RICE UNIVERSITY

Demand for Natural Gas in the US Power Generation Sector

Peter Hartley
Kenneth Medlock III
Jennifer Rosthal

James A. Baker III Institute of Public Policy
RICE UNIVERSITY

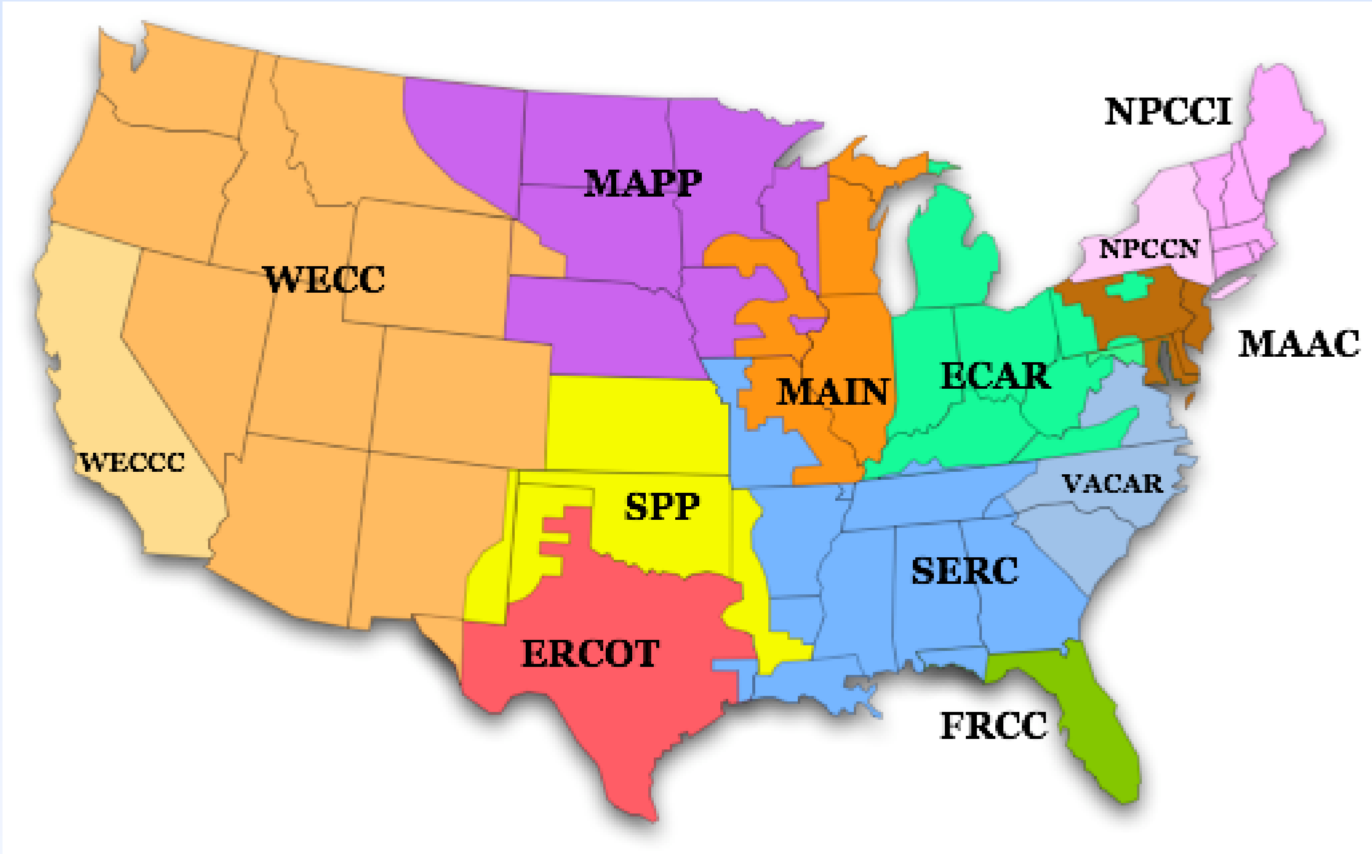
RICE UNIVERSITY

Natural Gas/Oil Product Substitution

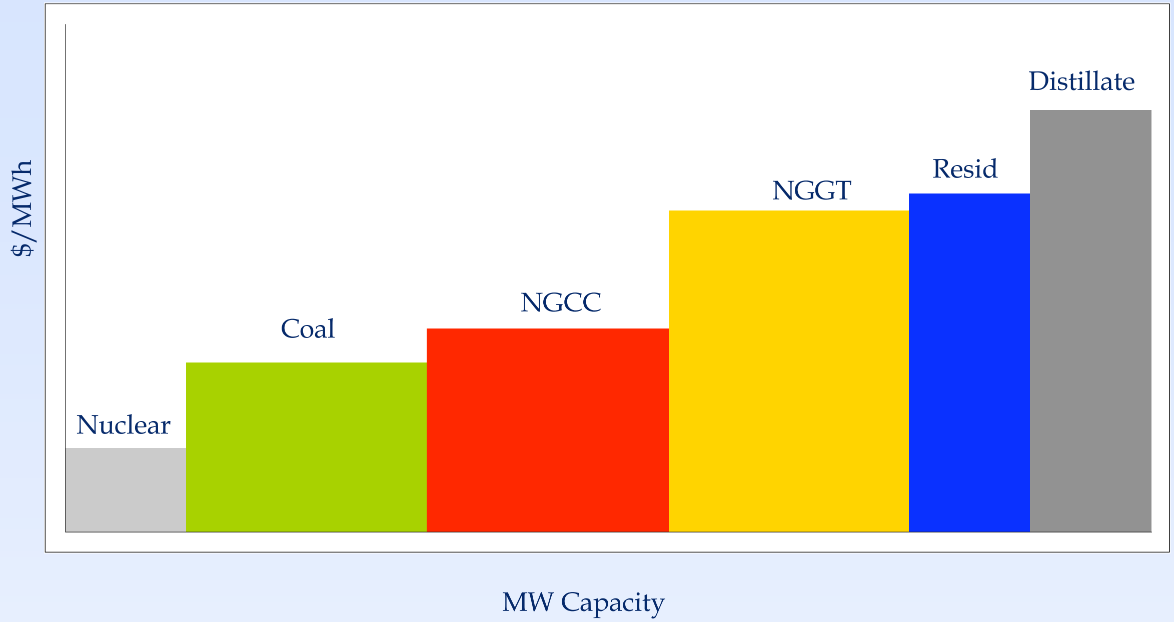
- Aggregate time series analysis supports the notion of a link between natural gas and fuel oil prices forged via electricity sector fuel substitution
- Does a micro-level analysis reinforce that conclusion? Theoretically:
  - Electricity generation is a major user of primary energy
  - There are opportunities for short and long run substitution between fuels
