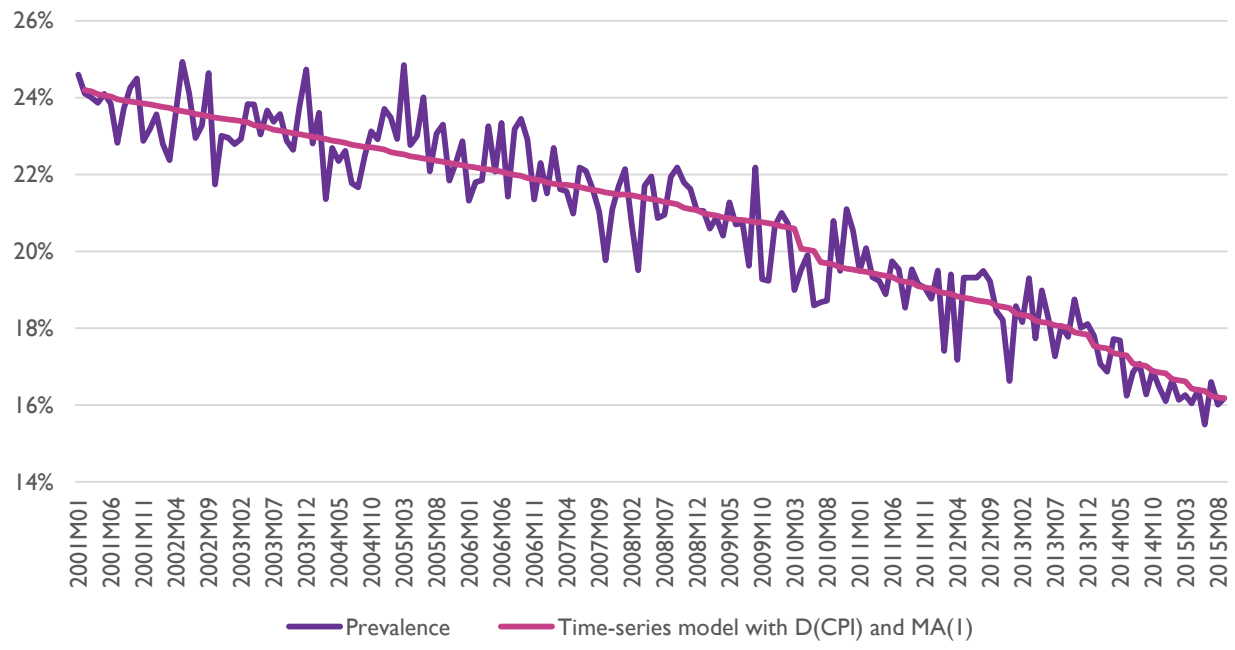
3.34 To cross-check whether controls for economic factors might change our results, we replicated the multiple breaks analysis described above in a model which includes additional control variables that are potentially relevant to smoking prevalence, namely:

- The relative price of tobacco products — this is calculated as the ratio between the seasonally adjusted consumer price index (CPI) for tobacco products in Australia and the seasonally adjusted consumer price index (CPI) for all product groups in Australia, as available from the Australian Bureau of Statistics.[14]

- GDP per capita — this is the seasonally adjusted GDP per capita (chain volume measure), as available from the Australian Bureau of Statistics.[15]

3.35 I explain in Section 5 that when we introduce additional controls for GDP per capita and relative tobacco prices, only the prices prove to be statistically significant. The best time series model with ARIMA components and relative tobacco prices is, again, an MA(1), which we illustrate in the following figure.

Figure 3.5: Time-series model with changes in the relative price of tobacco



[14] These variables are available from ABS Cat. 6401.0, series IDs A3604731A, and A3604506F.
[15] This variable is available from ABS Cat. 5206.0, series ID A2304404C.

3.36 We then tested whether the three dummy variables for our three breaks (December 2004, May 2012 and January 2013), when added sequentially to the model above, are statistically significant. The results are reported in Sections 4 and 5. We find that, in all three models, the only statistically significant break is the break in December 2004. Neither the May 2012 break nor the January 2013 break are significant.

