Deloitte Access Economics

Review of NLTF forecasting model

New Zealand Ministry of Transport

11 October 2012

Deloitte.

Contents

1 Introduction

Deloitte Access Economics (DAE) was contracted by the New Zealand Ministry of Transport (the Ministry) to provide a review of the National Land Transport Fund revenue forecasting model, which DAE previously prepared for the Ministry.

The previous model used data current to the end of 2010. With data now available through to mid-2012, the first step in the review process was to incorporate the 'correct' data into the original model and compare the results to those originally provided in early 2011.

The next step was to re-estimate the original model equations with the longer time span of historical data to determine if the originally specified models still provide a reasonable 'fit' in predicting the dependent variable. This is a particularly important task given the unexpected volatility that has befallen the global economy in the last few years.

For some models, econometric testing was undertaken to determine whether the relationship between the input variables had significantly changed from the original estimation.

The main deliverable of this project is thus an updated model incorporating changes as discussed with the Ministry over the course of the review.

This short report accompanies the revised model and documents some of the key changes and findings. Chapter 2 discusses the first stage of the review (comparing 'actual' data with the original forecasts), and Chapter 3 discusses the second stage (econometric evaluation of original model equations).

2 Model errors

2.1 Sources of error

In this modelling exercise, errors can occur in one of two ways: either the input data are wrong, or the model coefficients are wrong. To the extent that the model over or underestimates a variable due to incorrect inputs, that does not reflect a weakness with the model, but rather an unavoidable error caused by having to use estimates of key inputs.

Chart 2.1 shows how some of the key input variables to the various models have differed from their original forecasts.

Chart 2.1: Forecast errors for key input variables, 2010-11 and 2011-12
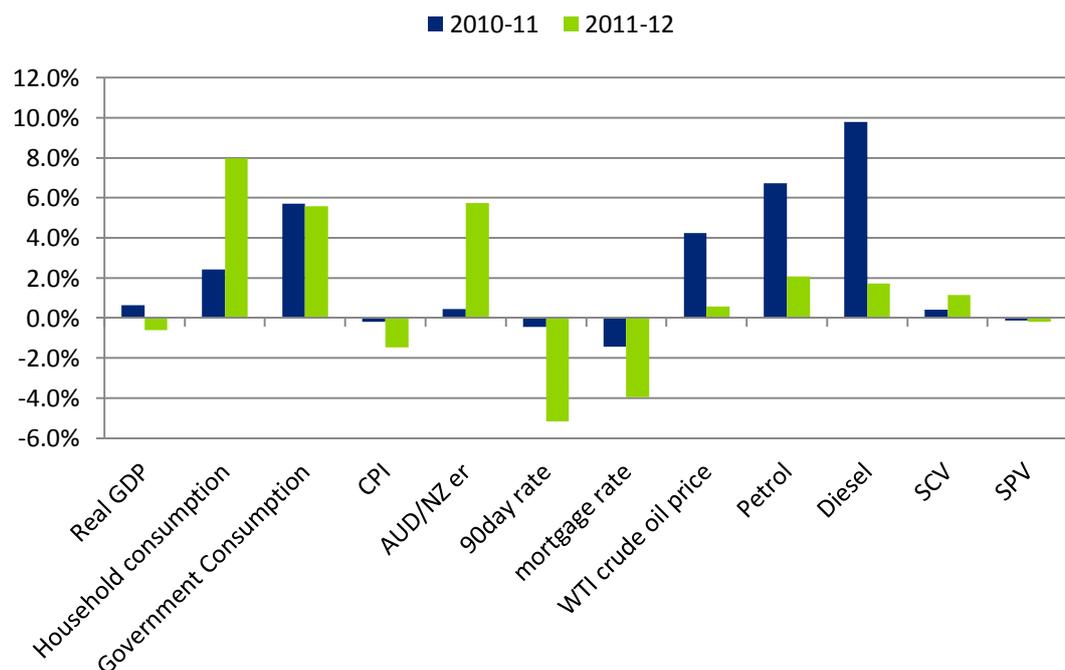
A positive error means the actual value turned out to be higher than forecast

The biggest deviations are in petrol and diesel prices, both of which were significantly underestimated. Crude oil prices have shown considerable volatility since the original forecasts, and this no doubt flowed through to significantly higher prices at the pump.

By running the original model with the 'correct' inputs we are able to estimate how much of the total error is caused by the model itself and how much of the error could potentially have been avoided had the data inputs been more accurate.

In the discussion below, 'original model' refers to the model originally sent to the Ministry in April 2011, while 'updated model' is the same as the original model except that known

values of input variables are used instead of estimated values for the period March 2011-June 2012.

2.2 Separating out the errors

The relative deviation between the updated model's (the original model with correct data inputs) revenue forecasts for 2010-11 and 2011-12 and the actual revenue outcomes for those years can be thought of as the model error. Even with perfect input data, the model itself is not perfect and so will always produce an error.

If the model's performance were to become substantially worse after adding true values for its inputs (i.e. the absolute deviations blew out), it would indicate that the model itself needed re-estimating. Table 2.1 and Table 2.2 show how updating key input data can affect the forecasts.

'Actual' is what we now know as the true value for the respective revenue head. Original forecast is the forecast originally made in April 2011. 'Original no RUC/PED indexation' shows what the forecasts would have been had RUC and PED rates not been indexed.[1] The 'updated forecast' is the forecast the model would have made had the data inputs been perfect.

Taking Light RUC as an example, revenue in 2010-11 ended up being $315 million. The original model forecast, made in 2011 with input data current only to December 2010 projected that revenue would be $322 million, a 2.2% error. But the original forecasts assumed that RUC rates would index as per usual, when in fact they remained constant through 2011. Had the original model correctly used constant RUC rates through 2011, the forecast would have been $319 million. So roughly half of the total model error is due purely to the RUC rate being incorrectly indexed in the original model. Finally, the forecast from the original model using perfect inputs was $317 million. So of the original $7 million error, $5.5 million was due to incorrect data inputs.

For most of the models, the above example holds, though this need not necessarily be the case. In some models, certain input variables turned out to be higher than originally forecast while others turned out to be lower than originally forecast.