    - Some NERC regions have plants with switching capability
    - Most regions have both natural gas and oil-fired plants that can be run for different periods of time depending on fuel costs
    - In the longer run, technical change has affected heat rates of CCGT encouraging investment in natural gas plants
    - Technology has also reduced non-energy differences between fuels
- To answer these questions empirically we examined monthly data on electricity production, plant capacities and fuel consumption in 13 NERC regions and sub-regions over January 1992–March 2006

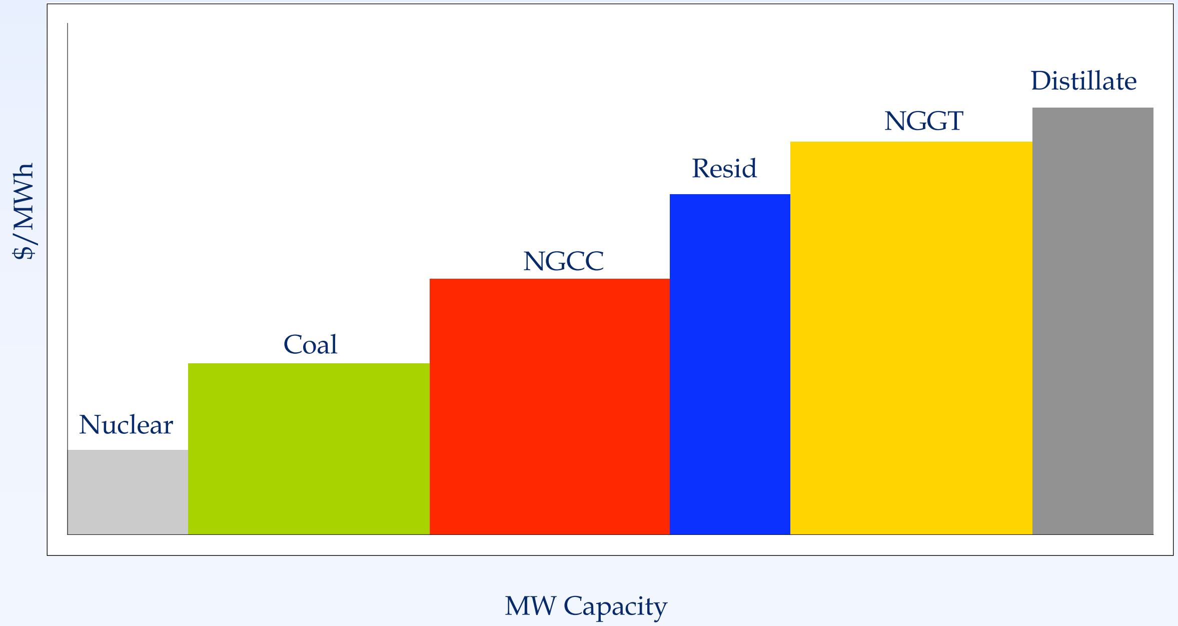
Rice University

North America Electric Reliability Council (NERC) regions prior to January 1, 2006



