Has the accuracy of energy demand projections in the OECD countries improved since the 1970s?

Jan Bentzen (jb@asb.dk)
Hans Linderoth (lin@asb.dk)

Department of Economics
The Aarhus School of Business
Fuglesangs Allé 20
8210 Aarhus V
Denmark

Abstract

Since the 1970s almost all OECD countries have published projections or forecasts of future energy consumption. By now, three decades later, the actual values of energy consumption are available for the same number of countries and thus a considerable amount of empirical data is available concerning formal hypothesis testing, e.g. whether there have been improvements in the forecasting accuracy during this period. Using data for sixteen OECD countries the empirical evidence weakly favors the hypothesis that these countries have made some advances in forecasting accuracy concerning projections at the aggregate level of energy consumption and to a lesser degree at sectoral levels. Also, in accordance with a priori expectations, the forecasting failure is increasing with the length of the forecasting horizon.

JEL Codes: Q48, C53.

Keywords: Energy demand projections; OECD countries; Forecasting accuracy.

1. Introduction

Since the oil price shocks of the 1970s much effort has been devoted to the task of forecasting energy prices and energy consumption in the industrialized countries in order to better plan and influence on the use of this important resource. Today, a couple of decades later, a considerable amount of empirical data concerning energy projections is available for formal testing of whether official energy forecasts have been reasonably accurate and also whether forecasting accuracy has improved over time. Based on the official figures from various issues of the annual publication *Energy Policies and Programmes of IEA-countries* from the International Energy Agency (IEA) these issues are analysed. Forecasts that appear in the IEA-publications are submitted by the member countries and are based on official energy plans which to a certain degree express the desired targets for energy consumption in the respective countries. It is demonstrated that in some specific areas the forecasting accuracy has improved when both the length of the forecasting horizon and the periods characterized by oil price shocks are properly included in the test methodology. The present analysis deals with energy consumption projections and does not include formal tests of energy price forecasting precision.

The analysis concerns the OECD energy forecasts from sixteen countries covering forecasts for the years 1985, 1990 and 1995 and with varying forecast horizons. The hypothesis to test is whether more recent forecasts are more precise than earlier ones, i.e. comparing the results for the three forecast periods (1985, 1990 and 1995), but also to what degree the length of the forecasting horizon influences on the ability to project energy consumption. The analysis will take place at both the aggregate level, defined as the total primary energy supply in the OECD statistics, and at sectoral levels, including industry, transport and “other sectors”. Also, the forecasting performances of the individual countries will be tested concerning the beforementioned variables, i.e. length of forecasting horizon and time of forecast.

2. Data sources and summary statistics

The data set covers these sixteen OECD countries: Austria, Belgium, Canada, Denmark, Germany (West), Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Spain, Sweden, Switzerland, United Kingdom and the USA. The energy projections refer to the years 1985, 1990 and 1995 and in all cases the forecast horizon varies between 2 and 7 years which consequently will be the lengths of the forecasting horizons dealt with in the statistical analysis. This means that the projections for 1985 are based on the review years 1978-1983, the 1990-projections have review years 1983-1988 and finally, for the projections for 1995 the review years are 1988-1993. These time periods are the so-called *review years* in the IEA reports, with a lag between the review year and the publication time, where the latter will typically be the next year. Hence, the 1978 review has been published in 1979 and data covering the period up until 1977 may have been used for making the projections.

With sixteen countries and six different forecast horizons for each of the three forecast years the data set comprises 280 observations in total, with only eight cases of missing information. A data set of this size is obtained for five different categories or sectors concerning energy consumption and this makes the total size of the data set 1440 observations.

Having collected the historical energy demand projections these can now be confronted with the actual levels of energy consumption - also derived from the OECD/IEA energy publications - in the respective cases. The forecasting accuracy will be measured as the relative deviation from the actual consumption level, as stated in (1):

$$A_{i,h,t} = \frac{F_{i,h,t}}{C_{i,h,t}} \tag{1}$$

F: forecast of consumption

C: actual consumption

i: 1,2,...,16 (Country)

h: 2,3,...,7 (Forecasting horizon)

t: 1,2,3 (Year of forecast; representing 1985, 1990 and 1995)

This set of forecasting accuracy coefficients will be calculated for the total primary energy supply (TPES), oil consumption, and for three components - the industry sector, the transport sector and “other sectors” (includes the household sector) - constituting total final consumption (TFC). Table 1 presents the unweighted mean values of these coefficients.