---

[1] In the original model it was assumed that RUC and PED rates would rise in line with CPI. As it turned out that indexation did not occur – hence the original forecasts would have been lower had the 'correct' RUC and PED rates been used.

Table 2.1: RUC, MVR and FED revenue ($m), actual vs forecast

	Actual		Original forecast		Original no RUC/PED indexation		Updated forecast	
	2010-11	2011-12	2010-11	2011-12	2010-11	2011-12	2010-11	2011-12
LRUC	315,174	327,606	322,164	348,768	318,850	336,480	316,591	335,804
HRUC	708,154	727,487	686,813	745,706	681,520	709,791	702,853	749,923
MR1	188,799	191,775	185,215	185,222	185,215	185,222	187,483	186,786
MR2	23,366	23,742	22,999	22,633	22,999	22,633	23,324	23,019
MR13	6,586	6,622	6,561	6,748	6,561	6,748	6,597	6,939
PED	1,483,186	1,471,965	1,531,709	1,578,126	1,531,709	1,532,081	1,472,885	1,553,326
LPG	27,209	29,686	27,978	32,052	27,978	32,052	29,710	32,958
Total	2,655,535	2,682,084	2,694,140	2,824,524	2,685,667	2,732,081	2,647,416	2,797,236

Table 2.2: Deviation from true value ($m) (forecast minus actual)

	Actual		Original forecast		Original no RUC/PED indexation		Updated forecast	
	2010-11	2011-12	2010-11	2011-12	2010-11	2011-12	2010-11	2011-12
LRUC			6,990	21,161	3,677	8,874	1,418	8,197
HRUC			-21,340	18,219	-26,634	-17,696	-5,300	22,436
MR1			-3,584	-6,553	-3,584	-6,553	-1,316	-4,989
MR2			-367	-1,109	-367	-1,109	-42	-723
MR13			-25	126	-25	126	11	318
PED			48,523	106,160	48,523	60,115	-10,301	81,360
LPG			769	2,366	769	2,331	2,501	3,272
Total			38,605	142,439	30,132	49,997	-8,119	115,151

Comparing the deviations in Table 2.2 with the 'actual values in Table 2.1, we are able to break up the total model error into two separate components.

The 'Model error' is the unavoidable error; it is the error caused by the model itself not being a perfect representation of the independent variable. In any model with an R^2 of less than 100% (meaning that 100% of the variation in the independent variable is explained by the model), some degree of model error will be present. The 'economic' error is the avoidable (or more accurately, reducible) error; it is the error caused by data inputs being imperfect.

The far right column shows what the total model error would have been had the RUC and PED rates been implemented correctly (i.e. if they hadn't been indexed). For light RUC and petrol excise (in 2011-12), we can see that up to two-thirds of the total error was caused by the prices being incorrectly indexed in the original model. For heavy RUC, it is clear that errors in other input variables counteracted the effects of the price error.

Table 2.3 Sources of total error

|	Total error		Model error		Economic error		Total error with correct prices	
	2010-11	**2011-12**	**2010-11**	**2011-12**	**2010-11**	**2011-12**	**2010-11**	**2011-12**
LRUC	2.22%	6.46%	0.45%	2.50%	1.77%	3.96%	1.17%	2.71%
HRUC	-3.01%	2.50%	-0.75%	3.08%	-2.27%	-0.58%	-3.76%	-2.43%
MR1	-1.90%	-3.42%	-0.70%	-2.60%	-1.20%	-0.82%		
MR2	-1.57%	-4.67%	-0.18%	-3.05%	-1.39%	-1.62%		
MR13	-0.37%	1.91%	0.16%	4.80%	-0.54%	-2.89%		
PED	3.27%	7.21%	-0.69%	5.53%	3.97%	1.68%	3.27%	4.08%
LPG	2.83%	7.97%	9.19%	11.02%	-6.37%	-3.05%	2.83%	7.85%
Total	1.45%	5.31%	-0.31%	4.29%	1.76%	1.02%	1.13%	1.86%

3 Updating the model

3.1 Changes to data inputs

3.1.1 Petrol Excise

The single largest change in the data inputs has been the adoption of a revised historical series for PED revenues, now on a cash basis. The net effect of the changes is an increase in the overall level of the series over the last four years, and an increase in overall volatility across that time. As a result of this volatility the overall fit of the PED model has fallen since the original forecasts were made.

Updating coefficients and/or undertaking a complete re-estimation of the model would likely improve the overall 'fit' of the model. Such an exercise is beyond the scope of the current review; nonetheless, any improvement may be relatively minor given the increased volatility observed in PED volumes in the post Global Financial Crisis (GFC) period (Chart 3.1). This heightened volatility will reduce the overall fit of any model until a sufficiently long post-GFC timeframe has been observed (by which time the pattern may have also settled somewhat).

The existing model was tested for existence of a structural break, and none were found to exist, indicating no immediate need to alter the model specification.

Chart 3.1: PED volumes, 2005-2012
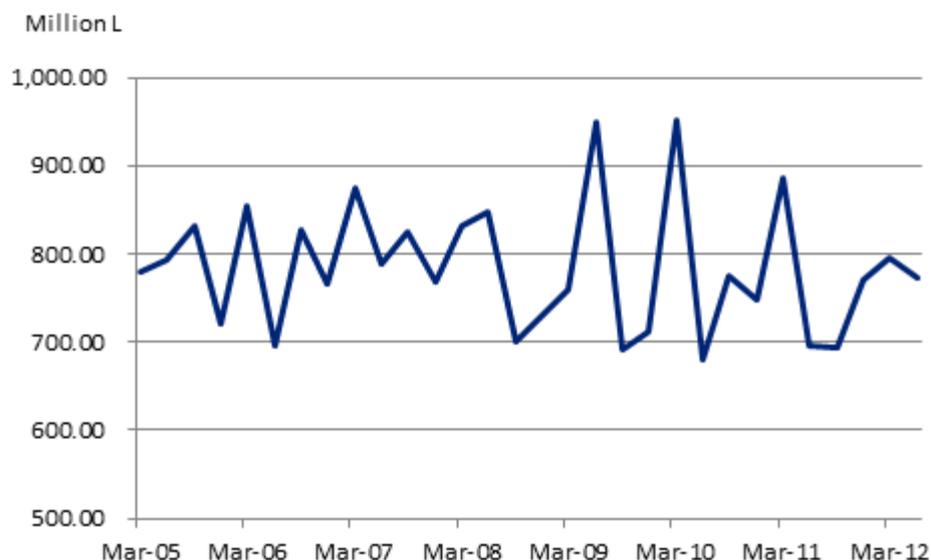
3.1.2 RUC Rates

The change in the RUC price series from a "policy price index" (or pure price index) to an "average user price" alters the economic interpretation of the price in the model, and thus also its expected impact in history and forecasts.

