

FoolWatch: A Case study of econometric analysis and 'evidenced-based-policy making' in the Australian Government*

Don Harding

Department of Economics and Finance La Trobe University

July 1, 2008

*This paper is somewhat outside my usual fields of research (macroeconomics and econometrics). I have been influenced by Stilgoe, Irwin and Jones (2006) who provide a useful discussion of 'evidence-based-policy' making in the UK. They provided links to many sources of information quoted in this paper including the haunting and beautiful sonet by Edna St Vincent Millay quoted in full on the on the next page of this paper.

Sonnet X

*Upon this age, that never speaks its mind,
This furtive age, this age endowed with power
To wake the moon with footsteps, fit an oar
Into the rowlocks of the wind, and find
What swims before his prow, what swirls behind —
Upon this gifted age, in its dark hour,
Falls from the sky a meteoric shower
Of facts . . . they lie unquestioned, uncombined.
Wisdom enough to leech us of our ill
Is daily spun; but there exists no loom
To weave it into fabric; undefiled
Proceeds pure Science, and has her say; but still
Upon this world from the collective womb
Is spewed all day the red triumphant child.
Edna Saint Vincent Millay (1939)¹.*

¹This sonnet was written to question the isolationist approach of US foreign policy in light of the 'meteoric shower of facts' about Nazi Germany in the 1930s. It remains relevant today because of its beauty and because Millay perceives the timeless nature of the essential problems of science and evidence in shaping policy: there are facts enough to spin the threads of wisdom but often there exists no loom to weave those threads into a fabric that is sufficiently undefiled by self interest as to provide the insights that shape good policy. Ultimately, the United States abandoned isolationism under FDR's leadership. Perhaps, that experience and this beautiful sonnet can inspire modern leaders to develop the loom needed to weave today's meteoric shower of facts into solutions to today's problems.

Abstract

The decision to introduce a national FuelWatch scheme provides a timely case study of ‘evidence-based- policy’ making in Australia.

The government based its decision on econometric work by the ACCC who refuse to release the data. They claim that their analysis is robust because it has been subject to scrutiny within the ACCC and by Treasury.

Experience with ‘evidence based policy’ making in the United Kingdom (UK) raises doubts about such claims. The UK experience led to the term ‘policy-based-evidence’ to describe the end result where government agencies filtered out information that was inconsistent with government policy.

The data used by the ACCC can be digitized from a graph in their report. I then use the data so obtained to assess the robustness of the ACCC analysis and econometrics. Since the government and their advisors did not realize that the data could be obtained in this way it provides a ‘natural experiment’ in which their claims can be evaluated and tested.

I find that the ACCC apply the wrong tests to the wrong variable. Specifically, they study the nominal retail margin when economic theory suggests that analysis of anything but the real retail margin to producers creates a misspecified model inconsistent with the econometric assumptions used.

The econometrics in Appendix S of the ACCC report on petrol is substandard in application of techniques and in reporting of what was done. This should not be seen as an adverse reflection on the econometricians employed by the ACCC. Rather, it is a reflection on the process through which their work is incorporated into reports and used by government.

The ACCC findings are not robust. When I apply the correct version of the procedures used by the ACCC to the correct variable I find that the data does not support the original ACCC finding. Specifically, it is not possible to conclude as the ACCC did that FuelWatch did not raise petrol prices in Western Australia.

FuelWatch of itself is unimportant. The important issues here relate to the integrity of the government’s ‘evidence-based’ approach to policy. The UK experience clearly shows that relying on the untested opinion of ‘experts’ leads to ‘fudging’ of the evidence or overstatement of the conclusions that can be supported by the data. This ultimately corrupts the evidence-based-policy approach. The FuelWatch experience shows that these dangers are present for Australian policy makers. Greater transparency by government, publication of data and analysis underpinning government decisions and independent review of econometric work are the only protections against this danger. Failure to implement these protections will invalidate claims about the evidence base of future policy decisions.

1 Introduction

On 24 November 2007 the Australian people elected a new government that promised to increase the extent to which evidence is used to inform its policy decisions.² Six months into the new Government's term it's decision to introduce a national FuelWatch scheme provides a timely case study of econometric 'evidence-based policy' making in the Rudd Labor government. Background information on FuelWatch is provided in section 2

The Australian government and the ACCC claim that Appendix S of the ACCC (2007) report on petrol pricing together with the ACCC's (2008) supplementary material constitute evidence of the success of FuelWatch in lowering petrol prices in Western Australia. In April 2008 the Australian Government based its decision to introduce a national FuelWatch scheme on the ACCC econometric analysis.

Section 3 shows how the data used by the ACCC in its econometric study of FuelWatch can be extracted from a graph in the ACCC report something that is entirely legal. The fact that I can obtain the data in this way contributes to making this an ideal case study as one is not distracted by legal issues as is often the case when information is 'leaked'. Also the empirical questions are simple. Did FuelWatch have any impact on retail prices in Western Australia? And if so what was the direction and magnitude of that effect? The simplicity of these questions also makes the FuelWatch decision ideal for a case study as attention can be focused on the issues associated with evidence based policy making.

There are three main issues.

First, what is 'evidence'? Most scientists would not view information as being evidence until it was shown that the results could be replicated independently - what this means in practice varies across the sciences but a minimum requirement is that the data be public and the paper must have been subject to independent public scrutiny. Often this scrutiny takes the form of refereeing and publication in a journal. In courts of law information is not regarded as evidence unless it is tested in court.

The government has claimed that the ACCC are the experts. At Senate estimates the ACCC claimed that the econometric work was made suffi-

²A preliminary draft of this paper was sent to appropriate government officials so that they could comment and correct any errors. These included the Chief Executive Officer of the ACCC, Executive Director Markets Group Treasury and Deputy Secretary Domestic Policy Department of Prime Minister and Cabinet.

I am grateful to Dr Stephen King for a comment that helped me to clarify an important issue. I am also grateful to Richard Hayes of the ACCC for his comments.

All responsibility for errors and omissions is mine alone.

ciently robust to be considered evidence because it was reviewed independently by other ACCC staff and by the Treasury.

The UK's experience with 'evidence-based-policy' led to the term 'policy-based-evidence' to describe the case where information released by government was filtered to remove that which was inconsistent with government policy. The UK experience summarised in Stilgoe, Irwin and Jones (2006) suggests that it is unwise to characterise bureaucrats as 'experts' unless one is really sure that the information is being supplied by the person who actually did the econometrics rather than by their manager or their manager's manager. The UK experience also suggests that internal review procedures and even review by other government agencies have little value in protecting against 'policy-based-evidence'. The reason is that it is not in the career interests of the bureaucrats to 'lean too heavily against the wind'.

The UK experience provides a testable hypothesis. If the government and ACCC are correct I should find no significant flaws in the ACCC analysis or econometrics. This hypothesis is tested in section 4 where the ACCC's econometric analysis is evaluated against the standard of what is regarded as good practice by experienced econometricians supervising honours and Ph.D. students.

The second issue discussed briefly in this paper is who are the experts? Expert advice is essential in evidence-based-policy making. It follows that the capacity to recognize who are the experts is essential to true evidence-based-policy making. This issue is discussed in section 5.

The third issue is how can the experts be used to assist in implementing 'evidence-based-policy' rather than 'policy-based-evidence'? I discuss this issue in section 6 which discusses assembles information about evidence based policy making in Australia and the United Kingdom.

Conclusions are presented in section 7.

2 Background to FuelWatch and the 2007 ACCC report on petrol pricing

The national FuelWatch scheme which was announced on April 15 and is to start on 15 December 2008 is based on the Western Australian government's scheme of the same name.

2.1 Western Australian FuelWatch

On 2 January 2001 the Western Australian government introduced the Fuel-Watch scheme. The key feature of this is the '24 hour rule' which is explained

as follows on the FuelWatch website

The 24-Hour Rule

Throughout most of WA fuel retailers are required to notify their next day's retail price for each fuel they sell by 2pm.

Retailers must stay at the notified price from 6am for 24 hours this means WA motorists are free from the frustrating intra-day price fluctuations, which motorists in other Australian States still experience.