The overall conclusion from table 1 seems to be that the precision of energy forecasts have improved over time as the deviations are diminishing from 1985 to 1995, except for the transport sector and “other sectors”. The relatively high values of the standard deviations also suggest that more formal test proving is needed as the shrinking values of the accuracy coefficients may not show up to be significantly valid and additionally, the length of the forecasting horizon should be included in the testing procedure. Finally, table 1 shows that especially for oil consumption and the transport sector - also an oil consuming activity - there is a tendency to underestimate the future levels of energy consumption, which is still a fact for many industrialized countries with persistent increases in economic activity in the transport sector. Energy savings or increasing energy efficiencies have proved difficult to implement with regard to this sector and there seems to be a general tendency to underestimate future energy needs for transport activities. This may also be caused by political considerations as strong increases in oil consumption may be in contradiction with the official energy policies and therefore the oil projections are biased downwards.

Table 1. Mean values of the forecasting accuracy coefficients.

	1985		1990		1995	
TPES	1.12	(*0.14*)	1.06	(*0.12*)	0.99	(*0.09*)
OIL	1.27	(*0.28*)	0.98	(*0.09*)	0.98	(*0.14*)
TFC:						
- Industry	1.27	(*0.24*)	1.05	(*0.13*)	1.12	(*0.11*)
- Transport	1.03	(*0.14*)	0.89	(*0.10*)	0.91	(*0.08*)
- Other sectors	1.08	(*0.15*)	1.01	(*0.13*)	0.94	(*0.12*)

Notes: The coefficients are calculated from (1) including all sixteen OECD countries with no distinction made for the length of the forecasting horizon. TPES is total primary energy supply and TFC is total final consumption - both as used in the OECD energy balances - and final consumption is represented by the three sectors mentioned in the table. Standard error in parenthesis and the number of observations varies between 89 and 96 for the respective cases.

Source: IEA: Energy Policy and Programmes of IEA Countries, var. issues, OECD/IEA.

3. Energy projection accuracy and empirical testing

Much of the literature related to the topic of forecasting accuracy of energy demand projections is related to specific countries or fuels and usually illuminating the many forecasting errors made in the past, especially in the area of oil prices. For the US economy Shlyakhter et al. (1994) uses 170 past forecasts to develop a method for improved uncertainty estimation in relation to future energy projections. Some of the studies paying attention to some general features and problems related to energy forecasting is Tussing and Harris (1992), Griffin (1993) and O'Dell (1996) discussing both historical facts and reasons concerning forecasting failure. Before the oil price shocks of the 1970s only little attention was paid to the questions of energy supply or energy price development. The following price upheavals in the oil market certainly changed this situation and especially oil prices were in the late 1970s often forecasted to reach future levels of \$50 to \$100 per barrel. History has proven that this was an overreaction by both official and private forecasting professionals. Additionally, the oil price crash in 1986 and the short-term peaks in oil prices at the time of the Gulf War have illustrated that forecasts of energy demand and energy prices ought to include both theoretical arguments and empirical evidence related to long-run considerations. One example of this is the model used by the OPEC Secreatriat, described in Ghanem et al. (1998), where the recent advances from the time series literature concerning cointegration and error-correction modelling have been used in order to make explicit distinctions between short-run and long-run effects.

Energy consumption is usually assumed to be influenced by primarily real energy prices and real income and therefore unexpected shocks to the economic development will also have significant effects on the amounts of energy consumed. If, additionally, the price and income elasticities in energy demand are relatively high this will reinforce the size of the forecasting failure of energy consumption, e.g. as measured in (1), in case of shocks to the projected path of the business cycle. Another fact explaining forecasting failure is the tendency in the official energy projections to only adjust the forecasts with relatively long time spans. When several issues of an IEA review publication report exactly the same number for a given component of energy consumption in a country, it is obvious that the forecast is not annually adjusted or updated. This will also influence the tests done in the present analysis and therefore only subsamples of a reasonable size (minimum 90 observations) are used in order to avoid too much dependency on this phenomenon. A similar problem is related to the fact that some of the energy projections more or less express politically decided targets for future energy consumption levels which further complicates forecast evaluation.

