OXFORD ECONOMICS

Estimation of households' demand for gas and electricity

December 2008

Oxford Economics
121, St Aldates, Oxford, OX1 1HB
☏: 01865 268900, 🖷: 01865 268906
: www.oxfordeconomics.com

1. Current equations

The current equations for gas and electricity demand by households are reported below.
Gas demand is a function of households income (hh_income), winter degree days (wint_deg_gas), the number of households with gas central heating (households*central_heating and gas prices in real terms (gas_prod_rpu). Oxford Economics recommended trying to include a trend in the equation. Moreover, the current equation does not put the number of households with gas central heating in the long-term part of the ECM, while the variable is non-stationary.

dln(gas_tot_dem) = 2.26- 0.18*(5.05*ln(gas_tot_dem(-1)) + 0.16* ln(hh_income(-1)))
+ 0.29 * ln(wint_deg_gas) + 0.49* ln(households*central_heating/100) - 0.15*ln(gas_prod_rpu)

Electricity demand is a function of similar variables. Here the total number of households is used (households). The income variable, hhe_income, is similar but slightly different to the one used in the gas equation. There is a clear case for harmonisation. Moreover, as is the case for gas, the number of households is not included in the long-run part of the equation.

dln(elec_tot_dem) = -5.59- 0.08*(7.41*ln(elec_tot_dem(-1)) + 0.23* ln(hhe_income(-1)))
+ 0.12 * ln(wint_deg_elec) + 0.99* ln(households) - 0.08*ln(elec_prod_rpu)

The long-run elasticities implied by these equations are reported in the table below. The income elasticity is actually negative, but very small, at 3% in both cases. The elasticity of electricity demand to the number of households seems large at 1.7. The price elasticites are slightly negative, as would be expected.

Elasticities in current equations

	Income	Number of households	Price
Gas demand	-0.03	0.5	-0.2
Electricity demand	-0.03	1.7	-0.1

2. Testing new equations: general form of the equations

For both gas and electricity demand, we tested a long-run relationship between fuel demand on the one hand and

- the number of households
- households real disposable income
- winter degree days
- fuel prices in real terms

and a time trend that is meant to capture possible structural changes such as on the average household size, average house size or quality of insulation.

$$\ln(fuel\_demand) = \alpha_0 + \alpha_1 \ln(nber\_hhlds) + \alpha_2 \ln(hh\_income) + \alpha_3 \ln(w\text{int}\_\text{deg}\_\ fuel)$$
$$+ \alpha_4 \ln(pfuel) + \alpha_5 trend$$

In the data set used, fuel demand is in Muths and prices are in p/utherm. The number of winter degree days used in the 6-month measure. Finally, the income measure is real disposable income in 1995 prices.

We also tested whether the above long-run relationship could be included in an ECM-type equation with short-term dynamics and adjustment to the long-run equilibrium.

Note that we need a constant when estimating the long-run relationships, to take care of the choice of units for the dependent and explanatory variables. This choice is arbitrary (for instance, whether the number of households is in thousands or millions) and affects the constant. Without constant, the other coefficients are not estimated properly.

3. Cointegration tests

Gas and electricity demand have trended upwards over the past 35 years (see chart). The number of households and household incomes are also clearly non-stationary. The number of winter degree days is stationary, while unit root test results for gas and electricity prices are mixed, showing some signs of non-stationarity when looked at since 1970 but with no evidence of non-stationarity when looked at since 1980.

Chart 1 Gas and electricity demand (MUTHs)