Implications of the time series tests with additional economic explanatory variables

3.37 The key difference between the models with additional economic explanatory variables and the pure time series models is that, with economic variables included, only the December 2004 break remains statistically relevant whereas in the pure time series models both the December 2004 and May 2012 breaks were statistically significant. In neither case was Dr Chipty's break statistically significant once the other statistically-preferred breaks were included.

4 The QLR Test

4.1 Testing for a single structural break at an unknown date, also known as the Quandt likelihood ratio (QLR), is based on the standard Chow test for a single structural break at a given (known) date. The concept of QLR test was proposed by Quandt in 1960 but, due to the lack of theories at the time the formal test procedure was developed by Andrews only in 1993. Below we provide a theoretical rational and the implementation guide for the QLR test.

4.2 Since the QLR test relies on the Chow test for a known break date, we first describe the regression and test set-up and the Chow test. We then present the QLR test and its practical implementation.

Regression and test set-up

4.3 Suppose we have a linear regression with a single break at time point τ:

$$y_t = x'_t\beta_1 + \varepsilon_t, \qquad t = 1, \dots, \tau$$

$$y_t = x'_t\beta_2 + \varepsilon_t, \qquad t = \tau + 1, \dots, T$$

where

$t = 1, \dots, \tau, \tau + 1, \dots, T$ – time index running from 1 to T,

τ – a single break point,

y_t – the dependent variable,

x_t – a vector of k explanatory variables, $k = 1, \dots K$.

β_1 and β_2 – coefficients for explanatory variables, before and after the time break, respectively,

ε_t – i.i.d. errors under normal distribution with zero mean and variance σ^2, $\varepsilon_t \sim N(o, \sigma^2)$.

4.4 In the absence of structural break, we would have the same coefficient vector β before and after the break. This assumption is known in the test procedure as the null hypothesis. $H_0: \beta_1 = \beta_2$. The alternative hypothesis is the existence of break at time point τ, $H_{alt}: \beta_1 \neq \beta_2$.

Single break at a known date

4.5 If we have some preliminary information about the break date, we can use the standard Chow test (1960) for verifying whether a break occurred at a given point in time. The Chow test calculates the following F-statistics:

$$F_T(\tau) = \frac{(SSR_{1:T} - (SSR_{1:\tau} + SSR_{\tau+1:T}))/k}{(SSR_{1:\tau} + SSR_{\tau+1:T})/(T - 2 * k)}$$

where

k – number of explanatory variables (as above).

SSR – sum of squared residuals in a given sample: $1{:}\,T$ – full time span, $1{:}\,\tau$ – before the break, and $\tau + 1{:}\,T$ – after the break.

4.6 When the date break is known, the F-statistics is distributed under the null hypothesis as chi-squared distribution with k degrees of freedom: $F_T(\tau) \sim \chi^2(k)$. The calculated F-statistics is compared to a tabulated value and a decision is made whether to accept or reject the null hypothesis of no break at a certain level of statistical significance.

Single break at unknown date

4.7 If we do not have information about the break date, we have to use the QLR statistics developed by Quandt (1960) and Andrews (1993) based on the Chow F-statistics:

$$QLR_T = \max_{\tau \in \{\tau_{min}, \ldots, \tau_{max}\}} F_T(\tau)$$

where

τ – an unknown break date,

τ_{min} and τ_{max} – lower and upper bounds, respectively, of a supposed time break.

4.8 The QLR statistics takes max over the F-statistics calculated at all possible time points where a structural break might occur. For test stability, the original time interval is trimmed at both sides, typically at 15 per cent, i.e. the first 15 per cent and the last 15 per cent of observations are excluded from testing. The QLR statistics has a non-standard distribution and the statistics value has to be compared against simulated critical values to make the decision when the break could most likely occur.[16]

Implementation of QLR test

4.9 The QLR theory provides a straightforward route for practical implementation. The test for a structural break at an unknown date proceeds as follows:

1. Calculate the Chow F-statistics for all time points in the given sample, excluding the first and the last 15 per cent of observations.
2. Calculate the QLR statistics as the max of all Chow F-statistics, to choose the likeliest break date.
3. Compare the QLR statistics value against the table of non-standard critical values and decide whether the break indeed occurred at the likeliest date, or reject the existence of break altogether.