Source: http://www.fuelwatch.wa.gov.au/about/dsp_how_fuelWatch_works.cfm

The FuelWatch scheme provides a website, telephone number and email service that consumers can use to locate the cheapest price on a given day.

2.2 The 2007 ACCC report on petrol pricing

Petrol pricing has been a contentious issue in Australia and has been the subject of a large number of inquiries. In many of these inquiries the ACCC has voiced suspicion that the Western Australian FuelWatch scheme might lead to increases in the average price of petrol. However, the ACCC did not have access to sufficient data to test this hypothesis.

In late 2007 the ACCC delivered a report on Petrol pricing that was very cautious about FuelWatch stating that³

Assessing any system in the style of FuelWatch that incorporates increased price information and price commitment requires great care due to the potential for anti-competitive as well as pro-competitive benefits. Although the inquiry gained a preliminary assessment of the impacts in Perth from the scheme, it is clear that a case-by-case approach is required to assess the potential impacts on competition of any similar scheme. In particular the ACCC has not analysed the application of such a scheme to rural and regional areas. Apparent extra considerations here include the increased potential for anti-competitive effects due to the more concentrated nature of the market, the extra cost in initialisation, administration and compliance and how to decide which areas to cover. In summary, there are potential benefits and potential costs of adopting a national price commitment arrangement that need to be carefully considered.

³ACCC, 2007, *Petrol prices and Australian consumers—report of the ACCC inquiry into the price of unleaded petrol*, 18th December 2007

Appendix S of the ACCC report is an econometric analysis of the Western Australian FuelWatch scheme. Changes to the relevant laws in 2006 and at the end of 2007 provided the ACCC with new powers to subpoena documents, data and witnesses.⁴ The data used in this analysis was obtained from Informed Sources using the ACCC's new subpoena powers. Mr. Graeme Samuel, Chairman of the ACCC explained the propose of that econometric work to the Senate Estimate inquiry

There has been an enormous focus on one particular half-page of our report, the econometric analysis and appendix S of that report. The econometric analysis which was described in appendix S was undertaken with one purpose in mind, and that was to determine whether or not FuelWatch had caused any harm to Perth motorists since its introduction. It was designed to assess whether FuelWatch had caused any increase in prices to Perth motorists.

Source: Senate Committee Hansard, Standing Committee on Economics, Estimates, Thursday 5 June 2008.

On 15 April the Australian Cabinet decided to implement a national FuelWatch scheme. The press release announcing that decision said that⁵

Econometric analysis undertaken by the ACCC last year concluded that under the WA FuelWatch scheme the "relevant weekly average price margin was around 1.9 cpl [cents per litre] less on average".

Note the change in use of the econometric analysis. Initially the analysis was designed to reassure the ACCC that Fuel watch was not causing WA motorists to pay higher prices for petrol. Now the Government has transformed its use to support a conclusion that FuelWatch had reduced the price

⁴Mr Samuel, Chairman of the ACCC observed that,

the powers that were conferred upon us by then federal Treasurer, Peter Costello, powers of subpoena, powers of ability to obtain evidence from Informed Sources that had never been available to us before and powers to obtain data that enabled us to undertake econometric modelling.

⁵The Prime Minister also observed that

"The National FuelWatch Scheme is a key part of the Rudd Government's response to the ACCC report into the price of unleaded petrol."

of petrol to WA motorists. This change in use is the first evidence that we have a case of 'policy-based-evidence' rather than 'evidence-based-policy'.⁶

On 29 April the ACCC released a document "Further FuelWatch econometric analysis undertaken by ACCC". The conclusion of that document is as follows:

The purpose of this econometric analysis has been to satisfy the ACCC that the introduction of a FuelWatch scheme nationally would not, based on the experience in Western Australia, lead to consumers paying higher prices for petrol.

From the econometric analysis, on a conservative basis, the ACCC can say that there is no evidence that the introduction of FuelWatch in Western Australia led to any increase in prices and it appears to have resulted in a small price decrease overall.

The ACCC conclusion up until the 12 words "and it appears to have resulted in a small price decrease overall" are unexceptional. They represent a competition authority doing its job by reassuring the public that the extension of a scheme that it had previously viewed as possibly anti-competitive was unlikely to hurt consumers. The only issue that one could take with this statement is that the econometric analysis released by the ACCC is not sufficiently well explained to support the statement.

The last 12 words in the ACCC statement go well beyond the ACCC's brief this and provide comfort to a government that was in political difficulty over its decision to introduce a FuelWatch scheme. It is this part of the ACCC's conclusion that supports the contention that it provided 'policy-based-evidence' rather than the reverse.

3 Reconstructing the FuelWatch data

Chart S1 on page 376 of the ACCC (2007) report shows the weekly price margin used by the ACCC. Since the appendix is available electronically I convert into a bitmap and import it into computer programs that can be used to reconstruct the data.⁷ The two computer programs that I used are discussed in Appendix A where I also explain how those programs were used.

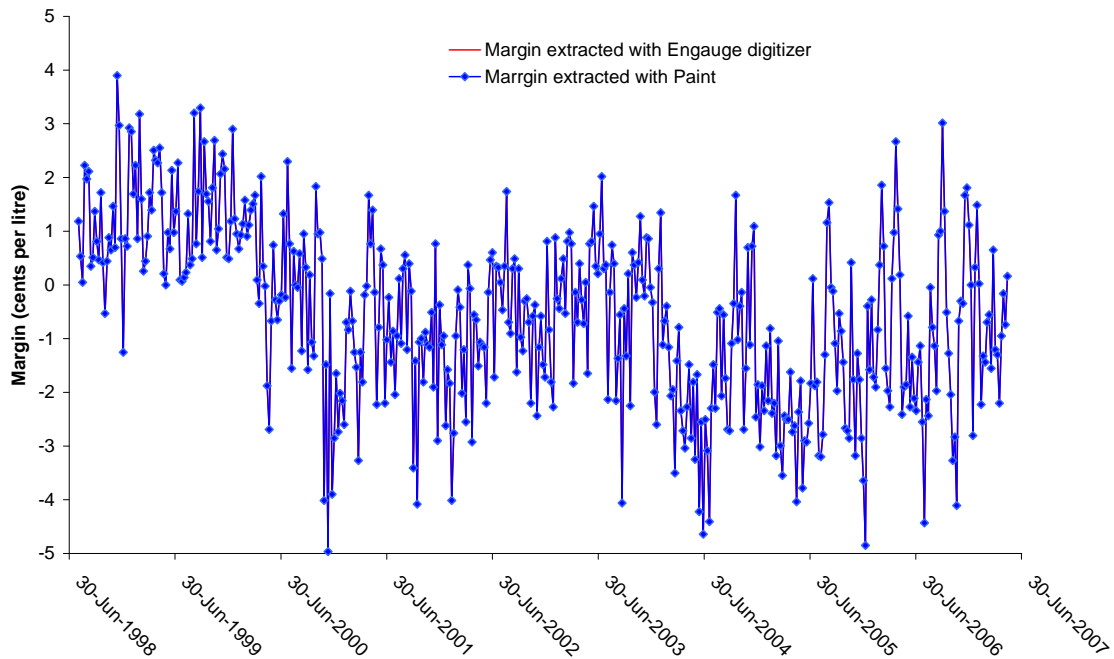
⁶The term 'policy-based evidence' comes from the UK's experience with 'evidence-based policy' under the Blair Labor government.

⁷This is an application of the work of the seventeenth century French philosopher Rene Descartes who showed that graphs and equations are two different representations of the same reality.

3.1 The reconstructed data

The data extracted using Paint and Engauge Digitizer are shown in figure 1. It is evident that the two series are so highly correlated that they cannot be distinguished on the graph. This visual information suggests that the data has been extracted with minimal measurement error.