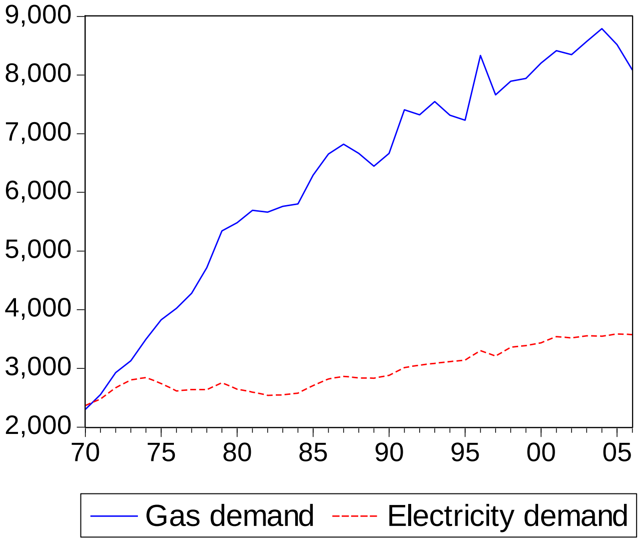
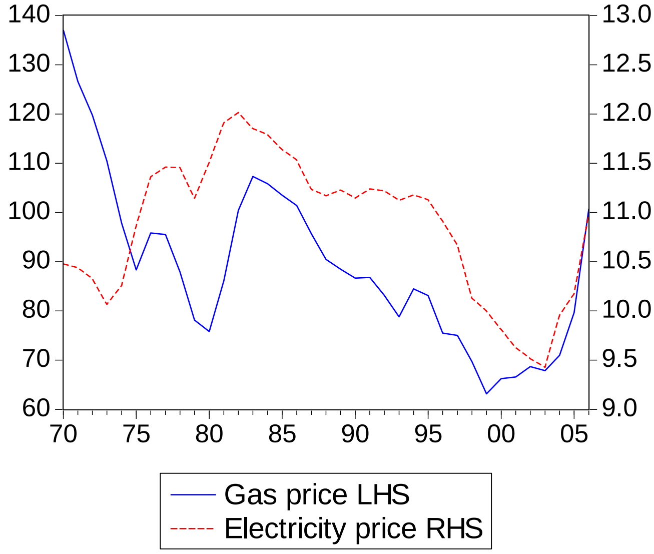


Chart 2 Gas and electricity prices (p/therm 1995£)



Cointegration tests do not provide any satisfactory results. Some evidence of cointegration is found but with a negative coefficient on households income. Given the limited amount of data available, the results of cointegration tests may not be reliable. One possible approach would be to use a first difference specification. However, we found that the explanatory power of such a specification is limited. We would recommend to rely on OLS estimation of a long-run relationship that is then embedded in an error correction form as long as it provides results that make economic sense.

4. OLS estimation

Note that **we recommend estimating the equations since 1980 rather than since 1970 as, especially in the case of electricity demand, the longer estimation period introduces some instability.** For the industry sector, we had also found that estimations from 1980 provided better results. For the last few years of the sample the estimated policy impacts (as supplied by BERR) are added back onto actual demand, so that we can estimate a hypothetical 'pre-policy' level of demand. We tested estimation up to 2003, i.e. before policy impacts are estimated and up to 2006.

We could not find equations of the form shown in section 2 where all the coefficients are significant. Typically, the trend is not significant and either the number of households or households' income is not significant. After testing various alternatives, we recommend the following equations because they are in line with economic theory and pass standard residual and stability tests.

The table shows the long-run elasticities in the equations that we recommend using. Compared with the current equations, there is some rebalancing of the income and number of households elasticities. The income elasticity of gas demand is zero while the elasticity to the number of households is smaller than in the current equation. For electricity, the income elasticity is much larger than in the current equation but the elasticity to the number of households is zero. The price elasticities are similar to those of the current equations.

Elasticities in recommended equations			
	Income	Number of households	Price
Gas demand	0.00	0.5	-0.2
Electricity demand	0.76	0.0	-0.1

4.1. Gas demand

For gas, **we recommend an equation where a long-run relationship is set between gas demand, the number of households with central heating, the number of winter degree days**[1] **and the real price of gas.** We found that we could not have both income and the number of households with central heating, probably because of colinearity between the two variables. Note that we also tested the number of households with central heating together with income per households but the latter was not significant. We recommend using the number of households with central heating rather than households' income, since the variable already captures changes in gas demand accounted for by take up of gas central heating.