In order to test for the degree of failure in the energy projections it is most reasonable to treat deviations above and below actual consumption values in a symmetric manner, which is stated in the left hand side of (2) used for hypothesis testing:

$$\left| A_{i,h,t} - 1 \right| = \alpha_0 + \alpha_1 D_{78\_80} + \alpha_2 D_{83\_85} + \alpha_3 h + \alpha_4 t + \varepsilon_t \tag{2}$$

D_{78_80}: dummy = 1 for 1978-1980, otherwise zero.

D_{83_85}: dummy = 1 for 1983-1985, otherwise zero.

The subscripts *i, h, t* refer to country, forecast horizon and time trend as defined in (1). The right hand side explanatory variables in (2) include dummy variables to take into account the special circumstances that may have ruled in the periods of sharp oil price changes in the late 1970s and the mid-1980s. Failure to correct for these events might bias the results concerning improved forecasting accuracy over time. Forecasts for 1995 - and the short-term forecasts for 1990 - are less influenced or “disturbed” by the beforementioned energy prices shocks and thus these forecasts must be expected to perform relatively better - which might be a wrong conclusion as far as the only reason for decreasing forecasting failures is the absence of “shocks” which automatically makes forecasting easier.

As a starting procedure only a common intercept term is included in (2) - alternatively, country specific effects may be taken into account by expanding the intercept term in this direction - and the explanatory variables are the length of the forecast horizon (h) and the time period forecasted (t). The parameter estimates of horizon (h) and time (t) are expected to show up with positive and negative signs, respectively, assuming a longer forecast horizon to make the projections more imprecise and finally, that the quality of forecasts have improved over time. The test is applied to the five categories from table 1 and the results are exhibited in table 2.

Table 2. Parameter estimates of model (2).

	TPES	OIL	Industry	Transport	Other sectors
Intercept	0.075* (*3.64*)	0.094* (*2.74*)	0.136* (*4.56*)	0.036* (*1.94*)	0.068* (*2.85*)
D_{78_80}	0.056* (*2.48*)	0.209* (*5.53*)	0.171* (*5.24*)	0.017 (*0.85*)	0.066* (*2.55*)
D_{83_85}	-0.023 (*-1.25*)	-0.099* (*-3.26*)	-0.082* (*-3.13*)	0.041* (*2.50*)	-0.027 (*-1.30*)
Horizon (h)	0.014* (*3.23*)	0.025* (*3.36*)	0.021* (*3.30*)	0.011* (*2.80*)	0.007 (*1.35*)
Time (t)	-0.022* (*-2.54*)	-0.037* (*-2.55*)	-0.034* (*-2.73*)	0.006 (*0.77*)	-0.003 (*-0.25*)
$\bar{R}^2$	0.21	0.41	0.39	0.12	0.09

Notes: The number of observations is 280 in all five cases. In parenthesis *t*-values for the parameter estimates, where * indicates an estimate significantly deviating from zero at least at the 5 per cent level of significance.

The simple model (2) seems to perform well in three of the five cases as the degree of explanation is relatively high and a majority of the parameter estimates significantly deviating from zero. The parameter estimates for the length of the forecasting horizon are positive and significant at all conventional levels in four of the cases, indicating that a longer time horizon is also associated with greater forecasting failure. For the time dimension, the trend (t) parameter is significantly negative in three of the cases - also in accordance with a priori expectations - showing that the precision of forecasts are likely to have improved during the analysed period. The transport sector and “other sectors” both seem difficult to forecast with an insignificant time parameter estimate. The first dummy variable for 1978-1980 is positive and significant in four cases indicating increased forecasting failure in the projections made in this period. This may be explained by the fact that when forecasts tend to overshoot future consumption levels unexpected energy price increases will induce increased forecasting failures as energy consumption levels are influenced downwards compared to the forecasts. Likewise, as oil prices collapsed in 1985-86, the second dummy variable shows up with a negative sign in most cases - although only significant in two of these cases - indicating a diminishing forecasting failure, which is probably caused by lower energy prices expanding energy consumption and thereby reducing the difference between actual and projected levels of energy consumption. The opposite effect holds for the transport sector where the dummy variable is found to be significantly positive which can be explained by the tendency to “undershoot” future energy consumption levels in this sector.

In order to further analyse the transport sector and “other sectors” as the quality of the energy projections did not seem to improve over time, a fixed effects model based on (2) with country specific intercept terms is applied in these two cases. Still, a common parameter is assumed for both the horizon (h) and the trend (t) dimensions of the model, with table 3 reporting the results.