Key idea 1: Changes in fuel prices alter the stack

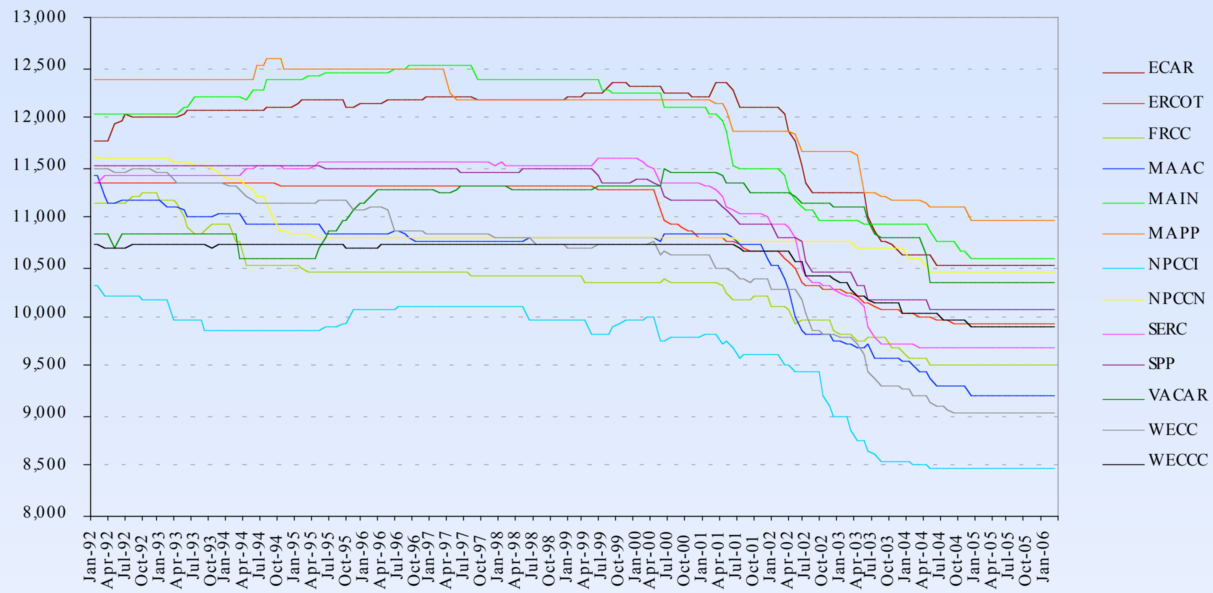




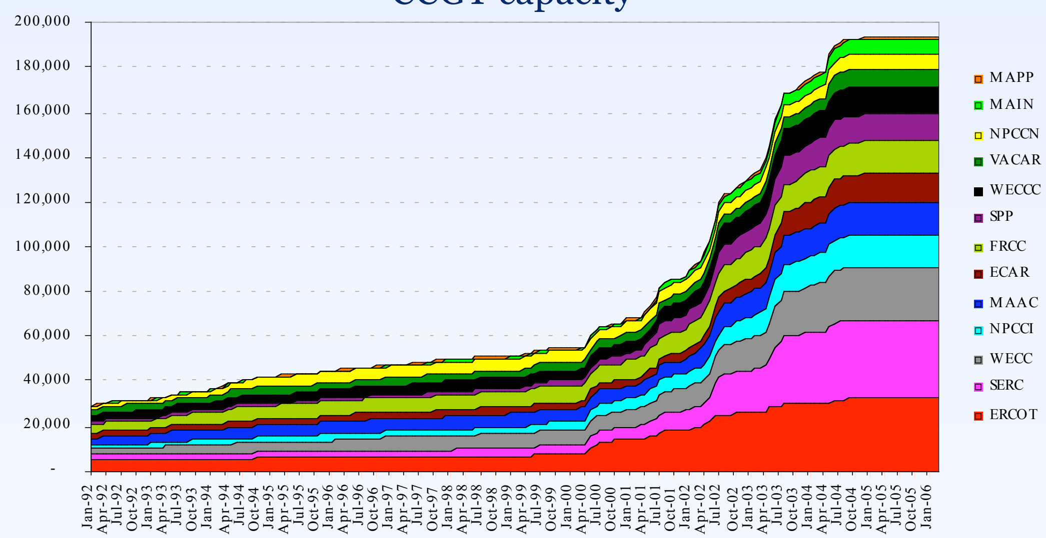
RICE UNIVERSITY

Key idea 2: CCGT have reduced NG heat rates

Capacity weighted average heat rates



CCGT capacity



RICE UNIVERSITY

Calculating costs

- For each NERC region *i* in period *t*, we form a capacity-weighted real cost of natural gas using the average electricity price as deflator

$$NGRCost_{it} = \frac{\sum_{j=1}^{N_i} K_{ij} HR_{ij} P_{ijt}^{NG}}{P_{it}^{E} \sum_{j=1}^{N_i} K_{ij}}$$

- Use city gate natural gas prices in a state for plants located in that state
- Electricity prices are a weighted average of state prices with weights given by the NERC capacity located in each state
- For oil costs, we used PADD prices in place of state-specific prices, which were not always available
- For coal, only national average delivered prices to generators were available
- Plant heat rates were obtained by matching two databases, one from EPA and one from EIA

RICE UNIVERSITY

Trends in real costs

- We first tested stationarity of the real cost variables in each region
- Oil cost was clearly non-stationary in every region
- Natural gas cost was non-stationary in six regions but looked stationary in ERCOT, NPCCI, NPCCN and perhaps WECC, WECCC, MAPP
  - But we find very strong evidence that the errors in the following regressions are stationary in every region

$$\ln NGRCost_{it} = \beta_{i0} + \beta_{i1} \ln OilRCost_{it} + \omega_{it}$$

  - Hence, the real NG cost in each NERC region must contain a non-stationary component that cancels the non-stationary component in oil costs
  - The β coefficients generally differ across NERC regions, reflecting adjustments to regional cost differences
  - The error terms represent deviations from the long run equilibrium cost relativities, which in turn influence short-run fuel choice decisions
- The real coal cost was stationary in all regions except WECC and FRCC, perhaps because of changes in coal plant heat rates in those regions