[16] A table of critical values is available, for example, in Chapter 14.6 of Stock, James H. and Watson, Mark W. (2011) *Introduction to Econometrics, 3rd Edition*. Pearson Education.

5 Time Series Tests

Pure time series tests

5.1 As noted in Section 3, in a pure time series model we found that monthly changes in smoking prevalence are best described by an ARIMA model with no autoregressive component and a moving average term of order 1 (i.e. a simple MA(1) process).[17] We provide the estimation output and the chart of this model below.

Table 5.1: Estimation output of MA(1) model

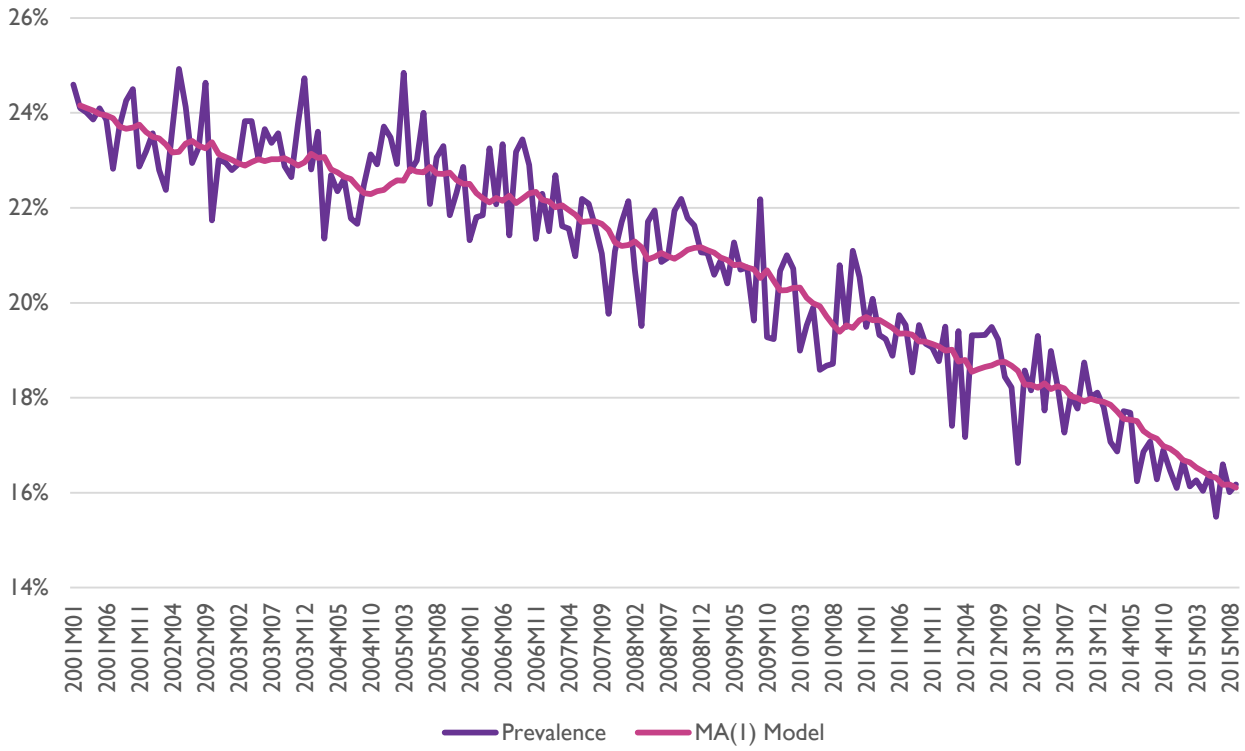
Dependent Variable: **Change in Smoking Prevalence**
Method: Least Squares
Sample (adjusted): 2001M02 2015M09
Included observations: 176 after adjustments
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 5.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	-0.000453	8.40E-05	-5.393837	0.0000
MA(1)	-0.873882	0.038344	-22.79056	0.0000