We report below the estimation results for the long-run and short-run parts of the equations. Note that the ECM coefficient is very close to 1 and the short-run elasticities are more or less identical to the long-run elasticities. This means that the short-run part of the equation does not add any value. **We therefore recommend using the long-run part of the equation only.**

Dependent Variable: LOG(GAS_TOT_DEM)
Method: Least Squares
Date: 06/12/08 Time: 16:40
Sample: 1980 2006
Included observations: 27

	Coefficient	Std. Error	t-Statistic	Prob.
C	2.65	0.75	3.53	0.00
LOG(HHG_GAS)	0.50	0.03	19.66	0.00
LOG(WINT_DEG)	0.30	0.08	3.66	0.00
LOG(PGAS)	-0.18	0.04	-4.31	0.00

R-squared	0.97	Mean dependent va	8.88
Adjusted R-squared	0.97	S.D. dependent var	0.15
S.E. of regression	0.03	Akaike info criterio	-4.25
Sum squared resid	0.02	Schwarz criterion	-4.06
Log likelihood	61	Hannan-Quinn crite	-4.20
F-statistic	255	Durbin-Watson sta	2.02
Prob(F-statistic)	0.00		

[1] We recomm ... alternatives and found no ... -month series.

Dependent Variable: DLOG(GAS_TOT_DEM)				
Method: Least Squares				
Date: 06/12/08 Time: 16:40				
Sample: 1980 2006				
Included observations: 27				
	Coefficient	Std. Error	t-Statistic	Prob.
C	2.86	0.57	5.02	0.00
DLOG(HHG_GAS)	0.48	0.19	2.47	0.02
DLOG(WINT_DEG)	0.24	0.05	4.44	0.00
DLOG(PGAS)	-0.11	0.07	-1.58	0.13
ECM(-1)	-1.08	0.22	-5.00	0.00
R-squared	0.72	Mean dependent va		0.02
Adjusted R-squared	0.67	S.D. dependent var		0.05
S.E. of regression	0.03	Akaike info criterior		-4.28
Sum squared resid	0.02	Schwarz criterion		-4.04
Log likelihood	62.73	Hannan-Quinn crite		-4.20
F-statistic	14.22	Durbin-Watson stat		1.95
Prob(F-statistic)	0.00			

The residuals show no sign of autocorrelation, no sign of heteroskedasticity. The equation also passes stability tests (CUSUM and recursive coefficient estimates). All the test results are reported in Appendix.

4.2. Electricity demand

For electricity, **we recommend an equation where a long-run relationship is set between electricity demand, real households' income, the number of winter degree days and the real electricity price. The short-term dynamics include the same variables.** As in the case of gas, we found that we could not have both income and the number of households or the number of households together with income per households. From an economic theory point of view, it seems less than ideal not to have the number of households in the equation. Indeed, the fact that the average household size has diminished (and hence population changes which are captured in total income changes have been smaller than changes in the number of households) should have led to a rise in the average electricity bill per person. However, since no satisfactory results can be found with both variables in the equation, **we recommend using households' income because it produces more stable estimation results.**

The electricity equation explains around 80% of the variance of electricity demand. The residuals show no sign of autocorrelation, no sign of heteroskedasticity. The equation also passes stability tests (CUSUM and recursive coefficient estimates). All the test results are reported in Appendix.

Dependent Variable: LOG(ELEC_TOT_DEM)
Method: Least Squares
Date: 06/12/08 Time: 16:40
Sample: 1980 2006
Included observations: 27

	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.66	0.60	-1.11	0.28
LOG(HH_INCOME)	0.76	0.03	22.43	0.00
LOG(WINT_DEG)	0.20	0.04	5.05	0.00
LOG(PELEC)	-0.06	0.03	-1.78	0.09

R-squared	0.99	Mean dependent var	8.03
Adjusted R-squared	0.99	S.D. dependent var	0.12
S.E. of regression	0.01	Akaike info criterion	-5.68
Sum squared resid	0.00	Schwarz criterion	-5.49
Log likelihood	80.68	Hannan-Quinn criter.	-5.62
F-statistic	722.54	Durbin-Watson stat	1.72
Prob(F-statistic)	0.00		

Dependent Variable: DLOG(ELEC_TOT_DEM)
Method: Least Squares
Date: 06/12/08 Time: 16:40
Sample: 1980 2006
Included observations: 27