Figure 1: Difference between the nominal retail margin on petrol in Peth and the eastern capital cities, reconstructed using Paint and Engauge Digitizer

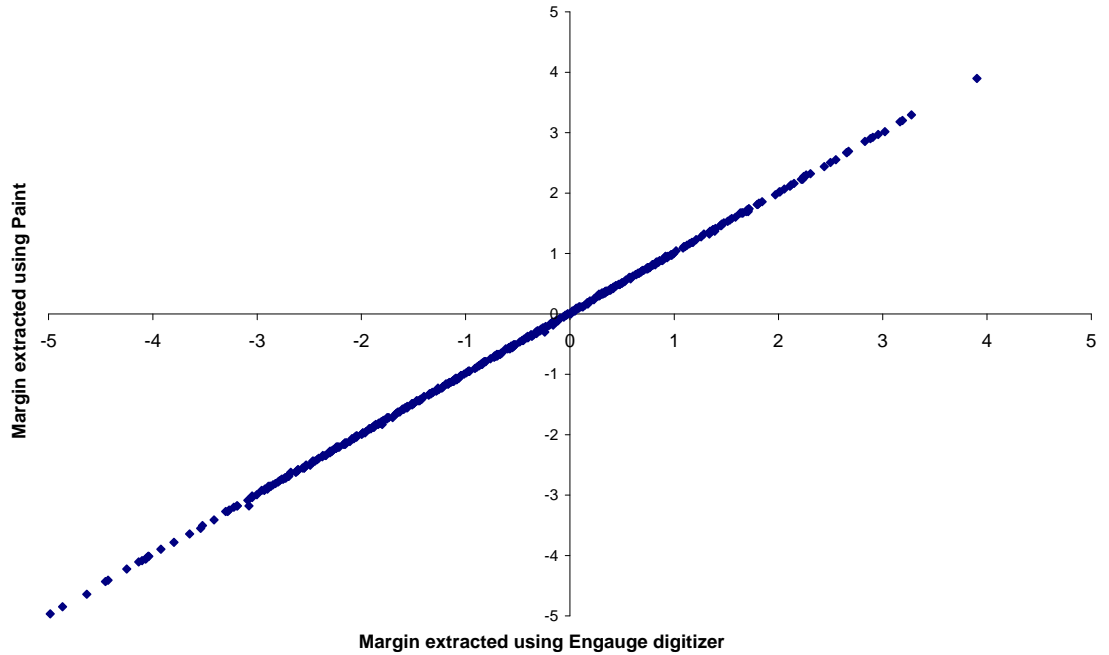


An important feature that is visible in figure 1 is that the margin seems to be more volatile towards the end of the period than it is at the beginning of the period. This feature will play an important role later in the paper.

Figure 2 shows an xy plot of the data where the data extracted using Engauge Digitizer are shown on the x axis and the data extracted using Paint are shown on the y axis. If there was no measurement error the data would lie along a 45 degree line through the origin. The extent of any measurement error is shown as deviations of the data from the 45 degree line through the origin. The correlation between the two series is 0.99997 confirming that the data lie almost on the 45 degree line. Clearly, the measurement error is

small.⁸

Figure 2: FuelWatch nominal retail margin reconstructed by two methods, x,y plot.



Given two series that are measures of the same concept, an improved estimate that has lower measurement error variance can be constructed as an average of the two series.⁹ This average is used as the dependent variable in this paper.

There is, however, one caveat that needs to be made here about the accuracy of the digitized data. The ACCC claims that their series runs from 1 August 1998 to 8 June 2007 a period of some 462 weeks whereas using both methods I can only distinguish only 459 weeks of data in the chart. I am reasonably confident that there is only 459 weeks of data in the Chart S1 of

⁸The average difference between the two series is fourteen thousandths of one cent with a standard deviation of of twelve one thousandths of one cent. This latter quantity is an estimate of the sum of the standard deviations of the measurement error in the two series.

⁹The extent of the reduction in measurement error achieved by averaging depends on the actual standard deviations of the measurement error. If the standard deviations are equal across the two series then the reduction in measurement error is about 29 per cent.

the ACCC report and that the data starts at 1 August.¹⁰ But the ACCC may have used different length of data in its econometric analysis. I doubt whether this difference affects my results.¹¹ But the ACCC can easily resolve the issue by doing what it should have done in the first place — release the data.

4 The ACCC's econometric analysis

In assessing the quality of the ACCC econometrics it is useful to set out objective criteria for determining good econometrics. I have chosen criteria that can be extracted from instructions that Professor Kenneth D. West, Ragnar Frisch Professor of Economics, at the University of Wisconsin gives to his students when preparing an empirical research paper.¹²

The objective of your project should not be to calculate many numbers, nor to perform a long series of tests, but rather to learn something about some interesting parameter(s). The paper will be evaluated with respect to clarity of exposition, thoroughness of description of the data and methods, competence in using the methods, and thoughtfulness in interpreting results. Complexity of economic theory and econometric methods does not carry weight in the evaluation. Appropriateness of the theory and methods to the project does carry weight. So does good writing.¹³

By way of background, low ranked academic journals, student essays and increasingly the applied work of government agencies are chock full of numbers and tests of every type. This approach might be labeled inoculation econometrics as researchers who apply this kind of approach are attempting to inoculate their work against criticism. However, in doing so the methods used and tests applied are often inappropriate. West is advising his students to avoid this approach as it rarely leads to understanding of the process being investigated. What West suggest's is that instead of the inoculation

¹⁰Chart S1 is poorly labeled as it does not report the period covered by the data.

¹¹It could marginally affect conclusions related to the timing of things.

¹²The course is for second year Phd students. Professor West teaches a similar course for 4th year undergraduates that sets similar standards. Standards comparable to those set by Professor West are set in the main Australian Economics Departments for work in 4th year Honours theses and above. It seems reasonable to hold that the econometric work of public agencies should meet these objective standards.

¹³Source: <http://www.ssc.wisc.edu/~kwest/teaching/econ880/econ880%20syllabus.pdf>

approach his students should focus on an investigative approach “to learn something about some interesting parameter(s) ”

4.1 The ACCC data

The ACCC constructed a series that represents the difference between the nominal retail margin on petrol in Perth (m_t^{Perth}) and the nominal retail margin on petrol in the eastern capital cities m_t^{East} . The ACCC described this data as follows

The data series was constructed using pricing information supplied by Informed Sources and Platts. The series tested was a measure of price margin that removes factors from the retail price that are beyond the scope of FuelWatch to affect, such as net taxes, fuel quality premiums and ex-refinery petrol prices. ACCC report p375.

Details of how the difference in the nominal price margin, which I denote as m_t , are calculated are provided on page 375 of the ACCC report and the ‘formulas’ are set out below.

$$m_t^{Perth} = (\text{Retail price-lagged Mogas95 price-net taxes - fuel quality premium})_{\text{Perth}}$$

$$m_t^{East} = (\text{Retail price-lagged Mogas95 price-net taxes - fuel quality premium})_{\text{Average of eastern capitals}}$$

$$m_t = m_t^{Perth} - m_t^{East}$$

The ACCC’s description of how the data is calculated is incomplete. The following are a selection of issues that are not addressed in the ACCC report.

How are the averages calculated? Are they geometric or arithmetic means? Are they weighted or unweighted?

What does the Mogas95 price represent? Specifically, does it include freight and freight related charges? Do such charges differ between the States? Might any differences in freight be accentuated by the increase in energy prices.

How are the taxes and charges calculated? What taxes are included? Have there been significant variations in those taxes? If so were the variations in taxes sufficiently large to have second round effects on petrol demand and thus petrol prices. Were these effects similar or different across the states? The new tax system was introduced in stages starting on 1 January 1999 and proceeding through until 1 July 2001. These included, among other things,

major cuts in income taxes that could have differential effects across states if the income distribution differs sufficiently across states. Similarly there were major changes to capital gains taxes which could affect petrol retail margins if there are differences across states in the intensity of land use in petrol stations or if there are differences in the expected capital gain on the land component of service stations. These data issues are not addressed in the ACCC appendix S.