The model captures the expected response by consumers in response to a price change – hence, the original index was volume weighted so that by definition, changes in the index reflect only price changes. The 'average user price' is not a pure price effect and will change any time RUC volumes change – model coefficients based on this series would potentially have a spurious interpretation, because they would implicitly be taking into account volume changes when the dependent variable itself is volumes. Users of the model wishing to better understand the purpose of the pure price index should review sections 4.2.1 and 4.2.2 of the Forecast Model Report.

Deloitte Access Economics recommend reverting to the previously used pure price index so that the model can assess the 'pure' price response in RUC kilometres that it was designed to capture. Alternatively, a single index for each of Light and Heavy RUC could be taken from the Cost Allocation Model (CAM) to represent each index's timing impacts, assuming the CAM provides sufficient history. Either way, as long as there are infrequent updates to official RUC rates and announcement of those changes in advance, the current lag structure of prices should be maintained.

An enhanced implementation of the Rate Testing tab has also been implemented, expanding its capacity to shock by c/L or % separately, and for each of RUC, PED and MVR. Expanding the tab to allow quarterly shocks could simplify calculations if price increases ever occur outside the July month, though is not a necessity.

3.2 Other matters

3.2.1 Newly added sheets

Forecasts – quarterly: All input data in the spreadsheet are linked correctly.

Graphs data: Fixed.

Fuel prices, columns AM and AN were only taking September values, rather than year average.

Incorrect column references in columns BB:BC

Diff from prev fct: All input data in the spreadsheet are linked correctly

Three Charts tabs by RH: All input data in the spreadsheet are linked correctly

3.2.2 Addressing additional forecast runs

The Ministry has requested advice on the feasibility and/or desirability of conducting 'out of period' forecast runs.

Though it may be desirable from time to time to run an ad-hoc forecast when certain data are released ahead of the 'official' data, DAE advise the Ministry to exercise caution in updating forecasts based on non-official inputs for any formal forecasts. Any update using non-government inputs should not be used for any formal purpose – the burden of that risk then falls entirely on the Ministry. Informal updates should do no more than inform internal debate on rough directions and magnitudes of changes.

Possible sources for 'non-official' information include:

- Asia Pacific Consensus Forecasts – though not formal or covering all required inputs, the high frequency and timeliness make them a useful source

- The RBNZ releases a quarterly Monetary Policy Statement which contains forecasts for some key inputs. Note, we would expect these to broadly align with the views of the Treasury, but need not do so.

A more useful guide could potentially be compiled out of periodic (monthly if available) updates to transport and revenue data.

- For each given series, the span history subject to the same policy and administrative conditions can be used to compile a *typical* year to date profile of collections.

- This then gives a % collected relative to the final outcome.

- Collections/levels in the year to date for each series can then be compared to the average profile and extrapolated to provide an estimate of the current year's final result.

- These final results can be used to override model estimates in the very short run using the *manual forecast adjustments* columns in each model.

Note, by definition, the extrapolated estimates improve in accuracy as the year progresses.

3.3   Structural break tests

To test whether the originally modelled relationships were still appropriate, all models were updated and extended backwards to their original starting point (if necessary) and forward to include all available data. We then examined the residuals (errors) of each model. The residuals of any model should be random, and centred around zero – that is, on average the error should be zero.

If any trend is apparent in the residuals, then it suggests that the model may be consistently under or over-estimating the dependent variable – i.e. there are potentially more or better explanators of the dependent variable that the model is not capturing. Where a model's errors appear random up to a certain point but then appear to display a trend, it may be evidence that a structural break has occurred.

A structural break simply defines any point in a time series where there is an unexpected shift in the long term trend. To the extent that such a break has occurred, and hence the explanatory variables have changed, then a model that does not correct for that break could generate significant forecast errors. In today's world with economic growth still struggling to return to its pre-GFC trend, a structural break is a real possibility.

Deloitte Access Economics identified three models where the residuals or seemingly large changes in the coefficients indicated strain on the model and the possibility of structural breaks. In addition, the Ministry identified both RUC series as deviating from the normal rule of thumb relationships, and asked DAE to specifically explore these models. In addition to the PED model (discussed earlier in section 3.1), the following models were formally tested for containing structural breaks:

- Heavy RUC
- Light RUC
- MR1
- MR13
- LPG

By and large, the results of these explorations show these models have been performing, and adapting, quite well.