R-squared	0.410397	Mean dependent var	-0.000478
Adjusted R-squared	0.407009	S.D. dependent var	0.010133
S.E. of regression	0.007803	Akaike info criterion	-6.857316
Sum squared resid	0.010594	Schwarz criterion	-6.821288
Log likelihood	605.4438	Hannan-Quinn criter.	-6.842703
F-statistic	121.1139	Durbin-Watson stat	1.920499
Prob(F-statistic)	0.000000		

[17] The ARIMA models have been applied to the first-difference of smoking prevalence in order to avoid non-stationarity problems. We have confirmed that the series is stationary in first differences.

Figure 5.1: Chart of MA(1) model



5.2 We then tested whether the MA(1) model described above has a structural break at December 2004, May 2012, and January 2013. We first tested for the significance of such breaks in isolation. The results are reported below and indicate that the most significant break is at December 2004, the second most significant break is at May 2012, whilst the least significant break is at January 2013.

Table 5.2: Testing for breaks in isolation in a pure time-series model

Break name	Break date	F-Statistics output of Chow Test	Interpretation
B1	Dec-2004	0.0003	The break is statistically significant at the 99 per cent confidence level
B2	May-2012	0.0222	The break is statistically significant at the 95 per cent confidence level
B3	Jan-2013	0.0584	The break is statistically significant at the 90 per cent confidence level

5.3 Next we tested for the significance of the three breaks sequentially, through the use of dummy variables. More specifically, we added sequentially dummy variables representing break B1, B2 and B3 to the MA(1) model of Table 5.1 to test whether these are statistically significant. The results of these models are reported in the following tables and indicate that:

- B1 is statistically significant when alone (Table 5.3);
- B1 and B2 are simultaneously significant when both are present (see Table 5.4);
- when break B3 is also added to the model, the only break to remain statistically significant is B1 (Table 5.5).

Table 5.3: Estimation output of MA(1) model with break B1

Dependent Variable: **Change in Smoking Prevalence**
Method: Least Squares
Sample (adjusted): 2001M02 2015M09
Included observations: 176 after adjustments
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 5.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	-0.000142	0.000128	-1.105804	0.2703
B1	-0.000386	0.000149	-2.592692	0.0103
MA(1)	-0.943725	0.029653	-31.82597	0.0000

R-squared	0.428392	Mean dependent var	-0.000478
Adjusted R-squared	0.421784	S.D. dependent var	0.010133
S.E. of regression	0.007705	Akaike info criterion	-6.876949
Sum squared resid	0.010271	Schwarz criterion	-6.822907
Log likelihood	608.1715	Hannan-Quinn criter.	-6.855030
F-statistic	64.82759	Durbin-Watson stat	1.848271
Prob(F-statistic)	0.000000	Wald F-statistic	6.722054
Prob(Wald F-statistic)	0.010338		

Table 5.4: Estimation output of MA(1) model with breaks B1 and B2

Dependent Variable: **Change in Smoking Prevalence**
Method: Least Squares
Sample (adjusted): 2001M02 2015M09
Included observations: 176 after adjustments
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 5.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	-0.000133	5.95E-05	-2.235912	0.0266
B1	-0.000331	8.14E-05	-4.061893	0.0001
B2	-0.000218	6.43E-05	-3.386948	0.0009
MA(1)	-0.999953	0.021177	-47.21923	0.0000

R-squared	0.462570	Mean dependent var	-0.000478
Adjusted R-squared	0.453196	S.D. dependent var	0.010133
S.E. of regression	0.007493	Akaike info criterion	-6.927240
Sum squared resid	0.009657	Schwarz criterion	-6.855183
Log likelihood	613.5971	Hannan-Quinn criter.	-6.898014
F-statistic	49.34728	Durbin-Watson stat	1.859259
Prob(F-statistic)	0.000000	Wald F-statistic	35.01894
Prob(Wald F-statistic)	0.000000		