	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.52	0.11	-4.86	0.00
DLOG(HH_INCOME)	0.28	0.17	1.64	0.12
DLOG(WINT_DEG)	0.17	0.02	7.42	0.00
DLOG(PELEC)	-0.11	0.04	-2.63	0.02
ECM(-1)	-0.80	0.16	-4.94	0.00

R-squared	0.81	Mean dependent var	0.01
Adjusted R-squared	0.78	S.D. dependent var	0.02
S.E. of regression	0.01	Akaike info criterion	-5.99
Sum squared resid	0.00	Schwarz criterion	-5.75
Log likelihood	85.91	Hannan-Quinn criter.	-5.92
F-statistic	23.93	Durbin-Watson stat	1.58
Prob(F-statistic)	0.00		

5. Comparison of forecasts with new and old equations

Charts 3 and 4 compare the forecasts with the recommended and current equations. For both electricity and gas, the current equations produce higher forecasts.[2] For electricity, the main difference between the two forecasts stems from the fact that the current equation includes a large elasticity to the number of households while in the proposed equation the elasticity to households is zero. This is partly offset by a higher income elasticity in the latter equation but not entirely. For gas, the long-run trends of the two forecasts are very similar. The difference in the forecast levels stems from a short-term gap, due to the fact that the current equation only adjusts gradually to the long-run, while in our re-estimation we found that there was no need or support for such a gradual effect.

In discussions between the CCC, BERR and Oxford Economics, there was a concern that the forecasts produced by the recommended equations may be too optimistic.

[2] The data differ for 2005 and 2006 because in the estimation based on the current equations that was carried out by BERR, the policy impacts have been added to the observed data.

Chart 3 Electricity demand

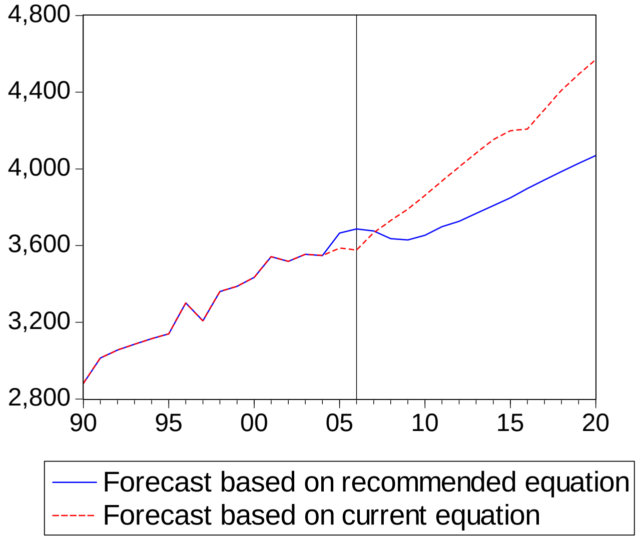
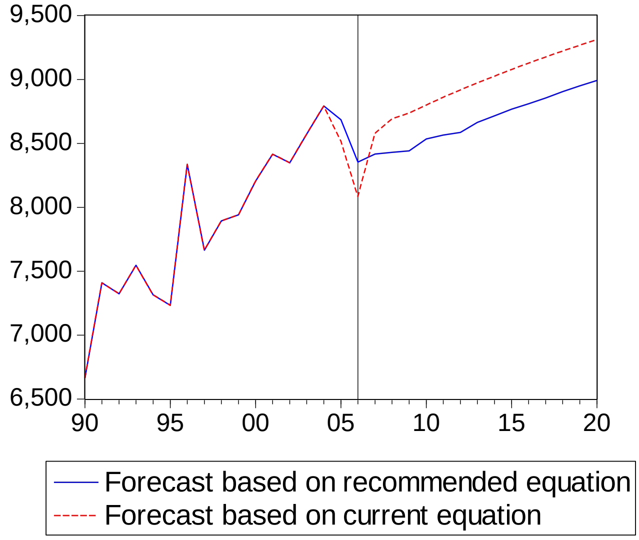


Chart 4 Gas demand



6. Conclusion

Having reviewed the estimation of gas and electricity demand by households, we cannot find any satisfactory specification that includes a trend. Changing the current equations would be an opportunity to correct some flaws. Our analysis suggests that the gas equation would be a long-run equation only, with as explanatory variables, the number of households with central heating, the number of winter degree days and the real price of gas. The electricity equation would be an ECM equation with as explanatory variables, households' income, the number of degree days and the real electricity price. These equations have the advantage over the current equations that they combine homogeneous variables in terms of degree of integration. Moreover, they pass the standard residual tests and are stable. We therefore think that they provide a sound basis for forecasts. This re-estimation would also be the opportunity to harmonise the households' income variable and the number of degree days across the two equations. These are small changes but they remove unnecessary complexity in the current model.