Table 3. Fixed effects version of model (2); parameter estimates for the transport sector and "other sectors".

	Transport		Other sectors	
D_{78_80}	0.018	(*0.94*)	0.067*	(*2.97*)
D_{83_85}	0.041*	(*2.67*)	-0.026	(-*1.44*)
Horizon (h)	0.011*	(*3.01*)	0.007	(*1.51*)
Time (t)	0.007	(*0.93*)	-0.003	(-*0.31*)
$\bar{R}^2$	0.24		0.31	

Notes: Sixteen country specific intercept terms are included in the model and they vary between 0.035 and 0.250 - most of them significantly deviating from zero. In parenthesis *t*-values for the parameter estimates, where * indicates an estimate significantly deviating from zero at least at the 5 per cent level of significance.

With country specific intercept terms and assuming identical dummy and slope parameters in the model the degree of explanation ($\bar{R}^2$) has increased remarkably compared with the similar results from table 2, but the parameter estimates are nearly identical with the former results. Therefore, it may be concluded, that for the transport sector and "other sectors" the forecasting failure does not seem to diminish when comparing the energy consumption projections from 1985, 1990 and 1995. The fixed effects model may be further expanded to include country-specific parameters for the horizon (h) variable and the time (t) variable, but such a procedure was not found to influence the conclusions already obtained from the results reported in table 3.

4. Country-specific performances in energy forecasting

The conclusion from part 3 seems to be that the forecasting failure is smaller when analysing primary energy consumption categories (e.g. TPES) compared with final energy consumption variables as represented by the three sectors included in the analysis. If total energy consumption forecasts are based on sectoral projections, this result may be explained by the fact that forecasting failures from sector-specific levels cancel out at the aggregate level and hence make TPES or total oil consumption forecasts appear relatively better. But, if it is assumed that the five different categories of energy forecasts (i.e. TPES, oil and the three sectors) in some sense represent independently calculated projections, there will be a reasonable number of forecasts for each of the sixteen countries analysed to allow an investigation of country-specific performances.

With these five categories of energy forecasts, with forecast horizons ranging from 2 to 7 years and a three period time dimension (1985, 1990 and 1995) there will be 90 observations for each of the OECD countries. The results form applying a fixed effects model with sector-specific intercept terms are reported in table 4.

Table 4. Estimates of country-specific performances.

Parameter:	D_{78_80}	D_{83_85}	Horizon (h)	Time (t)
Austria	0.082*	-0.060**	0.014**	-0.009
Belgium	0.217*	-0.013	0.013	-0.024
Canada	-0.050	-0.033	0.021*	-0.011
Denmark	0.038	0.063*	0.005	-0.002
Germany (West)	0.044	-0.102**	0.031*	0.005
Ireland	0.269*	-0.137*	0.028*	-0.056*
Italy	0.142*	-0.048	0.015*	-0.029*
Japan	0.189*	-0.095*	0.028*	-0.036*
The Netherlands	0.245*	-0.072	0.030*	-0.027
New Zealand	-0.059	-0.086*	0.016**	0.042*
Norway	0.129*	-0.016	0.006	0.001
Spain	0.200*	-0.024	0.008	-0.029*
Sweden	-0.011	0.000	0.013**	-0.048*
Switzerland	-0.082*	-0.012	0.009*	-0.028*
United Kingdom	0.119*	-0.014	0.014	-0.006
USA	0.065*	0.000	0.012*	-0.012

Notes: In parenthesis *t*-values for the parameter estimates, where * indicates an estimate significantly deviating from zero at least at the 5 per cent level of significance and similarly, ** indicating significance at the 10 per cent level. The number of observations is 90 in each case, except Germany (West) with missing values for 1995. Sector-specific intercept terms are included in the regression, but estimates are not reported as these are of minor relevance to the analysis.

Approximately half of the dummy variable parameter estimates are found to be significant and with positive and negative signs, respectively, with the interpretation already discussed in relation to table 2. In nearly all cases the horizon (h) parameter shows up significantly and positively signed, as expected. The more interesting results concerning the time (t) variable reveal that only six of the sixteen countries have a significant and negatively signed parameter estimate. Hence, the model estimates using country-specific data pick out these six countries as the ones with some improvements in forecasting accuracy over time. For New Zealand the implausible parameter estimate of the time variable is probably due to printing errors in the review reports from the late 1980s where some of the forecasted variables suddenly appear with values approximately halved compared with the information in earlier review reports.