4.2 Real versus nominal?

Nowhere in the ACCC appendix is it acknowledged that the price margins used in the report are nominal price margin whereas the relevant body of economic theory relates to real price margins.¹⁴ In the ACCC report the margin between the eastern states and Western Australia is just labeled as cents per litre. It is in fact a nominal margin so the unit of measurement 'cents' has a different purchasing power in every period. For later use I define the east-west difference real margin to service station owners (rm_t^P) and to consumers (rm_t^C) as follows

$$rm_t^P = \frac{m_t}{P_t^F}$$

$$rm_t^C = \frac{m_t}{P_t^C}$$

Where P_t^F is the price index for fuel in the CPI and P_t^C is the All groups headline measure of the consumer price index. These data are available quarterly so I interpolated the data using the formula

$$P_t^i = P_{t-1}^i \times \left(\frac{P_{Q(t)}^i}{P_{Q(t)-1}^i} \right)^{\frac{1}{n_{Q(t)}}}, \quad i = P, C$$

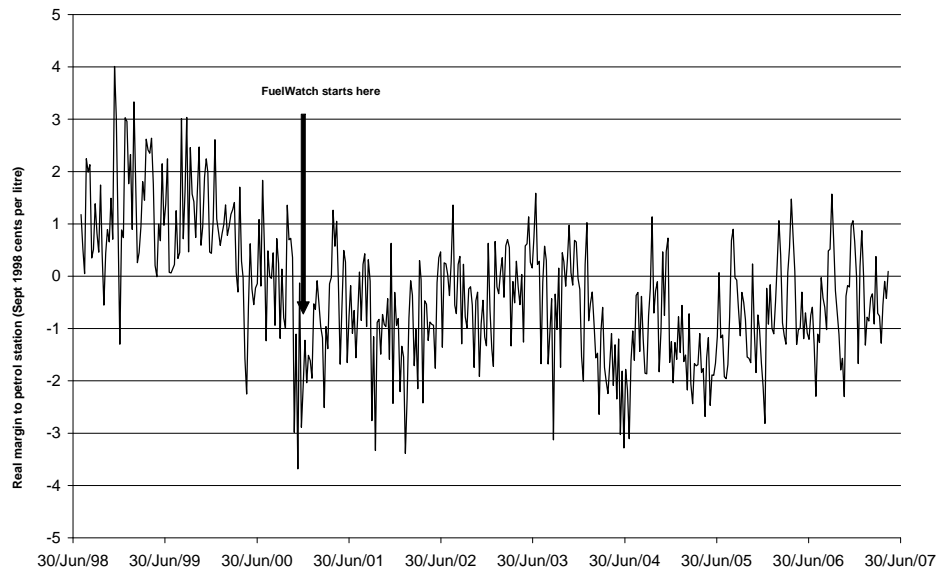
Where $P_{Q(t)}^i$ is the quarterly measure of the price level for the quarter corresponding to week t and $n_{Q(t)}$ is the number of weeks in the quarter.

¹⁴This feature of the ACCC's report falls short of best practice. For example, Professor Ken West's advice to students preparing an econometrics paper is

Be explicit about your data set. State the sample size. For time series, state whether the data are monthly, quarterly or annual, and whether or not they are seasonally adjusted. State the units of measurement. For example, if "income" is a variable, state whether it is measured in current dollars or constant 1992 dollars, and if it is per capita, say so.

The exact method used to converting the margin into real terms will matter empirically because the Consumer Price Index (CPI) rose by 29 per cent between August 1998 and June 2007 and the automotive fuel component rose by 78 per cent. Figure 3 shows rm_t^P the real (ie adjusted by the price index of automotive fuel) margin.

Figure 3: Difference between Perth and average of eastern capital cities in real retail margin to petrol station owners (rm_t^P), August 1998 to June 2007



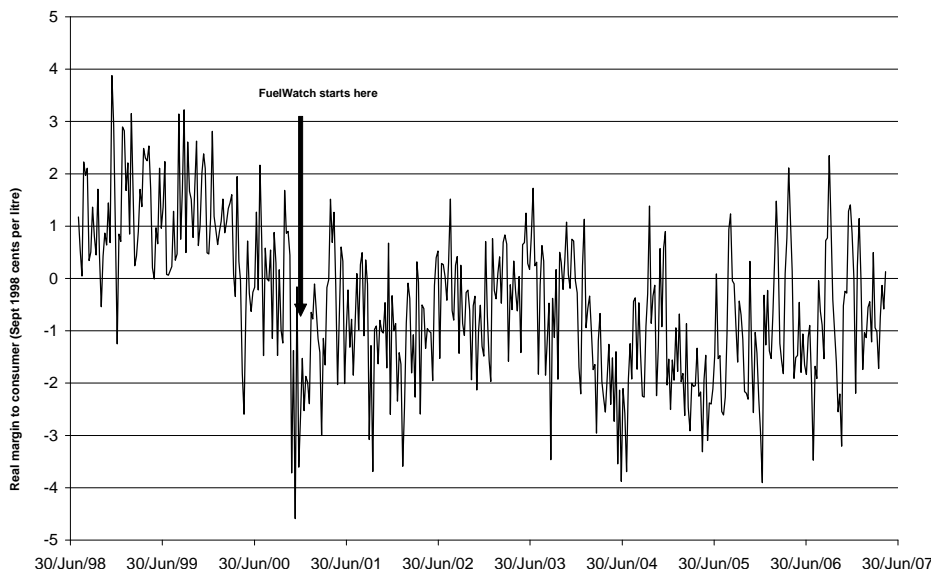
Notice that rm_t^P does not have the feature exhibited in the nominal retail margin (see Figure 1) that the variance seems to increase towards the end of the period.

Figure 4 shows rm_t^C the real (ie adjusted by the CPI) margin. Notice that there is some evidence that the variance in this series increases towards the end of the time period. But this effect is not as marked as it is in the nominal margin (see Figure 1).

What does economic theory say about the properties of these three series? Exit and entry of service stations is the economic process that can be expected to stabilise the real producer margin so that it fluctuates, without exhibiting a time trend, about a 'constant' that reflects the real retail margin at which two conditions are met

- entry of an additional service station is not profitable; and
- exit by one service station would make entry by another profitable.

Figure 4: Difference between Perth and average of eastern capital cities in real retail margin to consumers (rm_t^C), August 1998 to June 2007



I have put 'constant' in inverted commas for three reasons. First, changes in a variety of factors including the real interest rate, real (relative to petrol prices) wages, and the real rental rate's on land and capita will influence the real retail margin on petrol. Second the demand for petrol is heavily influenced by the calendar and this will cause the retail marginal to fluctuate with the calendar. The third reason is that if there is imperfect competition then it is possible that the petrol stations set prices that fluctuate through coordination.

If the real retail margin to service stations is approximately 'constant' then the nominal margin should vary with the price of fuel P_t^F . Specifically, the variance of the nominal margin should be proportional to the square of P_t^F . This feature makes the nominal margin (m_t) unsuitable for econometric analysis.¹⁵ A similar problem arises with rm_t^C as its variance will be proportional to the square of $\frac{P_t^F}{P_t^C}$.

Thus, the variable that is suitable for econometric analysis is the east-west difference in the real producer retail margin for petrol (rm_t^P). Factors such as interest rates and wage rates which are common between Perth and

¹⁵The essence of the problem is that P_t^F is almost certainly integrated of order one so that the variance of the nominal margin will be an integrated process. One can see this feature from a graph of the nominal margin.

the eastern capital cities should cancel out when this margin is calculated. Thus rm_t^P should be determined primarily by the following factors

- differences between Perth and the eastern capitals in calendar effects;
- differences between Perth and the eastern capitals in the effects of the new tax system that was introduced in 1999/2000; and
- the effects of FuelWatch.

Once these factors are controlled for the residual should be comprised of

- Measurement error arising because the data was digitized;
- independent and identically distributed shocks that represent primarily differences between Perth and the eastern capitals in petrol demand shocks
- a component that reflects differences between Perth and the eastern capitals in any coordination of petrol stations in setting prices.