Where a structural break was identified, attempts were made to 'correct' for that break by incorporating dummy variables into the estimated equation. The dummy variable takes a value of zero in the period leading up to the break, and one thereafter. Modelling it in this way thus effectively creates two separate models: one prior to the break and one after.
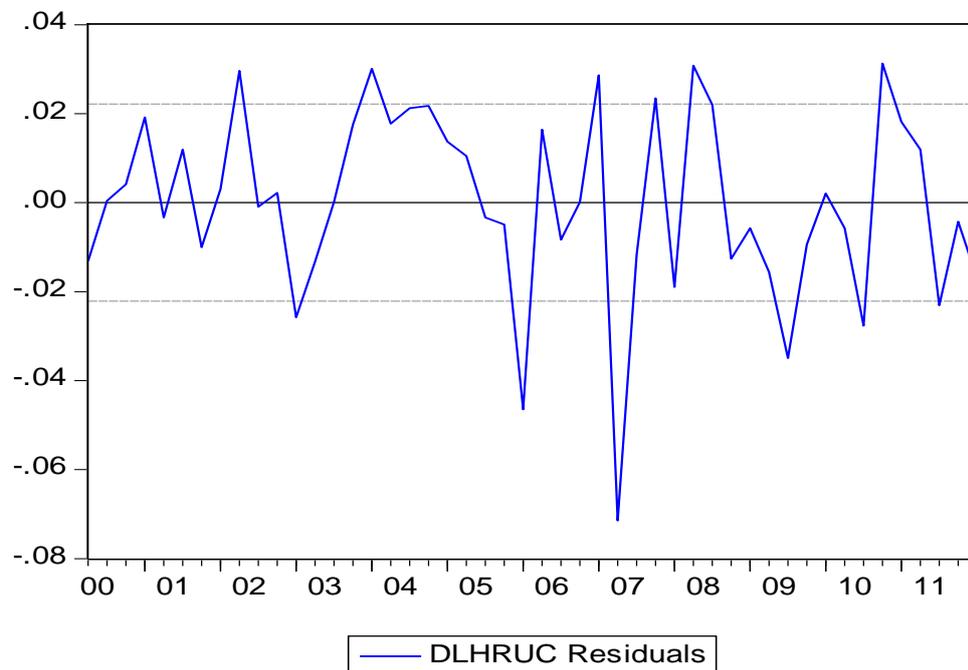
3.3.1   Heavy RUC

The modelling used the Pure Price Index, fixed in the final model, and the actual (unchanged) rates for 2011-12.

The long run trend specification failed the break-point test in 2007q2, though passed in later tests. On examination of the raw data however, large outliers became apparent in 2006 and 2007, which are likely the predominate driver of the apparent structural break. An apparent improvement in the model's performance since early 2009 led us to disregard break test results, and move on to the short run (dynamic) specification.

The dynamic specification failed the break-point test in 2007 and 2009, but passed in 2008. Again, this suggests the result is driven by those same outliers identified earlier.

For completeness, we proceeded to test for dummy variables working with a break at 2008q2. This should avoid any bias driven by the outliers, though we note the model passed the break test at this point.

The specification contains only two independent variables – GDP and the real Heavy RUC price. Two models were thus tested, one with a dummy variable on GDP, the other with dummy variables on price.

The updated GDP model improved errors in the last four quarters at the expense of the previous eight quarters. The updated price model did note appreciably improve recent errors in the model, though overall model fit did improve slightly. However, the estimated coefficients failed to produce a sensible economic interpretation.

Deloitte Access Economics recommend the model be left in its original form, due to the combination of unclear economic implications from the updated price model and negligible overall model improvement. In other words, we have left the model coefficients unchanged.

3.3.2 Light RUC

There was no evidence of a break in the long run model specification, either through changes in coefficients or examination of the residuals.

While not clearly a break, examination of the residuals in the dynamic specification suggested break points might exist in 2008Q2, 2009 Q2 and 2009Q4.
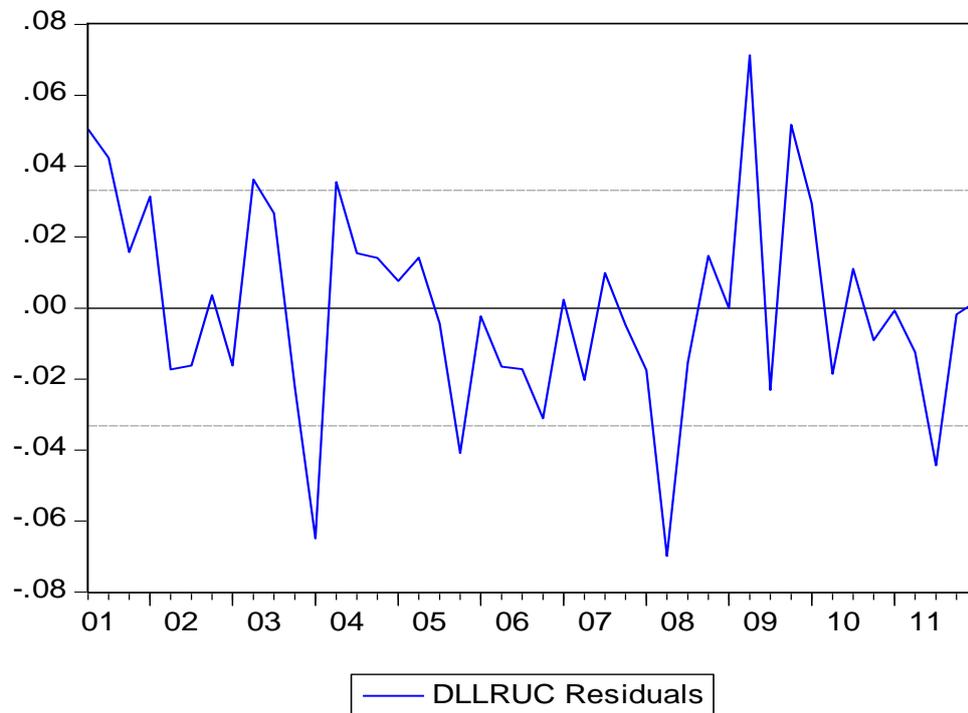
The model passed the break point test in 2009Q2, and 2009Q4 left the latter part of the equation too short to test. However, the model failed the break test in 2008. As with earlier examples, our *a priori* expectation was that this was a matter of outliers rather than real trend changes.

Nonetheless, attempts were made to improve the dynamic model's performance by imposing a dummy variable on both GNE and VPERC (vehicles per capita). The dummy variable is defined as being equal to one in the pre break period and zero thereafter. In effect therefore, the dummy variable takes the value of the base series in the period prior to the structural break.

Both dummy variables failed to improve the model appreciably, though this was not surprising given the already high adjusted-R^2, or to significantly alter the coefficients of existing variables. On these grounds, Deloitte Access Economics recommend no changes to the model.