RICE UNIVERSITY

Stationarity tests

NERC subregion	Test for *NGRCost* nonstationarity[a]	Test for *OilRCost* nonstationarity[a]	Test for *CoalRCost* nonstationarity[a]	β_0	β_1	Test for error nonstationarity[a]
FRCC	0.081	0.708	0.332	0.050	0.896	0.000
VACAR	0.427	0.906	0.0096	-0.075	0.961	0.001
MAAC	0.138	0.885	0.0156	0.346	0.757	0.000
MAIN	0.168	0.859	0.0048	-0.212	0.893	0.000
MAPP	0.093	0.888	0.0138	-0.146	0.864	0.000
NPCCN	0.003	0.790	0.0062	-0.054	0.918	0.000
ECAR	0.587	0.913	0.0657	-0.22	0.973	0.000
SPP	0.298	0.778	0.0028	-0.923	1.102	0.000
SERC	0.222	0.849	0.0079	-0.523	0.967	0.000
WECC	0.091	0.818	0.2861	-0.070	0.724	0.000
WECCC	0.072	0.732	0.0386	-0.902	1.012	0.000
ERCOT	0.000	0.738	0.0247	0.050	0.741	0.000
NPCCI	0.003	0.637	0.0009	0.710	0.628	0.000

[a] MacKinnon approximate p-value for the null hypothesis that the variable is nonstationary.

RICE UNIVERSITY

Translog expenditure function

$$\ln Exp = a + \sum_i b_i \ln RC_i + \sum_i c_i \ln\left(K_i \cdot HR_i\right) + d_1 \ln FE + d_2 \left(\ln FE\right)^2$$
$$+ \frac{1}{2}\sum_i \sum_j e_{ij} \ln RC_i \ln RC_j + \frac{1}{2}\sum_i \sum_j f_{ij} \ln RC_i \ln\left(K_j \cdot HR_j\right)$$
$$+ \frac{1}{2}\sum_i g_i \ln RC_i \ln FE + \frac{1}{2}\sum_i h_i \ln\left(K_i \cdot HR_i\right) \ln FE$$

$$\frac{\partial \ln Exp}{\partial \ln RC_i} = \frac{RC_i}{Exp}\frac{\partial Exp}{\partial RC_i} = S_i \quad \Rightarrow \quad \begin{aligned} S_t^{NG} &= \alpha_0 + \alpha_1 RC_{NG,t} + \alpha_2 RC_{oil,t} + \alpha_3 RC_{coal,t} + \alpha_4 \ln HR_{NG,t} \cdot K_{NG,t} \\ &+ \alpha_5 \ln HR_{oil,t} \cdot K_{oil,t} + \alpha_6 \ln HR_{coal,t} \cdot K_{coal,t} + \alpha_7 \ln FE_t \end{aligned}$$

- Real expenditure on fuel *i* was calculated as the product of the real cost RC_i times the corresponding amount of fuel consumed
- The natural gas expenditure share was taken as the dependent variable
  - Like real coal cost, it was stationary in all regions except FRCC and WECC
- Total fossil fuel generation in the region (*FE*) was taken as the output measure

RICE UNIVERSITY

Differences with prior literature

- We use a panel approach, treating each NERC sub-region as a cross-sectional unit
  - But we also examine time series results for each sub-region
  - Most prior translog analysis is cross-sectional only
- We adjust real costs for heat rates to account for technological change
- Since the real costs of natural gas and oil are cointegrated, we use a cointegrating error term in place of the individual real costs
  - The error term represents deviation from the long run equilibrium between real oil and natural gas input costs
  - Since the contemporaneous error term will be correlated with the dependent variable by construction, we use the lagged error term as an instrument
- We allow for lagged adjustment to changes in real costs by including the lagged dependent variable as a regressor
  - This also is instrumented in the panel estimation using lagged values
- We also included population weighted weather variables, *HDD*, *CDD*