4.3 Appropriateness of theory and method

4.3.1 Model

The model specified below encompasses the two models used by the ACCC¹⁶

$$m_t = \alpha_{ACCC} + \beta_{ACCC}d_{1t} + \delta_{ACCC}d_{2t} + \varepsilon_{ACCC,t} \quad (\text{Model 1 (ACCC)})$$

In responding to a draft of this paper an ACCC representative questioned whether model 1 is an accurate representation of the work of the. The argument put by the ACCC officer is that the analysis in Appendix S related a single structural break while Model 1 above has two structural breaks. This argument is incorrect for two main reasons. First, while the model in Appendix S did only have one structural break the document *Further Econometric Analysis Undertaken by ACCC* did contain a model with two structural breaks. Thus model one above encompasses these two cases and is an accurate characterisation of the ACCC model. Second, the single break case is easily obtained in from model 1 by setting $\beta_{ACCC} = 0$. This

¹⁶The two models I refer to are the one structural break model in Appendix S of the ACCC report and the two structural break model on the last page of the document *Petrol — Further Econometric Analysis Undertaken by ACCC* that was released on 29 May 2008.

is a restriction that can be tested and one will find that when d_{1t} breaks in early May 2000 the restriction $\beta_{ACCC} = 0$ is strongly rejected. Indeed from Table 1 one can see that the HACCC robust t-statistic is -4.2 . Another way of looking at this is that the single break model in Appendix S is not robust to the addition of an additional variable viz d_{1t} .

There is also an important logical reason for including the second structural break — the ACCC report relies on a form of logic which says ‘after this therefore because of this’. That’s how they can relate the shift in mean after the introduction of FuelWatch to FuelWatch. This form of logic is regarded with considerable suspicion in the sciences. A simple example will show why. Since Christmas cards always arrive before Christmas, this form of logic would say that Christmas cards cause Christmas. The fundamental problem here is that there is an omitted variable ‘expectation of Christmas’. Addition of a variable that measures ‘expectation of Christmas’ will solve the problem as after adding this variable one will no longer find that Christmas cards cause Christmas. This simple example shows how econometricians address their concerns about the ‘after this therefore because of this’ logic — they add further variables and see whether that changes the findings. That’s what I have done here.

In model 1, the model used by the ACCC, the nominal petrol retail margin m_t is the variable to be explained. But as discussed in the preceding section this specification is invalid and results in the variance of $\varepsilon_{ACCC,t}$ being proportional to to the square of P_t^F something that precludes using regression analysis. A sensible modification is to replace the dependent variable with rm_t^P which leads to model 2.^{17, 18}

$$rm_t^P = \alpha_P + \beta_P d_{1t} + \delta_P d_{2t} + \varepsilon_{P,t} \quad (\text{Model 2})$$

In model 2, rm_t^P is explained by our factors. The first of these is a constant (α_{ACCC}). The second and third factors are shift variables d_{1t} and d_{2t} .¹⁹ The

¹⁷One might ask why the margin is not set in constant real terms to the consumer. The reason suggested by economic theory is this. If petrol stations set the margin in this way then when the petrol price was increasing by more than the CPI they would go broke and when the petrol price was increasing by less than the CPI they would be earning excess profits that encouraged entry of new service stations thereby driving the margin down

¹⁸In this paper I have not sought to remove calendar effects because the ACCC did not do so and my objective is to replicate the ACCC work using the correct dependent variable. Omitting calendar effects will not bias the conclusions unless they are correlated with the introduction of FuelWatch something that seems unlikely. Incorporating calendar effects would, however, contribute to improving the efficiency of estimators.

¹⁹ d_{1t} takes the value 0 until week $t_1 - 1$ and it takes the value 1 from week t_1 through to week $t_2 - 1$.

d_{2t} takes the value 0 until week $t_2 - 1$ and it takes the value 1 from week t_2 through to

coefficients β_{ACCC} and δ_{ACCC} measure the extent to which the mean nominal retail margin for petrol is changed after dates t_1 and t_2 respectively. It is important to note that in this specification the ACCC are implicitly assuming that there are no other deterministic facts such as the calendar that explain the petrol prices. The fourth factor is the residual $\varepsilon_{ACCC,t}$ which represents the part of the margin that is explained by stochastic factors other than the constant and the two shift variables. The residual has unconditional mean of zero but may exhibit significant serial dependence. The unit of measurement for the coefficients is cent's per litre at fuel prices prevailing in the week starting 25 August 1998 - a real quantity. The residual $\varepsilon_{P,t}$ is also expressed in real terms.

4.3.2 Seasonal adjustment and calendar effects

Decisions about when people drive, how far they drive and the timing of their proximity to service stations are all affected by the calendar and by events such as public holidays and school holidays that are related to the calendar. Thus, demand for petrol shifts with the calendar causing the price of petrol moves with the calendar. To the extent that retail petrol prices are set above marginal cost through (imperfect competition) retail margins will also vary with the calendar.

The data supplied to the ACCC by Informed Sources was raw data that measures the difference between petrol retail margins in Perth and the eastern capital cities of Australia. This data will exhibit calendar effects unless the calendar effects on retail margins in Perth are exactly the same as those in the eastern capital cities something that is unlikely to be true. Thus it is important to check whether calendar variation is evident in the data. This is more tricky with weekly data than with monthly or quarterly data as the number of weeks in a year varies from 52 to 53 whereas the number of months and quarters in a year is constant across years. The simplest way of checking for seasonal effects is to construct a variable s_t that measures the proportion of a year that has elapsed (proportion of days to 31 December). A simple regression of rm_t^P on s_t and constant returns a coefficient of -2.24 with robust t-statistic of -2.6 which allows us to reject the null hypothesis of no calendar effects.²⁰

the end of sample.

²⁰In commenting on a draft of this paper an ACCC officer questioned whether the test for calendar effects is useful and suggested replacing the variable s_t with weekly dummy variables. One can certainly do this but it is not the most effective way of dealing with this issue. I have experimented with alternative methods and I am satisfied that the statement made in the paper about calendar effects is a conservative one. I am also convinced on the basis of that experimentation that allowing for those calendar effects in a more

The ACCC makes no allowance for calendar effects in their modelling however as mentioned earlier this misspecification will mainly affect the efficiency of the estimator.

In commenting on a draft of this paper an ACCC officer questioned whether the test for calendar effects is useful and suggested replacing the variable s_t with weekly dummy variables. One can certainly do this but it is not the most effective way of dealing with this issue. I have experimented with alternative methods and I am satisfied that the statement made in the paper about calendar effects is a conservative one. I am also convinced on the basis of that experimentation that allowing for those calendar effects in a more sophisticated way will not recover the original ACCC conclusion. Indeed, in these experiments I found that the data gives even less support to the ACCC conclusion.²¹

4.3.3 Testing for stationarity

In essence a data generating process is stationary if the moments do not change over time.²² Much econometric work focuses on the first two moments: the mean and the variance. One can think of non-stationarity as arising when these moments are functions of time. Two types of non-stationarity are potentially present in the FuelWatch data. The first of these are structural breaks in the mean. These could arise for example if

1. Changes in tax policy had different effects on petrol station margins in Western Australia to Eastern Australia. Importantly, the tax systems was changed extensively through the period 1999 to 2001 with major changes being made in mid 2000.
2. New entrants to the industry such as Coles caused a reduction in the margin obtained by Western Australian service stations;
3. The FuelWatch scheme had its intended effect to reduce service station margins.

sophisticated way will not recover the original ACCC conclusion. Indeed, it suggests that the data gives even less support to the ACCC conclusion.

²¹I have not included this additional material because the paper is already technical and making it more technical would make it less accessible to a broad audience.

²²This statement assumes a) that moments are finite; and b) that the probability distribution of the data is determined completely by the moments. A more precise, but less readable, statement can be applied to the more general case where neither of the assumptions hold.

The second form of non stationarity arises where the variance of a series is a function of time. This arises in two main way. The first is if the data is a deterministic function of time so that $rm_t^P - d_t$ is stationary where d_t is deterministic function of time. Because of the effect of the calendar on petrol demand this is an entirely plausible specification.

The second way that the variance of rm_t^P could be a function of time arises where $rm_t^P - d_t$ can be written as the cumulation of a stationary series.²³ In this case the series can be written as²⁴

$$\Delta (rm_t^P - d_t) = \pi (rm_{t-1}^P - d_{t-1}) + \sum_{i=1}^q \gamma_i \Delta (rm_{t-1}^P - d_{t-i}) + \eta_t \quad (1)$$

Where π is the parameter of interest. If $\pi = 0$ this means that the series rm_t^P has a unit root and the variance of rm_t^P will have a time trend.