3.3.3   MR1

Examination of MR1 residuals showed a gap had opened up in the long-term trend with actual values exceeding the fitted values of the past few years. Due to the construction of the error correction variable, the model was trying to drive down forecasts to restore the trend.
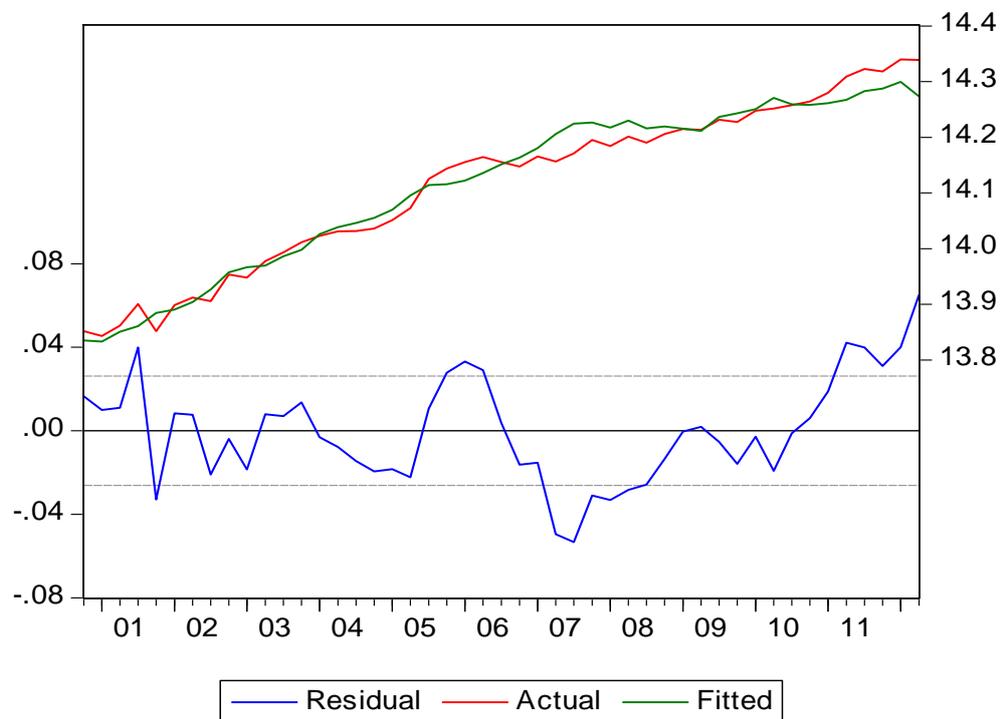
Break point tests revealed a structural break in 2007Q2, but the test statistics continued increasing until 2009Q4, suggesting that this is where the break 'peaked'. Further modelling was therefore undertaken from this point.

Again, the long run model was a simple relationship between only three variables: GDP and the passenger and commercial vehicle stocks. Dummy variables were imposed on all three explanatory variables but in all cases, the model failed to improve appreciably.

Despite the apparent break in the long run trend, it is worth noting that the dynamic model has actually adapted quite well, with positive and negative errors roughly offsetting each other since the long term break appeared.
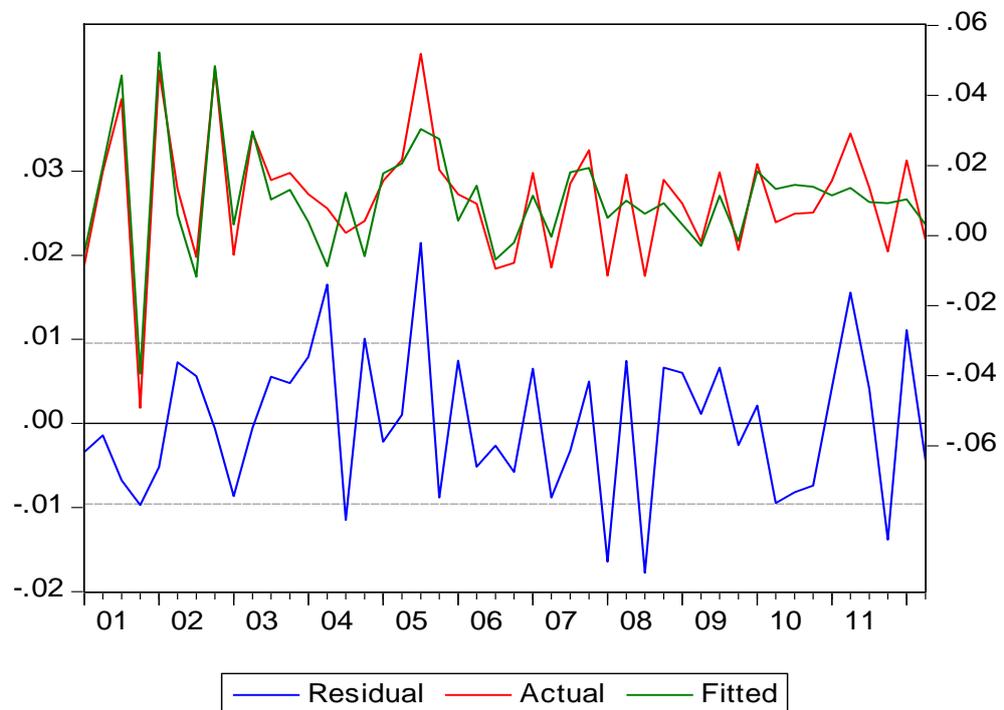
For completeness, we proceeded to test the dynamic model based on the 'corrected' long run models to see if they improved the short run fit of the model. However, since none of these specifications were able to improve the dynamic model, Deloitte Access Economics recommend leaving the models in their current form.

3.3.4   MR13

Examination of the residuals revealed possible structural breaks around 2008 and 2009. Further formal testing revealed a break in 2009 but was inconclusive about 2008. This deviation from trend has the potential to affect the error correction component in the dynamic specification: essentially, the model thinks the estimated value is too far away from trend, and so over or under corrects the estimated value.