Table 5.5: Estimation output of MA(1) model with breaks B1, B2, and B3

Dependent Variable: **Change in Smoking Prevalence**
Method: Least Squares
Sample (adjusted): 2001M02 2015M09
Included observations: 176 after adjustments

HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 5.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	-0.000131	6.41E-05	-2.040871	0.0428
B1	-0.000339	9.06E-05	-3.741429	0.0002
B2	-3.67E-06	0.000387	-0.009487	0.9924
B3	-0.000318	0.000450	-0.706599	0.4808
MA(1)	-0.991212	0.009763	-101.5269	0.0000

R-squared	0.462476	Mean dependent var	-0.000478
Adjusted R-squared	0.449902	S.D. dependent var	0.010133
S.E. of regression	0.007516	Akaike info criterion	-6.915700
Sum squared resid	0.009659	Schwarz criterion	-6.825630
Log likelihood	613.5816	Hannan-Quinn criter.	-6.879168
F-statistic	36.78128	Durbin-Watson stat	1.875727
Prob(F-statistic)	0.000000	Wald F-statistic	25.14108
Prob(Wald F-statistic)	0.000000		

Testing for multiple breaks in a time series model with additional economic explanatory variables

5.4 Since the relative price of tobacco products and GDP per capita are available on a quarterly basis (i.e. at a lower level of frequency than the monthly smoking prevalence data), there are two main approaches that can be used in order to include these variables in a model that aims to explain smoking prevalence:[18]

- The first approach is to run the model with monthly data and to assume that the relative price index for tobacco products and the GDP per capita have the same value across the three months of each quarter.
- The second approach is to run the model with quarterly data by defining quarterly smoking prevalence as the average smoking prevalence across the three months of each quarter.

5.5 We present below the results obtained following the first approach but, we note that with the second approach, we have obtained results that are the same in all key substantive ways.

5.6 We first estimate a model which includes both changes in the relative price of tobacco product and changes in GDP per capita as explanatory variables. We enrich such model by adding also AR and MA components, the order of which is determined by a model selection process based on the outcome of BIC statistics. The estimation output of the selected model (which includes an AR(1) and a MA(1) component) is provided below.

[18] A third approach would be to interpolate the quarterly data by assuming a linear growth or contraction within the quarter. Such an approach would have the considerable drawback that the data for a number of months would be constructed by the modeller.

Table 5.6: Time-series model with changes in the relative price of tobacco and GDP per capita

Dependent Variable: **Change in Smoking Prevalence**
Method: Least Squares
Sample (adjusted): 2001M03 2015M09
Included observations: 175 after adjustments
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 5.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	-0.000393	0.000112	-3.511535	0.0006
Change in relative CPI	-0.037761	0.012656	-2.983656	0.0033
Change in GDP per capita	3.36E-06	4.09E-06	0.820793	0.4129
AR(1)	0.082509	0.072682	1.135209	0.2579
MA(1)	-1.048871	0.038079	-27.54486	0.0000

R-squared	0.487332	Mean dependent var	-0.000453
Adjusted R-squared	0.475269	S.D. dependent var	0.010156
S.E. of regression	0.007357	Akaike info criterion	-6.958121
Sum squared resid	0.009202	Schwarz criterion	-6.867698
Log likelihood	613.8356	Hannan-Quinn criter.	-6.921443
F-statistic	40.39968	Durbin-Watson stat	2.047172
Prob(F-statistic)	0.000000	Wald F-statistic	7.816179
Prob(Wald F-statistic)	0.000566		

5.7 From Table 5.6 we notice that, whilst a change in the relative price of tobacco is statistically significant and, as we would expect, has a negative coefficient (this implies that an increase in the price of tobacco relative to other goods is associated with a reduction in smoking prevalence), a change in GDP per capita is not statistically significant.