RICE UNIVERSITY

Panel and NERC translog results

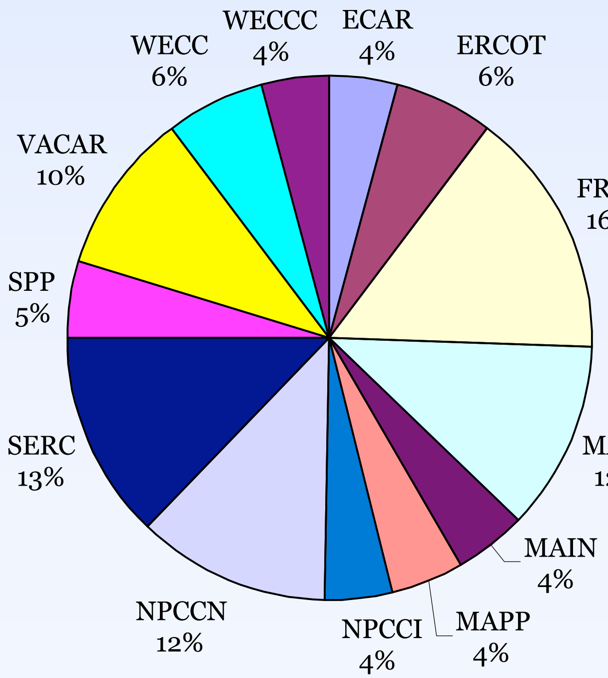
NERC subregion	S_{t-1}^{NG}	$\hat{\omega}_t$	$RC_{coal,t}$	$lnNGCap$	$\ln DfoCap$	$\ln RfoCap$	$\ln CoalCap$	$\ln FE$	CDD	HDD	Number of obs.	R^2 (overall)	R^2 (within)	R^2 (between)
Panel	0.8793*** (0.0160)	-0.0286*** (0.0094)	-0.0372** (0.0172)	0.0200*** (0.0040)				0.0067* (0.0037)	0.0002*** (0.00003)		2184	0.9610	0.8499	0.9880
NPCCN	0.5395*** (0.0860)	-0.1479** (0.0708)	0.1811** (0.0901)							-0.00018** (0.00008)	169	0.6889		
SERC	0.5934*** (0.0616)	-0.0813*** (0.0272)	-0.2210*** (0.0638)		0.2608** (0.1264)			0.1058*** (0.0301)		0.00009** (0.00004)	169	0.9076		
SPP	0.5822*** (0.0694)	-0.0621* (0.0359)	-0.2668*** (0.0846)	-0.1500** (0.0733)	0.8950** (0.4162)		-50.624*** (12.746)		0.0004*** (0.00009)		169	0.9305		
NPCCI	0.6698*** (0.0693)	-0.0613 (0.0692)		0.1152*** (0.0267)				0.0303*** (0.0117)		-0.0002*** (0.00007)	169	0.9069		
WECC	0.8484*** (0.0452)	-0.0577* (0.0299)		0.0215 (0.0137)				0.1263*** (0.0409)	0.00047*** (0.00014)	0.00013** (0.00005)	169	0.9593		
VACAR	0.4176*** (0.0662)	-0.0546* (0.0312)		0.0640** (0.0254)	0.8358 (0.5128)		-0.4758*** (0.1743)		0.0006*** (0.0001)		169	0.8339		
ERCOT	0.5136*** (0.0904)	-0.0413 (0.0513)	-0.1796** (0.0700)	0.1782*** (0.0648)	-0.3183*** (0.0843)		-0.9734** (0.4525)	0.1098*** (0.0314)	0.00016* (0.00009)		169	0.7837		
FRCC	0.7243*** (0.0548)	-0.0299 (0.0329)	-0.2735** (0.1252)	0.0637*** (0.0237)							169	0.8743		
MAIN	0.4226*** (0.0538)	-0.0237 (0.0231)			-0.0931*** (0.0337)			0.0688*** (0.0170)	0.0007*** (0.00008)		169	0.8171		
MAPP	0.6187*** (0.0468)	0.0060 (0.0113)				-0.0076*** (0.0013)			0.0004*** (0.00005)		169	0.8745		
MAAC	0.5660*** (0.0721)	0.0086 (0.0305)		0.0663*** (0.0215)		-0.5839** (0.2611)		-0.0336** (0.0156)	0.0005*** (0.0001)		169	0.8756		
ECAR	0.5083*** (0.0582)	0.0134 (0.0173)		0.0184*** (0.0035)					0.0003*** (0.00005)		169	0.8020		
WECCC	0.2653*** (0.0777)	0.0159* (0.0096)		0.0454*** (0.0165)							169	0.3228		

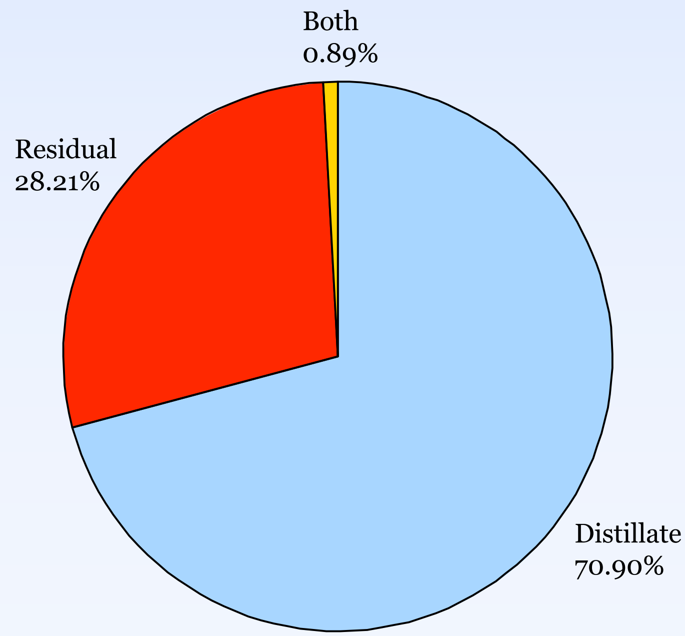
*** indicates significance at the 1% level, ** indicates significance at the 5 % level, and * indicates significance at the 10% level. Statistically insignificant variables are reported in grayed font.