Chart 3.6: Actual, fitted and residuals chart, original LR MR13 specification
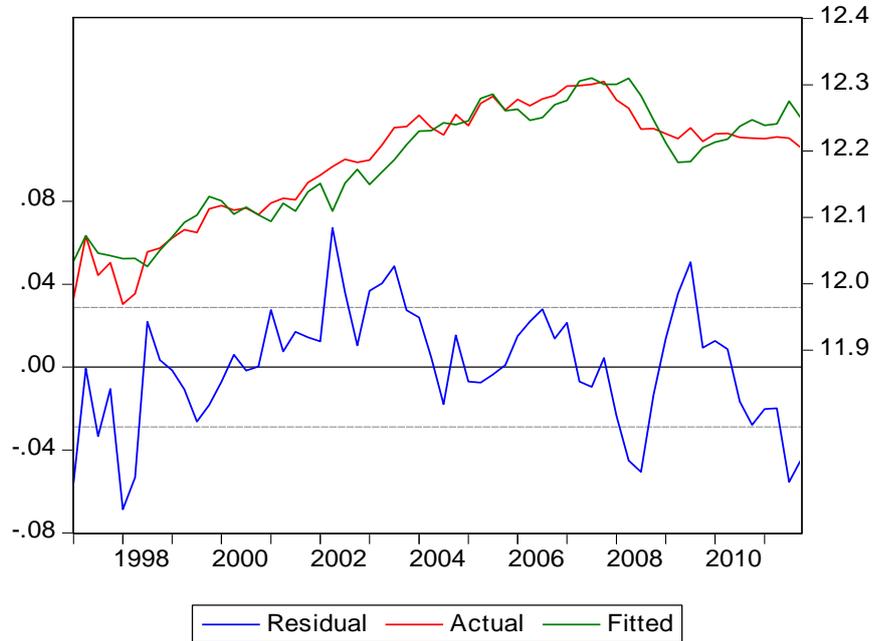
Dummy variables were added individually to both of the explanatory variables in the long run specification, with no noticeable improvement in results. A dummy variable was subsequently added to both of the explanatory variables, and the fitted (green) line shifted noticeably closer to the actual (red) line in the post break period, as shown below.

**Chart 3.7: Actual, fitted and residuals chart, 'corrected' LR MR13 specification**
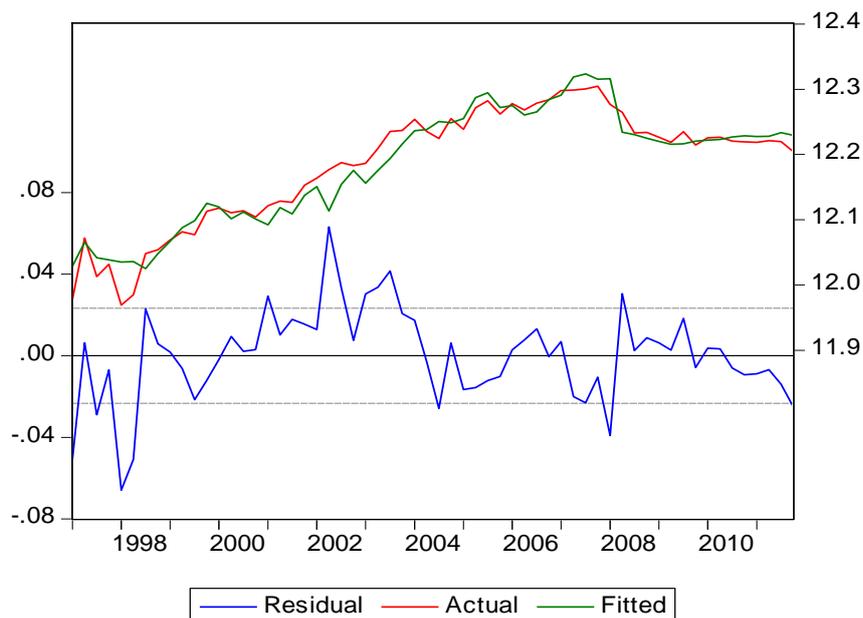
In the very short run this improved model specification may well produce improved results. However, the downside with adopting this 'new' model is that it implicitly assumes that the post break trend is that which will carry into the future. That is, it assumes there has been a sustained behavioural change rather than a one off 'GFC' style change.

Chart 3.8 shows year-to growth in MR13 volumes in the original model specification as well as the 'new' model with dummy variables to account for the structural break. In the early forecast period, both models show a notable upturn as GNE growth picks up from around 2013. Following this, both models' error correction terms 'push' the forecasts down, before stabilising at trend growth around 2015-16.

As shown below, growth in MR13 volumes did indeed drop quite notably in the GFC period. However, the last two quarters of actual data (March and June 2012) have shown year-to growth of roughly 2%, suggesting a recovery may already be in hand. On balance, Deloitte Access Economics believe that the downturn in growth was a GFC inspired outlier.

Chart 3.8: Year-to volume growth, MR13, 'original' vs 'new' model specification
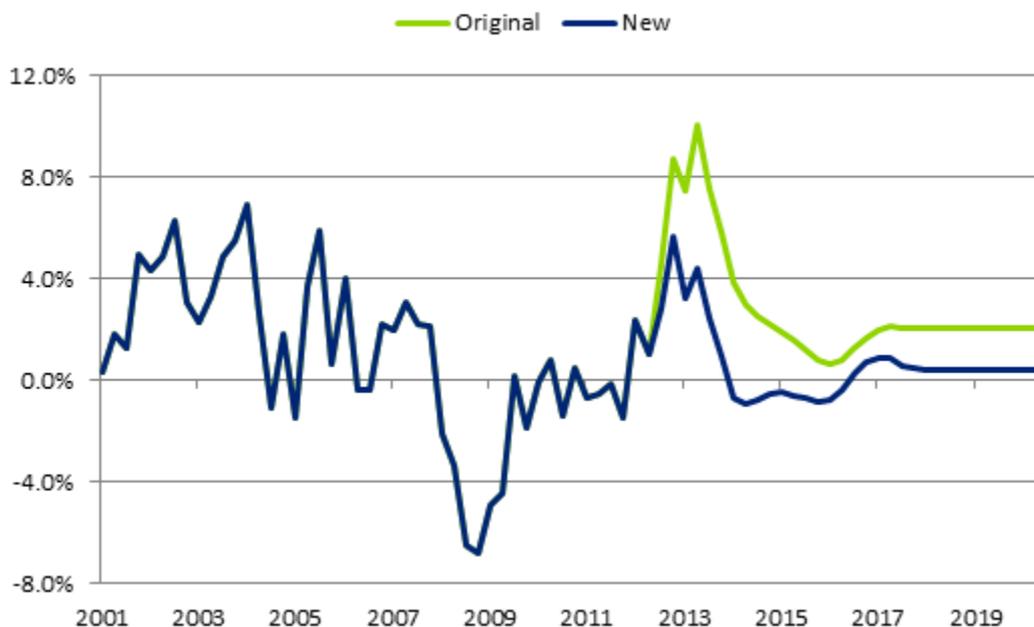
On these grounds, it is difficult to justify changing models. Given the high likelihood that the growth downturn in the GFC period was an outlier, it is inappropriate for this to affect the model's estimated long term growth path. Though the original model may have missed the mark in the last few years, its performance in the pre-GFC period was strong, and once the New Zealand economy returns to health there is every reason to believe these historical trends will continue.