The new equations overcome some of the problems with the existing ones, but are still not perfect themselves in terms of either the econometric specification (requiring use of cointegrating regressions without evidence of cointegration) or the economic interpretation (e.g. no income effect on gas use, and no household numbers impact on electricity). This is largely a function of the limited data (only 27 data points are used in the preferred specifications) and the probable changing form of the true relationships. Whilst we believe the new equations offer an improvement over the existing ones, and their stability in particular is encouraging, there can be no guarantee that they actually will be closer to the true data generating process going forward. As such there is inevitably judgement involved in the choice of model.

Following discussions with BERR and the CCC it was judged preferable to keep the current equations, at least in the short term. The larger demand increases in future implied by the existing equations were considered to be more in line with expectations for domestic energy demand in the absence of further policy and before application of a carbon price to generation. Furthermore, the existing estimates offer the advantage of being consistent with existing policy impact estimates. Without the opportunity to review and re-estimate policy impacts so that they are consistent with the new equations, there would a risk to bias the forecasts, and underestimate future demand. Given the difficulty in finding an ideal equation and the clear uncertainty over the true data generating process retaining the existing equations is deemed preferable. The implication of our analysis and of the results of the new equations is that retaining the current equations may produce an overly cautious forecast (i.e. higher energy demand), particularly while household numbers are forecast to grow more quickly in the future than historically. We therefore recommend that BERR continue to monitor this area closely, taking particular care in the setting of residuals for the domestic sector, and look to update the equations as new data becomes available.

Appendix: test results

Gas demand

1. Autocorrelation in residuals

There is no evidence of autocorrelation up to order 6.

Date: 06/11/08 Time: 16:04
Sample: 1970 2006
Included observations: 37

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. |*. |	. |*. |	1	0.083	0.083	0.2784	0.598
.*| . |	.*| . |	2	-0.166	-0.174	1.4189	0.492
.*| . |	.*| . |	3	-0.132	-0.105	2.1564	0.541
.*| . |	.*| . |	4	-0.096	-0.11	2.5632	0.633
. | . |	. | . |	5	-0.023	-0.051	2.588	0.763
. |*. |	. |*. |	6	0.193	0.159	4.3279	0.632

2. Heterosckedasticity

The null hypothesis of homoskedasticity is not rejected.

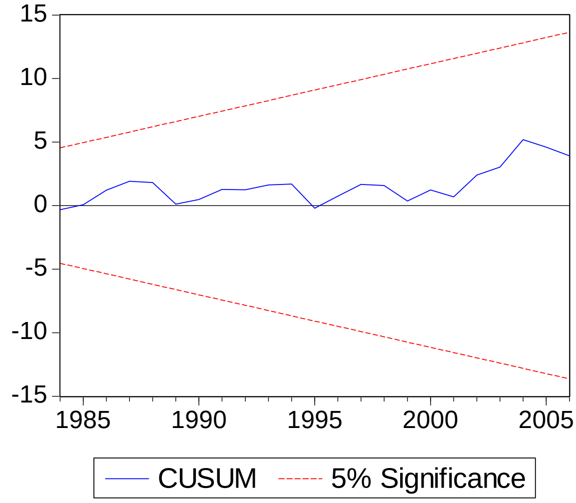
Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	0.219	Prob. F(3,33)	0.883
Obs*R-squared	0.722	Prob. Chi-Square(3)	0.868
Scaled explained SS	0.416	Prob. Chi-Square(3)	0.937