Obviously, countries with significant and negatively signed trend parameter estimates are not necessarily the ones with the best overall forecasting performance as the time variable in the test only deals with advances in forecasting precision. For example, a country with zero forecasting failure for all three time periods (1985, 1990 and 1995) would show up with a zero parameter estimate for the trend variable.

In order to further investigate this part of the forecasting performances the *root mean square error* (RMSE) - i.e. the square root of the average of the squared values of the forecast errors -

has been calculated for each of the OECD countries covering the three periods of forecasting, see table 5.

Table 5. Root mean square errors: 1985, 1990 and 1995.

	1985	1990	1995
Austria	0.206	0.087	0.125
Belgium	0.314	0.123	0.135
Canada	0.183	0.133	0.116
Denmark	0.153	0.190	0.120
Germany (West)	0.206	0.120	•
Ireland	0.485	0.098	0.156
Italy	0.255	0.085	0.075
Japan	0.365	0.122	0.137
The Netherlands	0.423	0.109	0.135
New Zealand	0.137	0.192	(0.268)
Norway	0.241	0.157	0.151
Spain	0.284	0.097	0.077
Sweden	0.223	0.140	0.083
Switzerland	0.073	0.115	0.063
United Kingdom	0.261	0.139	0.101
USA	0.171	0.097	0.117

Note: For New Zealand the value for 1995 is in parenthesis as there probably - as mentioned in connection with table 4 - are some printing errors in the IEA publications making the calculation invalid.

The RMSE-values have been calculated for each country covering the six forecasting horizons and the five sectors with the forecasting errors measured as the left hand side of (2). The results in table 5 reveal that four of the six countries from table 4 with an improving forecasting performance over time also show up with a relatively low value of the RMSE for 1995 - the exceptions being Ireland and Japan. The USA has a relatively stable and low value of the forecast error measure in all three cases and hence do not show up in table 4 with a significant and negatively signed trend parameter. For the UK the forecasting performance seems to improve relatively much, but when including the dummy variables and forecasting horizon in the test (cf. table 4) the decreasing forecast errors are due to these variables and not caused by forecasting improvements over time as the trend parameter is statistically insignificant.

5. Conclusions

Data for energy forecasts from sixteen OECD countries have been collected from the official International Energy Agency annual review reports concerning the forecasts for 1985, 1990 and 1995. In total, 280 projections were obtained for total primary energy supply (TPES), oil and three components of final energy consumption (industry, transport and “other sectors”), respectively. These projections were compared with the actual consumption levels as also published in the IEA/OECD energy publications. Taking into consideration the special problems concerning making forecasts in periods of extreme energy price fluctuations, in the present case 1979-1981 and 1985-1986, the empirical results favour a hypothesis of increased forecasting accuracy over time. Especially for TPES and oil consumption the parameter representing the time dimension in the testing procedure indicates that more recent forecasts are better that earlier ones. Also, the forecasting horizon was included in the model and in accordance with expectations, the projections exhibit an increasing degree of failure when expanding the length of the time period forecasted. Finally, for six of the analysed sixteen OECD countries the hypothesis of decreasing forecasting failure over time was confirmed.

References

Ghanem, S. and R. Lounnais, D. Ghazemzadeh, G. Brennand (1998): Oil and energy outlook to 2020: implications of the Kyoto Protocol. *Opec Review*, vol. XXII (2).

Griffin, J.M. (1993): Methodological Advances in Energy Modelling: 1970-1990. *The Energy Journal*, vol. 14 (1).

International Energy Agency (IEA): Energy Policy and Programmes of IEA Countries, Review, 1978-1989, var. issues.

International Energy Agency (IEA): Energy Policies of IEA Countries, Review, 1990-1998, var. issues.

O'Dell, S. (1996): Long-term energy demand: Projections of the International Energy Agency. *The Journal of Energy and Development*, vol. 19 (2).

Shlyakhter, A. I., D. M. Kammen, C. L. Broido and R. Wilson (1994): Quantifying the credibility of energy projections from trends in past data. *Energy Policy*, vol. 22 (2).

Tussing, A. R. and M. J. Harris (1992): Energy shocks, surprises and forecasting failures. *Opec Review*, Vol. XVI (Winter1992-supplement).