RICE UNIVERSITY

Plant-level switching

- 167 plants with 136,000 MW of capacity – 18% of generation capacity in the United States – was found to be multiple fueled
  - Approximately 13% of that capacity is industrial cogeneration
- Much of it is on the east coast (FL, GA, NC, SC, NY, NJ, PA, VA)
  - This raises doubts about the translog results which do not show strong substitution in FRCC, MAAC and to a lesser extent VACAR





RICE UNIVERSITY

Plant level model

- The model we estimate:

$$NGPct_{i,t} = \frac{\text{NG Consumption}_{\text{i,t}}}{\text{NG Consumption}_{\text{i,t}} + \text{Oil Consumption}_{\text{i,t}}}$$

$$= a_0 + a_1 NGPct_{i,t-1} + a_2 \ln\left(\frac{P_{it}^{NG}}{P_{it}^{E}}\right) + a_3 \ln\left(\frac{P_{it}^{oil}}{P_{it}^{E}}\right) + b_1 CDD_{i,t} + b_2 HDD_{i,t} + \sum_j c_j Month_{j,t} + v_i + \varepsilon_{i,t}$$

- Fuel consumption is in MMBtu
- A plant is included in period *t* only if it generates that month
- The lagged dependent variable again allows for dynamic adjustment
- Heat rates are not included since we have them by plant only
- Monthly variables allow for systematic seasonal outages
- If only one fuel is used in a month $NGPct_{i,t}$ would be zero or one
- We use a panel data Tobit model to allow for this truncation

RICE UNIVERSITY

Tobit results:

	All switching plants	Natural gas and distillate	Natural gas and residual
$NGPct_{t-1}$	0.6635*** (0.0083)	0.7108*** (0.0112)	0.6317*** (0.0126)
$\ln p_t^{NG}$	–0.0590*** (0.0112)	–0.0560*** (0.0151)	–0.1055*** (0.0157)
$\ln p_t^{rfo}$	0.0715*** (0.0210)		0.0921*** (0.0174)
$\ln p_t^{dfo}$	0.0269 (0.0221)	0.1150*** (0.0178)	
CDD	–0.00015*** (0.000027)	–0.00015*** (0.000035)	–0.00003 (0.00004)
HDD	–0.00011*** (0.000014)	–0.00010*** (0.000018)	–0.00013*** (0.00002)
σ_υ	0.3262*** (0.0182)	0.3476*** (0.0237)	0.2669*** (0.0299)
σ_ε	0.3133*** (0.0020)	0.3465*** (0.0029)	0.2492*** (0.0026)
observations	26290	17558	7230
left-censored	2610	2271	346
uncensored	12885	7291	4921
right-censored	10795	7996	1963
number of plants	167	110	43
$\ln L$	–9604.41	–7198.39	–1376.54
χ^2 (d.f.)	8251.6 (17)	5412.6 (16)	3187.1 (16)

*** indicates significance at the 1% level, ** indicates significance at the 5% level, and * indicates significance at the 10% level. The chi-square tests the joint significance of the explanatory variables.

RICE UNIVERSITY

Some implications of the results

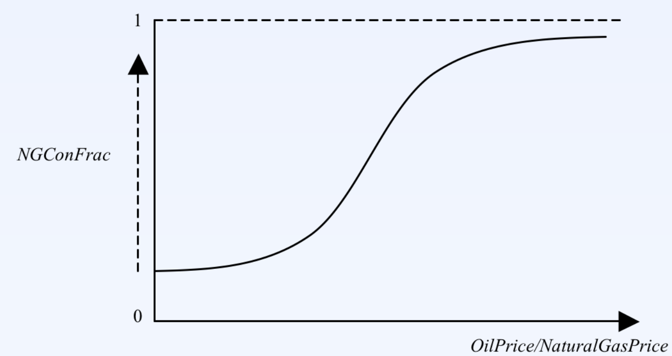
- The strong and statistically significant coefficients on real natural gas and oil product prices, indicate that plants switch fuels in response to price
- The response is stronger for plants that burn residual fuel oil than for plants that burn distillate as the substitute fuel
- The larger coefficient on the real residual fuel oil price in the full sample regression also suggests residual fuel oil is a stronger substitute overall
  - But residual fuel oil and distillate prices are themselves highly correlated
- *CDD*, *HDD negatively* affect switching (holding monthly effects constant)
  - This result is contrary to the more aggregate translog findings
- The strong results for switching plants raise doubts about the translog results for FRCC, MAAC and VACAR
- Some other results in the translog model, such as the effects of some of the capacity variables, are surprising
- Given our focus on natural gas demand, a key benefit of the translog approach, namely enforcing consistency across fuels, is not relevant

Alternative model of fuel use

- Change the dependent variable to

$$\ln\left(\underbrace{-\ln\left(\underbrace{\frac{NGCon}{NGConMax}}_{\in[0,1]}\right)}_{\geq 0}\right) = \ln\left(-\ln NGConFrac_t\right)$$

- *NGConFrac* is actual natural gas use divided by the maximum feasible use if all natural gas plant were operated 24/7 throughout that month
  - Natural gas usage then must lie within the bounds of what is feasible
- The transformation ln(–ln(.)) allows a variable constrained to lie in (0,1) to take any values, as would be required for a standard error term
  - Dependent variable was stationary in all regions except perhaps SERC where the test almost rejects non-stationarity at the 10% level
- As gas usage approaches either limit (0 or the theoretical maximum), it becomes less responsive to further changes in the explanatory variables



RICE UNIVERSITY

Estimated equations

- We estimated the following equations for each region

$$\ln\left(-\ln NGConFrac_t\right) = b_0 + b_1\hat{\omega}_t + b_2 \ln FE_t + b_3 CDD + b_4 HDD + \sum_i Month_i + \varepsilon_t$$

- For a stable system, we expect a positive ω to reduce natural gas demand, which implies its coefficient should be positive
- We expect an increase in total fossil fuel generation (*FE*) to increase natural gas demand, implying its coefficient should be negative
- *CDD* may affect gas demand (holding *FE* and monthly effects fixed) by altering the shape of the load curve, and *HDD* by curtailing natural gas demand by utilities with interruptible contracts
- Monthly influences include other weather effects, seasonal price movements, planned outages and the number of days in a month
- We also allowed the error terms ε to be autocorrelated (ARMA) to allow for slow adjustments, contracting behavior, omitted variables etc.