This is a specification that is often found in macroeconomics but we should question its plausibility for this data. If $\pi = 0$ it means that the variance of $(rm_t^P - d_t)$ is proportional to time. So that at June 30 2007 the variance of $(rm_t^P - d_t)$ would be 459 times as high as in the week beginning 1 August 1998 something that does not seem consistent with free entry into and exit from petrol retailing.

There are a number of technical deficiencies in the ACCC testing for unit roots. These are discussed in appendix B where I show that if one corrects these technical deficiencies then the test applied by the ACCC would actually find that the east-west difference in the nominal margin (m_t) has a unit root as is suggested by the theory. However, this finding is moot as the ACCC applied the wrong test, incorrectly to the wrong series.

One error that is so striking that it requires some discussion is the fact that the ACCC, in a paper exploring possible structural breaks, used a test that is valid only where there are no structural breaks. An augmented Dickey-Fuller test allowing for two known structural breaks can be conducted via the ADF regression (1) where the deterministic component is

$$d_t = \alpha_0 + \alpha_1 d_{1t} + \alpha_2 d_{2t}$$

I applied this test to rm_t^P which is the correct variable. Critical values for the ADF test with two known structural breaks can be obtained via simulation. For series with 459 observations the 1%, 5% and 10% critical values are -4.08 , -3.44 and -3.12 respectively. A lag length of 4 weeks was

²³Strictly speaking the stationary series must also have positive spectral density at frequency zero.

²⁴ Δ is the difference operator. So $\Delta x_t = x_t - x_{t-1}$.

the smallest that produced a reasonably flat spectrum for rm_t^P . With $q = 4$ and two structural breaks allowed for the ADF test statistic is -6.82 leading to a rejection of the null hypothesis of stationarity. So we can conclude that once structural breaks in the mean are allowed for the real producer retail margin on petrol is integrated of order zero as is suggested by economic theory.

4.3.4 Thoroughness of description of methods

As discussed above and in appendix B, ACCC Appendix S leaves out much of the information that is necessary to judge whether the unit root testing has been applied correctly. A similar problem arises with the Newey West standard errors that are reported in Table S2. The ACCC recognize that serial correlation is present in the data and seek to correct that using Newey West standard errors. There are two issues here,

1. The key parameter in calculating the Newey West standard errors is the window width used in the calculation. The ACCC does not tell us what window width was used or how it was determined. This departs from best practice in this area.
2. The ACCC does not provide diagnostics that might help us understand how important is the problem of serial dependence in the data.

The ACCC report says that they searched for endogenous structural breaks. However, no information is provided on how this search was conducted. A search that attempted to minimize the sum of squared errors in Model 1 would be incorrect as it would not allow for the presence of serial correlation in those errors.

4.3.5 Estimating the effect of FuelWatch for known break dates

Here I replicate the ACCC analysis using the correct dependent variable (rm_t^P). I assume that there are two break dates one in the first week of May 2000. The other in the first week of January 2001 reflecting the introduction of FuelWatch in WA. The estimates, robust standard errors²⁵ and robust confidence intervals are reported in Table 1.

²⁵The standard errors are calculated making use of the fact that the regressors are deterministic. So that the covariance matrix for the estimators is $(X'X)^{-1}X'\Omega X(X'X)^{-1}$ where $\Omega_{ij} = E\varepsilon_{P,i}\varepsilon_{P,j}$. Consistent estimators of the first q autocovariances were used to construct $\hat{\Omega}$ resulting in a feasible and consistent estimate of the covariance matrix using a $q = 17$. The covariance matrix hardly changes for values

Table 1: Estimated coefficients of Model 2

Coefficient	Estimate	Robust	
		Standard Error	95% Confidence Interval
α_P	1.23	0.22	(0.79, 1.67)
β_P	-1.68	0.40	(-2.48, -0.88)
δ_P (FuelWatch)	-0.29	0.36	(-1.01, 0.43)

Since the robust 95% confidence interval includes zero it is not possible, based on this data, to say as the ACCC did that the WA FuelWatch scheme did not act to increase the real retail margin for petrol in Perth. The mean estimated effect is for a reduction of 0.29 1 August 1998 cents which translates into about 0.52 of a cent in mid 2008.

One possibility is that the estimated effects of FuelWatch are biased downwards because although the scheme came into effect on 2 January 2001 it may not have had an effect until some weeks later. To check this I allowed the introduction of the FuelWatch to vary from 13 January 2001 to 3 March 2001. The resulting estimates for the FuelWatch coefficient ($\widehat{\delta}_P$) are -0.24, -0.21, -0.18, -0.14, -0.14, -0.14 and -0.16. These figures suggest that -0.29 of a cent per litre is the maximum reduction in the real margin generated by the FuelWatch scheme.

The analysis above does not allow for calendar effects or for additional structural breaks nor does it make the location of the structural breaks endogenous. My assessment is that such extensions to the analysis would serve to cast further doubt on the ACCC conclusion that FuelWatch did not act to increase the real retail margin for petrol in Western Australia.

4.3.6 Overall assessment

The approach of the ACCC does not stand up well against the criteria set out at the beginning of this section. Specifically,

1. The ACCC uses the wrong dependent variable (the east-west difference in the nominal retail margin for petrol) rather than the east-west difference in the real producer retail margin for petrol.
2. Use of the wrong variable results in more than misspecification because it means that the assumption of stationarity necessary for the econometrics cannot hold.

of q between 10 and 100. Similar standard errors are obtained using the Newey West procedure.

3. The ACCC procedures cannot detect such non stationarity because they apply the wrong test to the wrong variable. Moreover, their application of the test is flawed because they seem to use an incorrect lag length.
4. The ACCC approach is one in which it has sought to use tests to 'inoculate' itself against criticism rather than trying to "learn something about some interesting parameter(s)". This is the source of many of the econometric flaws in the ACCC analysis.
5. The reporting of the ACCC econometrics is substandard.
6. When I redo the ACCC analysis using the correct dependent variable and allowing for two structural breaks the data does not support the ACCC conclusion that the FuelWatch scheme did not lead to increased petrol prices in Perth.

That said Appendix S of the ACCC report provides a thoughtful discussion of the caveats to their work. However, some important caveats such as the issue of nominal versus real margin and the possible effects of the tax system introduced in 1999-2001 are neglected.

A natural but incorrect response is to blame the econometricians who undertook this work. The public service has many very able econometricians and economists and I draw no adverse conclusion about those who undertook this work or those who reviewed it. Econometrics is a difficult art. The quality of econometric work is almost always improved by public scrutiny and particularly by seminars and workshops involving academics. By the same token the quality of econometric work is almost always reduced by factors such as secrecy, the issues being politicised, ambitious junior ministers seeking to make a name for themselves and when the 'evil trio' of the 'managers manager', lawyers and 'econometric guns for hire' are involved.

5 Who are the experts?

Stilgoe Irwin and Jones (2006) observe that

The physicist Werner Heisenberg defined an expert as 'someone who knows some of the worst mistakes that can be made in their subject and who manages to avoid them'. Expert wisdom is about navigating uncertainty, reminding people in power what we still might not know, in addition to what we think we know, and cautioning against complacency.

This is a reasonable starting point for discussion but Heisenberg has defined someone who has expertise rather than someone who has sufficient knowledge of the particular issue as to be regarded as an expert.

For econometric analysis the expert is the person who actually undertook the econometrics and thus will have an intimate knowledge of the data and the methods used. Such a person should also have an appreciation of the qualifications to the analysis, the mistakes that can be made and what was done to make the analysis robust to these potential qualifications. Whether the person actually has the knowledge described can only be discovered by asking them. It is rare to see such people give evidence at public inquiries because they are typically very junior in the bureaucratic hierarchy. They are, however, experts on the particular matter.

The second category of expert is the manager who supervised the person(s) who performed the econometrics. Such people are experts because they know the procedures and protocols that were used to ensure that the econometrics were undertaken according to best practice. It is also relatively rare for such people to appear at public inquiries.