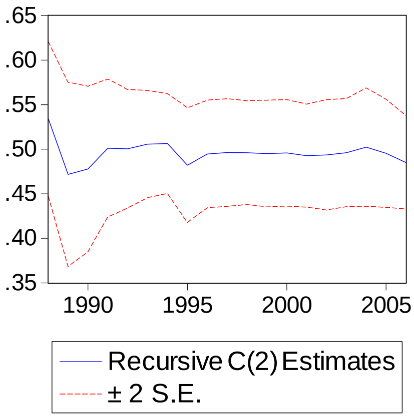
3. Stability tests: CUSUM

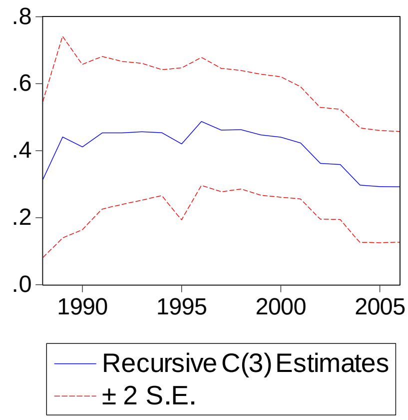
We show the results of stability tests on the long-run part of the equation, since this is the most significant.

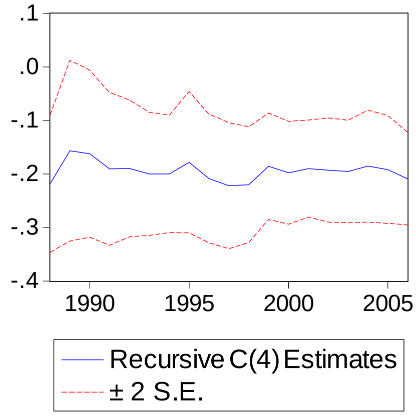
The CUSUM test show no signs of instability.



Recursive coefficient estimates also show that the equation is stable.







Electricity demand

1. <u>Autocorrelation in residuals</u>

There is no evidence of autocorrelation up to order 6.

Date: 06/11/08 Time: 09:40
Sample: 1980 2006
Included observations: 27

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
.**| . |	.**| .	1	-0.262	-0.262	2.0629	0.151
. | . |	. *| . |	2	-0.024	-0.099	2.0812	0.353
. |* . |	. | . |	3	0.091	0.063	2.3524	0.503
. *| . |	. *| . |	4	-0.11	-0.076	2.7673	0.597
. | . |	. | . |	5	-0.009	-0.056	2.7704	0.735
.**| . |	.**| .	6	-0.253	-0.313	5.1531	0.524

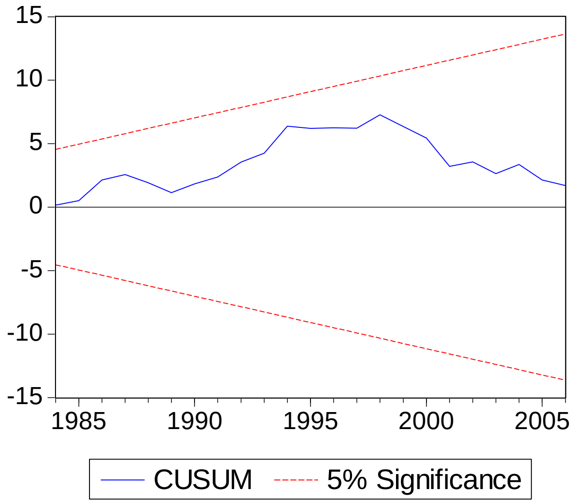
2. <u>Heterosckedasticity</u>

The null hypothesis of homoskedasticity is not rejected.

Heteroskedasticity Test: Breusch-Pagan-Godfrey			
F-statistic	0.339	Prob. F(4,22)	0.849
Obs*R-squared	1.568	Prob. Chi-Square(4)	0.815
Scaled explained SS	0.492	Prob. Chi-Square(4)	0.974

3. <u>Stability tests: CUSUM</u>

The CUSUM test show no signs of instability.



Recursive coefficient estimates are broadly stable. The elasticity to the number of winter degree (c(3)) days has decreased in recent years while the price elasticity (c(4)) has become less negative. We think that these variations are acceptable. We checked other stability tests (recursive residuals, one-step ahead forecast tests) and they all provide satisfactory results.