RICE UNIVERSITY

Estimated adjustment equations

NERC subregion	$\hat{\omega}$	ln *FE*	*CDD*	*HDD*	CAcrisis	AR(1)	MA term	Q-statistic (6 lags)	Q-statistic (12 lags)
Panel	0.0858*** (0.0195)	-0.3897*** (0.0222)	-0.0012*** (0.00007)	-0.00008** (0.00003)	-0.1299*** (0.0360)	panel-specific			
FRCC	0.2997*** (0.0812)	-0.3179*** (0.0773)	-0.0006*** (0.0002)	-0.0006*** (0.0002)		0.9240*** (0.0377)	$u_t - \underset{(0.0937)}{0.2878^{***}} u_{t-1}$	4.8774 (0.5596)	15.684 (0.2061)
VACAR	0.2179** (0.0988)	-0.1820* (0.0953)	-0.0019*** (0.0004)			0.6251*** (0.0628)	$u_t - \underset{(0.0888)}{0.2145^{**}} u_{t-5} + \underset{(0.0846)}{0.1994^{***}} u_{t-10}$	7.6000 (0.2689)	14.080 (0.2957)
MAAC	0.1727** (0.0719)	-0.2320*** (0.0404)	-0.0020*** (0.0004)			0.6863*** (0.0624)		3.5886 (0.7321)	10.221 (0.5966)
MAIN	0.1358** (0.0551)	-0.5167*** (0.1251)	-0.0020*** (0.0003)	-0.0002*** (0.0001)		0.9542*** (0.0296)	$u_t - \underset{(0.0878)}{0.3617^{***}} u_{t-1}$	8.0131 (0.2371)	16.225 (0.1812)
MAPP	0.1079** (0.0505)		-0.0020*** (0.0002)	-0.0001** (0.00004)		0.5049*** (0.0726)		5.266 (0.5102)	11.134 (0.5175)
NPCCN	0.0926*** (0.0321)	-0.2003*** (0.0665)	-0.0008** (0.0003)			0.8408*** (0.0456)		1.7551 (0.9408)	7.1401 (0.8482)
ECAR	0.0477 (0.0750)	-0.5512*** (0.2023)	-0.0017*** (0.0002)			0.9613*** (0.0302)	$u_t - \underset{(0.0980)}{0.4453^{***}} u_{t-1}$	8.2623 (0.2195)	17.385 (0.1357)
SPP	0.0334 (0.0564)	-0.6012*** (0.1287)	-0.0013*** (0.0002)	-0.0003*** (0.0001)		0.9204*** (0.0289)	$u_t + \underset{(0.0819)}{0.2492^{***}} u_{t-11} + \underset{(0.0898)}{0.4300^{***}} u_{t-13}$	3.5442 (0.7381)	8.8185 (0.7184)
SERC	0.0318 (0.0520)	-0.5853*** (0.0804)	-0.0007*** (0.0002)	-0.0001* (0.00006)		0.9412*** (0.0285)		7.8639 (0.2482)	10.156 (0.6023)
WECC	0.0125 (0.0649)	-0.7045*** (0.1078)	-0.0013*** (0.0004)	-0.0002*** (0.00009)		0.8908*** (0.0355)		1.9780 (0.9217)	12.204 (0.4294)
WECCC	0.0109 (0.0228)	-0.4172*** (0.0080)	-0.0002*** (0.00009)		-0.1027*** (0.0251)	0.9558*** (0.0340)	$u_t - \underset{(0.1052)}{0.1786^{*}} u_{t-1}$	2.9401 (0.8163)	6.6784 (0.8781)
NPCCI	0.0057 (0.0830)	-0.4918*** (0.0495)		0.0004*** (0.0001)		0.9545*** (0.0388)	$u_t - \underset{(0.1367)}{0.3331^{**}} u_{t-3} - \underset{(0.1029)}{0.2471^{**}} u_{t-6} - \underset{(0.0972)}{0.2291^{**}} u_{t-9} + \underset{(0.0837)}{0.2990^{***}} u_{t-11}$	7.2234 (0.3007)	15.811 (0.2000)
ERCOT	0.0009 (0.0320)	-0.4000*** (0.0345)	-0.0007*** (0.0001)	-0.0002*** (0.00006)		0.9663*** (0.0262)		7.2234 (0.3007)	15.811 (0.2000)