The third category is the manager's manager. Such people may or may not have expertise as defined by Heisenberg but they definitely are not experts because they have little direct knowledge of what was done or how what was done was supervised. It is usually the manager's manager or some person even further removed from direct experience with what was done who gives the evidence at bureaucratic inquiries. Because they are so removed from what was done the testimony of such people is largely useless in evaluating whether the process under study was capable of turning information into robust and tested evidence.

The ACCC's econometric work was discussed in the Senate estimate's economic committee on 5 June 2008 which heard evidence from a number of officials including the ACCC Chairman, the Commissioners who were responsible for the ACCC report and senior Treasury officials. An interesting question that is as yet unresolved is whether any of these officials that appeared before the Senate Committee as experts actually did the econometrics for appendix S of the ACCC report. A related question is whether any of these officials was the immediate manager of the person(s) who did the calculations. If the answer to these questions turns out to be no then these officials may or may not have expertise but they cannot be said to be experts in the econometric analysis of FuelWatch.

There is a fourth category of people with expertise that needs mentioning - academics. Academics were excluded from the process because the ACCC refused to release the data used in its report. In evidence to the ACCC Mr. Samuel effectively dismissed academics as a public nuisance stating that:

I am not in a position to be able to say that we would make our data and our methodology available to anyone out in the public arena. We are not prepared to make all this available for any economic modeller or any economic student to simply go through and then to engage the already heavily worked staff of the ACCC in debate on these issues. The commission of inquiry needed to satisfy itself that the work that was done was robust and they have done just that.

In essence Mr. Samuel is saying that the ACCC staff who defend ordinary Australians against anti-competitive behaviour by business need to be protected from competition in ideas and analysis that is provided by scrutiny from academics, students, journalists and the general public.²⁶ It is evident from the previous section that the ACCC's would benefit from greater exposure to academic econometricians. It is also evident that public policy would benefit from such scrutiny.

6 Reflections on evidence-based policy making in Australia

Shortly after winning the 2008 election the new Prime Minister reaffirmed his commitment to evidence based policy making:

I believe in evidence-based policy not just sort of grand statements. Prime Minister Kevin Rudd, *7.30 Report* 27/11/2007

²⁶Mr Samuel also said that

.... if there is an economic consulting firm that wants to do its own analysis of the impact of FuelWatch in Perth then they can approach Informed Sources. Not that it is our right to do so anyhow, but we have said to Informed Sources, 'You are absolutely free to make whatever data you want available to whomever you want on whatever terms and conditions you want to make it, so they are entitled to do their own research and use whatever test they want to use and whatever methodology they want to use.

This statement is disingenuous as the ACCC had to subpoena the data from Informed Sources who are unlikely to make the data publically available. Professor Joshua Gans states on his Blog

I can now confirm that Informed Sources is unable to provide the same data the ACCC had access to, to third parties such as academics. Source: <http://economics.com.au/?p=1439>

Implicit in the new Prime Minister's statement is the assertion that previous Australian governments were somehow deficient in not making adequate use of evidence in policy making. This claim is not true - even before federation state governments were making extensive use of evidence in decision making. For example, Sir Timothy Coghlan the New South Wales Statistician between 1886 and 1904 developed a macroeconomic data collection that is similar to national accounts; See Haig (2006). Haig observes that in the late 1920s Brigden, Giblin and others were using "Coghlan's quantitative approach in a study of the effects of the Australian tariff". In the Harvester Case in 1907, Justice Henry Bournes Higgins set out to use empirical evidence to determine what constituted "fair and reasonable" wages.

The Campbell Committee of inquiry final report was delivered on November 1981 assembled a vast body of evidence on the financial system and the need for change. The Hawke Labor government that implemented the findings of this report was arguably the pinnacle of evidence-based decision making in Australia. But even the Hawke Labor government experienced problems with the evidence-based approach. Most notably it found the work of the Bureau of Labour Market Research politically unpalatable and scrapped that body.

Experience suggests that several issues arise with the evidence-based approach.

First, public policy decisions create winners and losers who have considerable incentives to distort the evidence or its interpretation. These winners and losers may be commercial interests who have financial incentives to pay considerable sums of money to influence or shape 'expert' opinion. The battle of 'experts' that ensues rarely sheds light on how the public interest can be furthered.

Political interests also have strong incentives to influence or shape expert opinion so as to achieve a political agenda. Ambitious junior ministers, for example, have a strong incentive to achieve quick 'results' thereby enhancing their promotion prospects. As Stilgoe, Irwin and Jones (2006) observe it is very difficult for bureaucrats to 'lean against the wind' created by the ambition of a minister. In the FuelWatch case there is a cadre of well trained public servants in Treasury, Finance, Prime Minister and Cabinet, ACCC and the various energy, resource and industry departments who have sufficient econometric expertise to see through the flaws in the ACCC's published work. Indeed the leaked departmental coordinating comments suggests that at least four departments applied their econometric expertise to reach the correct conclusion. So the problem was not one of knowledge or training it was one of incentives. Any bureaucrat who stood up to point out the errors in the ACCC report would be doing more than 'leaning against the wind'.

In standing up to a new government that may be in office for two or more terms they would be making what is known as ‘a career limiting move’.

Second, after having made a decision for narrow political purposes governments have a strong incentive to ‘dress that decision up’ in the clothes of respectability by selectively using facts that support the decision. This is clearly what happened when the ACCC appendix was converted from saying that there was no evidence that FuelWatch increased petrol prices to the additional document which said that it may have lowered petrol prices

The problem is that once a government engages in such practices it becomes difficult for voters to distinguish between true evidence-based policy making and the sham that is ‘policy-based evidence’. The result is a general loss of faith in government and in its claims to base decisions on evidence.

Such financial and political incentives make the ‘realpolitik’ of evidence-based-policy making much more difficult than it might seem. Some of the problems experienced in the United Kingdom with evidence-based-policy are summarized in the following paragraph from Stilgoe, Irwin and Jones (2006)

In November 2006, a report from the Commons Select Committee on Science held a mirror to the fashion for evidence-based policy. In some areas, the committee argued, ‘evidence-based’ has become a way to justify policy rather than a way to make policy – the evidence is found to suit the decision. Evan Harris, a committee member and Liberal Democrat science spokesman, said that the way some policies claimed to be evidence-based was a ‘fraud which corrupts the whole use of science in government’.

Commitment to evidence-based-policy has a natural attraction to voters it conveys the hope that public policy decisions can be objective, rational and informed by careful quantification of the evidence. But some government decisions have to be made on the basis of ideology, convenience or whim because the evidence is inconclusive. The problem with the unqualified mantra of ‘evidence-based-decision’ making is that it creates a false hope that governments and their advisors cannot possibly meet. When they are inevitably seen to fail in reaching this goal governments run the risk of damaging their credibility and that of the evidence based approach.

Based on the preceding analysis I conclude that if the Australian government is serious about ‘evidence-based-policy’ making it should:

1. Adopt the standard that it will not claim that a decision is evidence-based unless:

- (a) It releases the data and analysis on which that decision was based;
and
 - (b) It has subjected the analysis to truly independent review.
2. Heed the lesson in the fact that the advice from his department that the Treasurer dismissed as ‘too academic’ proved ultimately to be correct;
 3. Recognize that Cabinet coordinating comments which are often partly anonymous serve both to protect a government against error and also partly protect bureaucrats who provide accurate but unpalatable advice.
 4. Recognize that some decisions cannot be evidence-based because there is insufficient evidence on which to base a decision or the evidence is inconclusive. In these cases ‘fudging’ the evidence or overstating what conclusion the evidence supports acts to corrode and corrupt the use of evidence in shaping policy; and
 5. Recognize that academics provide an invaluable resource capable of the independent reviews necessary to catch cases of ‘policy-based-evidence’. Providing such academics with access to resources and the data is fundamental to implementing evidence-based-policy making.

7 Conclusion

Experience shows that good decisions are based on evidence that is tested. The ACCC econometric ‘evidence’ used to justify the FuelWatch decision was not adequately tested or evaluated. The actual econometrics was deeply flawed. When the econometrics is redone the ACCC conclusion can no longer be supported.