Therefore Deloitte Access Economics recommend sticking with the original model. Should the Ministry wish to manually adjust forecasts in the short term, it can do so using the space provided in column AQ.

3.3.5   LPG

Examination of the residuals from the long run trend specification for LPG volumes strongly suggested a structural break. As shown in the chart below, logged LPG revenues (the red line) appear to have reached a turning point in the post GFC period.

The change has clearly been too sudden and too strong for the model coefficients to adapt sufficiently well, with the model's estimates (the green line) showing considerable deviation from the red line since the shift occurred.

There are two broad explanations for this. First, it could be that these changes are largely a behavioural shift and that the old trend is no longer applicable. There may thus be a 'correction' in behaviour before a new trend carries the series forward. Second, the structural break may simply be a one-off, GFC inspired outlier and eventually, the old trend behaviour may return.

**Chart 3.9: Actual, fitted and residuals chart, original LR LPG specification**
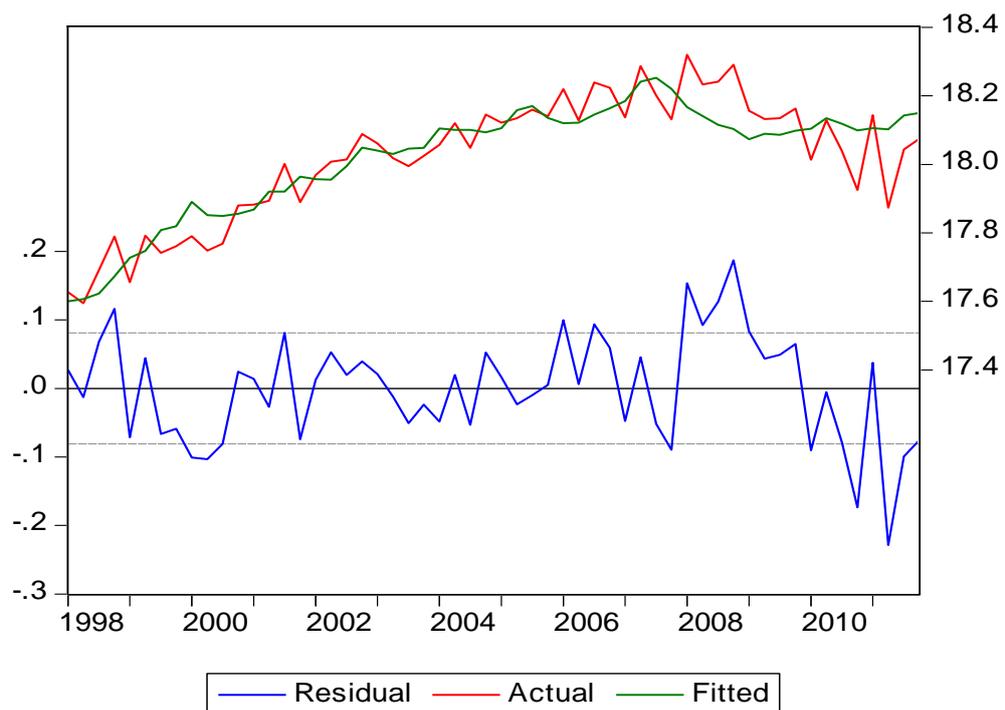
Even with updated coefficients, the long run LPG model misses the mark in recent years. A break point test showed the model broke down, and allowed us to narrow the range over which the break occurs to late 2008. As elsewhere, we tested a year either side of our 'inspected' range, and the central point presented the strongest evidence of a break.

The long run model is driven by GDP, the pump price of petrol, and the working age population. Given that population itself is not subject to rapid change, we focused our attention on the former two variables.

Testing dummies based on either of these series in isolation produced significant model improvement as measured by the standard statistics, but failed to have a sensible economic interpretation. In response, we attempted a model with dummy variables correcting for the structural break in both GDP and the petrol price – something which produced a vastly improved and more sensible model, with the fitted value (green line) successfully tracking the actual value (red line) even after the structural break.