5.8 We have, therefore, estimated another model which includes only the relative price of tobacco as an explanatory variable. Again, the model includes AR and MA components selected based upon the outcome of BIC statistics (thus, the model that follows uses AR and MA components selected independently, not developed from Table 5.6). When only the change in the price of cigarettes is included as an explanatory variable, the selected model is (as in the previous section) one with only a MA(1) component. The estimation results and the chart for such model are provided below.

Table 5.7: Time-series model with changes in the relative price of tobacco

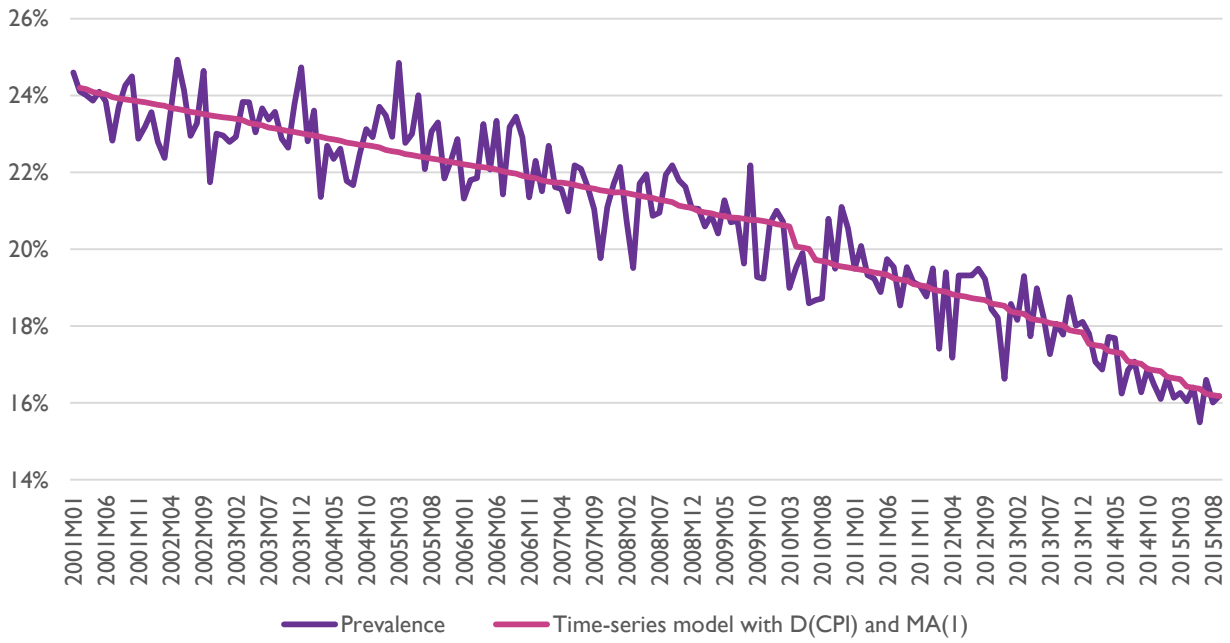
Dependent Variable: **Change in Smoking Prevalence**
Method: Least Squares
Sample (adjusted): 2001M02 2015M09
Included observations: 176 after adjustments
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 5.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000289	2.54E-05	-11.39119	0.0000
Change in relative CPI	-0.044237	0.005033	-8.789216	0.0000
MA(1)	-0.999957	0.014038	-71.23429	0.0000

R-squared	0.441553	Mean dependent var	-0.000478
Adjusted R-squared	0.435097	S.D. dependent var	0.010133

S.E. of regression	0.007616	Akaike info criterion	-6.900242
Sum squared resid	0.010035	Schwarz criterion	-6.846200
Log likelihood	610.2213	Hannan-Quinn criter.	-6.878323
F-statistic	68.39387	Durbin-Watson stat	1.804220
Prob(F-statistic)	0.000000	Wald F-statistic	77.25033
Prob(Wald F-statistic)	0.000000		

Figure 5.2: Time-series model with changes in the relative price of tobacco



5.9 We notice from Table 5.7 that changes in the relative price of tobacco remain statistically significant and negatively associated to changes in smoking prevalence.