This case study replicates the UK experience where it was found that ‘expert’ analysis was often misleading unless it was made robust by being exposed to public scrutiny and scientific analysis. For econometric analysis the key criteria is that the data is publicly available and clearly described so that the analysis can be replicated and extended if it is found to be unsatisfactory or incomplete. Until the government modifies its procedures to reflect this experience it cannot legitimately claim that its decisions are evidence based.

Appendices

A Programs used to digitize the FuelWatch data in Chart S1

Two computer programs, Paint and Engauge Digitizer, were used and the extracted data compared to check for errors. Both methods are described below.

A.1 Paint

Paint is found on most personal computers and has several important properties that are useful in manually extracting data:

- when one clicks on the line icon the screen pointer turns into cross hairs;
- It is possible to place a one-pixel width grid over the graph.
- the resolution can be magnified up to 8 times so as to help resolve difficult cases.
 - For most cases a magnification of 2 times was sufficient.
- by placing the cross hairs over a point it is possible to read the location in xy plane of each point on a graph.

The features outlined above mean that by placing the cross hairs:

- at the edges of the axis it is possible to read the scale of the graph in pixels; and
- on each data point it is possible to read the (x, y) coordinates in pixels.

A simple linear equation converts the data expressed in pixels back into the price margin expressed in cents per litre of petrol.

A.2 Engauge Digitizer

Engauge Digitizer is shareware that allows users to digitize graphs.²⁷ To use digitizer one simply,

- imports the graph;
- places the cross hairs of the pointer over the maximum and minimum values of the x and y axis and click the mouse to record the location of the axes.
- for each data point, place the cross hairs over the data point and click the mouse to record the (x, y) coordinates on the graph;²⁸
- if an error is made it can be easily corrected; and
- the data can be copied onto the computers clipboard and then pasted into a spreadsheet.

The two programs provide independent ways of extracting the data from a graph. For this reason the error made with one program should not be correlated with the error made from the other program.

²⁷ Available at <http://digitizer.sourceforge.net/>.

²⁸ Engauge Digitizer can also be used in automatic mode where it is "trained" to identify certain datapoints and then extract the data automatically. This feature was tried but did not produce results as accurate as those from the manual location of datapoints.

B Technical aspects of the ACCC stationarity tests

The ACCC tested for the presence of a unit root using the regression (1). This involves choosing the deterministic component d_t and an integer q to represent the number of lags. The choices made by the ACCC were

1. To assume that d_t is a constant and there is no structural break;
2. To set the maximum q according to the formula $q_{max} = Int \left[12 * \left(\frac{T}{100} \right)^{0.25} \right]$. Where $Int[a]$ refers to the integer part of the number a . In this case $q_{max} = 17$.
3. To choose q by a sequential t-test starting at q_{max} and using a significance level of 10-percent. Unfortunately the ACCC does not say what the actual value of q was chosen. Nor do they say in which direction the sequential testing was run: one can either start from $p = 0$ and stop when the t-statistic on the longest lag is insignificant at the 10 per cent level — called the 'bottom up' approach. Or one can start at $p = p_{max}$ and keep reducing p by one until the t-statistic on the longest lag is significant at the 10 per cent level — often called the 'top down' approach. The latter approach when used with the particular t-test can be shown to overestimate the number of lags something that produces inefficiency.²⁹ The former approach of starting at $p = 0$ runs the risk of selecting p that is too small and thus not controlling adequately for serial dependence. The result of using the 'bottom up' approach is a test that is invalid.

The so called augmented Dickey Fuller (ADF) test then gathers statistical evidence gathered as to whether $\pi = 0$ or $\pi < 1$.

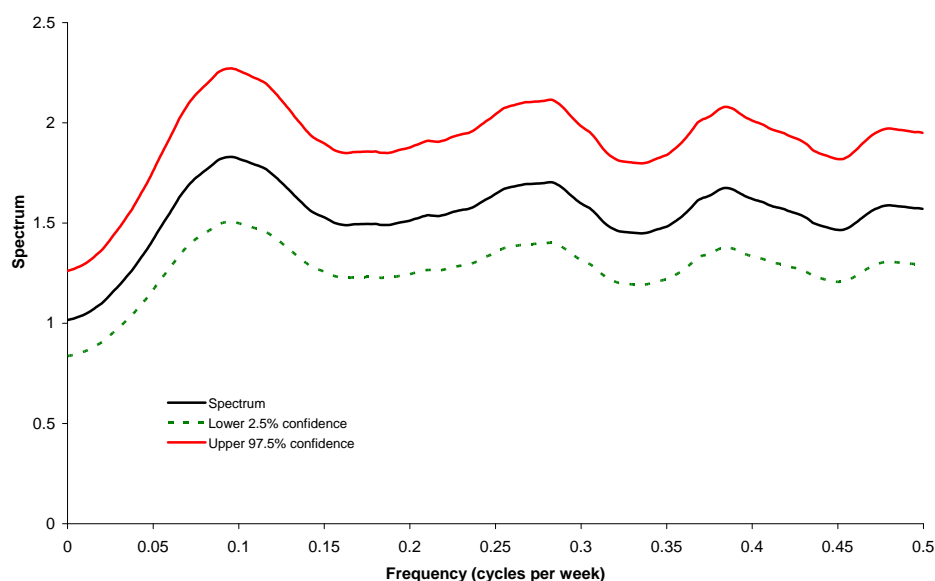
When I repeat the exercise as described by the ACCC I only find that I can reject the null hypothesis (using a 10 per cent significance level) if I use the invalid bottom up approach to selecting q . If I do this I find that $\hat{q} = 5$ and $\pi = -0.163$ with standard error 0.043 so the ADF t-statistic is -3.8 which would lead to rejection of the null hypothesis of a unit root.

For the reason given above one might suspect that q was chosen too small and thus there remains unaccounted serial dependence in the residuals from the ADF regression. One can check this by inspecting the spectrum of the

²⁹Using a t-test with 10 per cent significant level is approximately the same as using the Akaike information criteria.

residuals which is shown in Figure below together with confidence bands. If there were sufficient lags in the ADF regression (and the deterministic component was correctly specified) then the spectrum would be flat. As can be seen from Figure this is not the case confirming that the bottoms up approach yields too small an value of q to adequately purge the serial correlation from the residuals of the ADF regression.

Figure 5: Spectrum of residual from ADF regression for nominal retail margin on petrol when $q = 5$



The procedure outlined above of using some model selection criteria to chose q is frequently observed in the literature. But this approach confuses parsimony (choice of a small q) with selection of the smallest q that is consistent with a flat spectrum. It is the latter feature that is necessary to ensure that the statistical tests under which the ADF test is valid are met. For the reason just stated I will use the 'flat spectrum' criteria to inform my choice of q .

Starting from $q_{max} = 17$ one finds that the spectrum of the ADF residuals is 'flat' until $q = 12$ so this suggests using $q = 13$ weeks (the smallest value of q that yields a flat spectrum). With this choice of q the ADF statistic is -2.1 and since this is greater than the critical value of -2.57 we cannot reject the null hypothesis that the nominal margin contains a unit root. Clearly this is at odds with the ACCC finding and suggests that the nominal margin m_t contains a unit root as would be expected if Model 2 is the true model and

P_t^F contains a unit root. The existence of a unit root in m_t is consistent with the visual in evidence in Figures 1 and 4 which suggests that the volatility of the series is increasing over time. The leading hypothesis about the cause of this increasing volatility is that Model 2 is indeed the true model.

C Bibliography

ACCC 2007, *Petrol Prices and Australian Consumers: Report of the ACCC inquiry into the price of unleaded petrol*, 18 December 2007.

ACCC 2008, *Petrol—Further Econometric Analysis Undertaken by ACCC*, 29 May 2008.

Haig B., (2006). “Sir Timothy Coghlan and the Development of National Accounts”, *History of Political Economy* 2006 38(2):339-375

Saint Vincent Millay E. (1939), Sonnet X, *Huntsman, What Quarry?* (New York: Harper, 1939), p. 697.

Stilgoe, J., A. Irwin and K., Jones 2006, *The Received Wisdom: Opening up Expert Advice*, Demos.