Chart 3.10: Actual, fitted and residuals chart, 'corrected' LR LPG specification
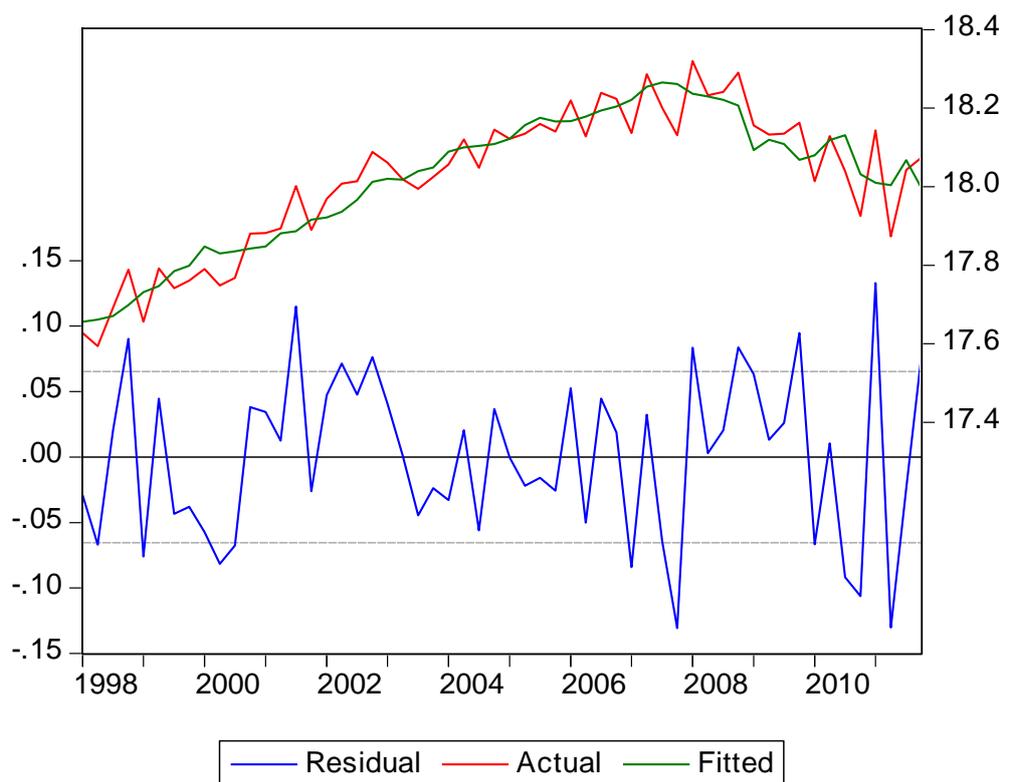
The Ministry has indicated a preference towards adopting a simple trend estimation. In light of the current difficulties with forecasting LPG, such an approach is a sensible alternative. However, when implementing the TREND function in Microsoft Excel, it becomes apparent that the only discernible 'trend' in the post structural break period is a flat line – i.e. the formula results in what is essentially a zero growth assumption.

In a time of rising petrol prices and increasing environmental consciousness of the general public, a zero growth assumption may be too low. As the economy returns to stable growth, we expect that LPG volumes (and revenue) will equally return to growth.

That said, while a structural break has clearly occurred since 2007-08, it is impossible at this stage to ascertain for certain whether this represents a new trend of declining LPG usage or whether it's a one-off behavioural change in response to tough economic times.

Chart 3.11 shows the forecasts generated for LPG volumes from the original model, the corrected model and the TREND formula. In our opinion, the corrected model produces the most sensible results.

On balance, Deloitte Access Economics recommend adopting the 'corrected' model specification. However, given the uncertainties surrounding LPG usage at present, we recommend the Ministry keep a close eye on LPG data as and when they become available in order to ascertain whether the structural break does indeed represent a new trend.

Chart 3.11: Alternative LPG forecasts
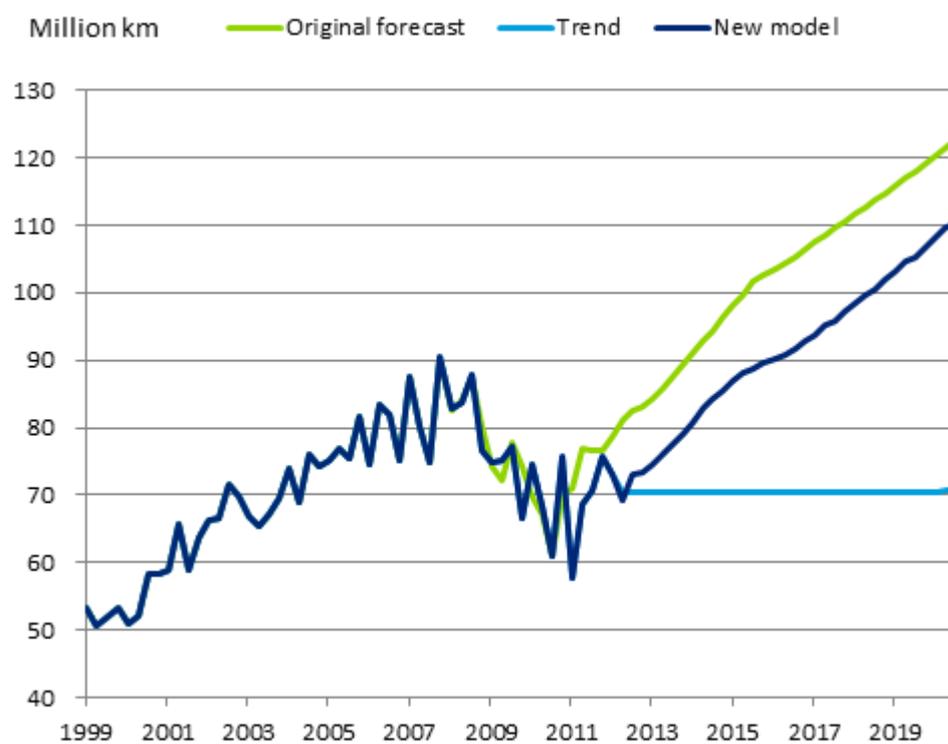
Limitation of our work

General use restriction

This report is prepared solely for the internal use of the Ministry of Transport. This report is not intended to and should not be used or relied upon by anyone else and we accept no duty of care to any other person or entity. The report has been prepared for the purpose of Ministry of Transport. You should not refer to or use our name or the advice for any other purpose.

Deloitte Access Economics is Australia's pre-eminent economics advisory practice and a member of Deloitte's global economics group. The Directors and staff of Access Economics joined Deloitte in early 2011.

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee, and its network of member firms, each of which is a legally separate and independent entity. Please see www.deloitte.com/au/about for a detailed description of the legal structure of Deloitte Touche Tohmatsu Limited and its member firms.

About Deloitte

Deloitte provides audit, tax, consulting, and financial advisory services to public and private clients spanning multiple industries. With a globally connected network of member firms in more than 150 countries, Deloitte brings world-class capabilities and deep local expertise to help clients succeed wherever they operate. Deloitte's approximately 170,000 professionals are committed to becoming the standard of excellence.

About Deloitte Australia

In Australia, the member firm is the Australian partnership of Deloitte Touche Tohmatsu. As one of Australia's leading professional services firms. Deloitte Touche Tohmatsu and its affiliates provide audit, tax, consulting, and financial advisory services through approximately 5,400 people across the country. Focused on the creation of value and growth, and known as an employer of choice for innovative human resources programs, we are dedicated to helping our clients and our people excel. For more information, please visit our web site at www.deloitte.com.au.