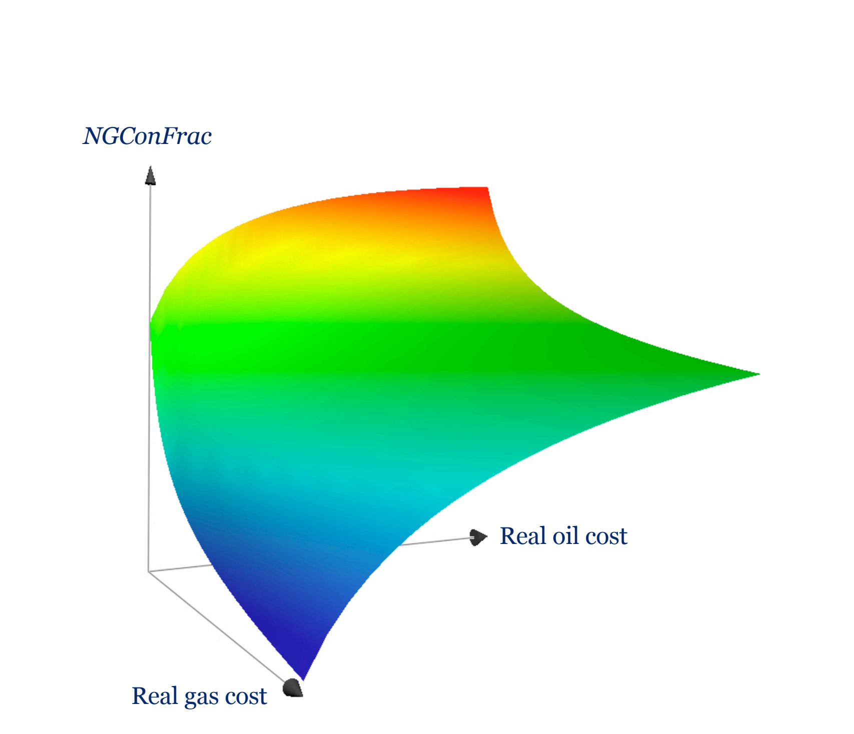
*** indicates significance at the 1% level, ** indicates significance at the 5% level, and * indicates significance at the 10% level. Statistically insignificant variables are reported in grayed font

Implications of the "average" equation

- Estimated panel equation (excluding region-specific monthly effects)

$$\ln(-\ln NGConFrac_t) = b_0 + 0.0858\omega_t - 0.3897\ln FE_t - 0.0012CDD_t - 0.00008HDD_t$$

- A positive deviation from the long run cost relationship (relatively high gas prices) would reduce natural gas demand as hypothesized
- Estimated response surface for the range of cost variations in the data
  - A decline in gas costs holding oil costs fixed increases the use of gas capacity at an increasing rate
  - An increase in oil costs holding gas costs fixed increases the use of gas capacity at a decreasing rate
  - This is similar to the response of cost share in the translog function



Rice University

Effects of other variables

- The average value of *NGConFrac* is 11.8%, with a minimum of 0.088% and a maximum of 71.53%
  - However, there are large and statistically significant differences across regions
- At the mean values for *NGConFrac,* and holding other effects fixed,
  - A 1% increase in *FE* above its mean would raise *NGConFrac* by around 0.098 percentage points
  - A one standard deviation increase in ω would reduce *NGConFrac* by around 0.39 percentage points
  - A one standard deviation (around 150.6) increase in *CDD* above its mean (111.6) would raise *NGConFrac* by around 4.8 percentage points
  - A one standard deviation (around 387.6) increase in *HDD* above its mean (371.6) would raise *NGConFrac* by around 0.8 percentage points
- Quite a few of the regions have strong seasonal patterns (holding *CDD* and *HDD* fixed), but there is no uniform tendency for the same months to be correlated with higher or lower natural gas demand in all regions

RICE UNIVERSITY

Implications of the results

- Key regions for arbitrage between gas and oil prices:
  - FRCC, VACAR, MAAC, MAIN, MAPP, NPCCN i.e.. east coast from Florida to NY and PA and Mid-Western states centered on IL, WI, IA and MN
  - ECAR, SPP and SERC also have reasonably large responses, but the standard errors of the estimates make the conclusion less certain
  - The results for the second specification imply that WECC, WECCC, NPCCI and ERCOT play a very minor role in arbitraging gas/oil price differences
    - However, ERCOT and WECC each contain about 6% of the flexible fuel plants
    - ERCOT and WECC are estimated to have stronger responses to price differences under the translog specification
  - The translog results suggest that coal may be a complement to gas in SERC, SPP, ERCOT and FRCC, while it is a substitute in NPCCN
- The second specification implies that all regions except MAPP increase natural gas use in response to increased fossil fuel generation overall
  - The translog results also imply this for SERC, NPCCI, WECC, ERCOT and MAIN, but they suggest higher *FE reduces* gas share in MAAC
- Higher CDD increases NG input in all regions except NPCCI in the second specification and all regions except NPCCI, NPCCN, SERC, FRCC and WECCC in the translog specification
  - HDD has small effects in eight regions in the second specification and four regions in the translog specification
- Autocorrelated errors could represent lagged adjustment or omitted variables

RICE UNIVERSITY

Concluding Remarks

- The disaggregated evidence supports the hypothesis that the power sector plays a strong role in maintaining a long run relationship between gas and oil prices
  - Our overall results are consistent with a market focus on price relativities between natural gas and oil being reestablished over the longer run
- Residual fuel oil appears to be the primary competitor for natural gas
- The micro analysis showed that weather shocks play an important role in affecting gas demand for power generation and hence price movements
  - These observations are consistent with the market focus on weather as a critical variable influencing energy price movements
  - They also are consistent with the aggregate results showing weather as important sources of deviations in price relativities in the short run