5.10 We then tested whether the three dummy variables for breaks B1, B2, and B3, when added sequentially to the model of Table 5.7 are statistically significant. The results are reported in Tables 5.8-5.10. We find that, in all three models, the only statistically significant break is B1. Neither B2 nor B3 are significant.[19]

Table 5.8: Time-series model with changes in the relative price of tobacco and B1

Dependent Variable: **Change in Smoking Prevalence**
Method: Least Squares
Sample (adjusted): 2001M02 2015M09
Included observations: 176 after adjustments
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 5.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	-0.000174	6.19E-05	-2.805169	0.0056
B1	-0.000178	0.000108	-1.656937	0.0994

[19] I note that, under the QLR procedure I have used, breaks should be introduced in the sequence the QLR test identifies them. So, B1 should be introduced first, then B2 then B3. However, for interest, I note that if B1 is introduced then B3 is introduced next, without B2 (noting that B2 is not statistically significant in Table 5.9), B3 is not statistically significant.

Change in relative CPI	-0.034913	0.010555	-3.307750	0.0011
MA(1)	-0.999985	1.78E-06	-562530.5	0.0000

R-squared	0.465041	Mean dependent var	-0.000478
Adjusted R-squared	0.455711	S.D. dependent var	0.010133
S.E. of regression	0.007476	Akaike info criterion	-6.931848
Sum squared resid	0.009612	Schwarz criterion	-6.859792
Log likelihood	614.0027	Hannan-Quinn criter.	-6.902623
F-statistic	49.84004	Durbin-Watson stat	1.879037
Prob(F-statistic)	0.000000	Wald F-statistic	30.12105
Prob(Wald F-statistic)	0.000000		

Table 5.9: Time-series model with changes in the relative price of tobacco, B1, and B2

Dependent Variable: **Change in Smoking Prevalence**
Method: Least Squares
Sample (adjusted): 2001M02 2015M09
Included observations: 176 after adjustments
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 5.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	-0.000142	5.99E-05	-2.378345	0.0185
B1	-0.000262	0.000105	-2.507968	0.0131
B2	-8.06E-05	0.000130	-0.622003	0.5348
Change in relative CPI	-0.020236	0.015753	-1.284579	0.2007
MA(1)	-0.999758	0.024445	-40.89819	0.0000

R-squared	0.468832	Mean dependent var	-0.000478
Adjusted R-squared	0.456407	S.D. dependent var	0.010133
S.E. of regression	0.007471	Akaike info criterion	-6.927596
Sum squared resid	0.009544	Schwarz criterion	-6.837525
Log likelihood	614.6284	Hannan-Quinn criter.	-6.891064
F-statistic	37.73301	Durbin-Watson stat	1.887277
Prob(F-statistic)	0.000000	Wald F-statistic	26.94264
Prob(Wald F-statistic)	0.000000		

Table 5.10: Time-series model with changes in the relative price of tobacco, B1, B2, and B3

Dependent Variable: **Change in Smoking Prevalence**
Method: Least Squares
Sample (adjusted): 2001M02 2015M09
Included observations: 176 after adjustments
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 5.0000

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	-0.000138	6.53E-05	-2.108781	0.0364
B1	-0.000273	0.000111	-2.467030	0.0146
B2	0.000129	0.000379	0.340486	0.7339
B3	-0.000311	0.000430	-0.722576	0.4709
Change in relative CPI	-0.020038	0.015281	-1.311365	0.1915
MA(1)	-0.991759	0.009856	-100.6278	0.0000

R-squared	0.468494	Mean dependent var	-0.000478
Adjusted R-squared	0.452862	S.D. dependent var	0.010133
S.E. of regression	0.007495	Akaike info criterion	-6.915597
Sum squared resid	0.009550	Schwarz criterion	-6.807512
Log likelihood	614.5725	Hannan-Quinn criter.	-6.871758
F-statistic	29.96921	Durbin-Watson stat	1.901671
Prob(F-statistic)	0.000000	Wald F-statistic	21.26937
Prob(Wald F-statistic)	0.000000		