

Appendix C

Statistical Considerations

The monthly sales (volume and price) and monthly deliveries (volume) of natural gas to residential, commercial and industrial consumers presented in this report by State are estimated from data reported on the Form EIA-857, "Monthly Report of Natural Gas Purchases and Deliveries to Consumers." Monthly prices in select states (currently Georgia, Maryland, New York and Ohio) are supplemented with data from the Form EIA-910 "Monthly Natural Gas Marketer Survey". (See Appendix B for a description of these Forms.) These estimations must be made from the reported data since the Form EIA-857 is a sample survey. A description of the sample design and the estimation procedures is given below.

Sample Design

The Form EIA-857 is a monthly sample survey of companies delivering natural gas to consumers. It includes inter- and intrastate pipeline companies, and producers, as well as local distribution companies. The survey provides data that are used each month to estimate the volume of natural gas delivered and the price for onsystem sales of natural gas by State to three consumer sectors—residential, commercial, and industrial. Monthly deliveries and prices of natural gas to the electric power sector are reported on the Form EIA-906, "Power Plant Report, and the Form FERC-423, "Monthly Report of Costs and Quality of Fuels for Electric Plants."

Sample Universe. The sample currently in use was selected from a universe of 1,556 companies. These companies were respondents to the Form EIA-176, "Annual Report of Natural and Supplemental Gas Supply and Disposition," for reporting year 2001 who reported sales or deliveries to consumers in the residential, commercial or industrial sectors. (See Appendix B for a description of the Form EIA-176.)

Sampling Plan. The goal was a sample that would provide estimates of monthly natural gas consumption by the three consuming sectors within each State and the District of Columbia. A stratified sample using a single stage and systematic selection with probability proportional to size was designed.

The measure of size was the volume of natural gas physically delivered in the State to the three consuming sectors by the company in 2001. There were two strata—companies selected with certainty and companies selected under the systematic probability proportional to size design.

Initial calculations showed that a 25 percent sample of companies would yield reasonably accurate estimates. The sample was selected independently in each State, resulting in a national total of 405 respondent companies.

Certainty Stratum. Since estimates were needed for each of the 50 States and the District of Columbia, the strata were established independently within each State. In 16 States and the District of Columbia where sampling was not feasible due to small numbers of companies and/or small volumes of gas deliveries, all companies were selected. The 16 States were: Alaska, Connecticut, Delaware, Hawaii, Idaho, Maine, North Dakota, New Hampshire, New Jersey, Nevada, Oregon, Rhode Island, South Dakota, Utah, Vermont, and Washington.

For each of the remaining States, the total volumes of industrial sales and deliveries and of the combined residential/commercial sales and deliveries were determined. Companies with natural gas deliveries to the industrial sector or to the combined residential/commercial sector above a certain level were selected with certainty. Since a few large companies often account for most of the natural gas delivered within a State, this ensures those companies' inclusion in the sample. The formula for determining certainty was applied independently in the two consumer sectors—the industrial and the combined residential/commercial. These selected companies, together with the companies in the jurisdictions discussed where sampling was not feasible, formed the certainty stratum.

All companies with natural gas deliveries in sector j greater than the cut-off value (C_j) were included in the certainty stratum. The formula for C_j was:

$$C_j = \frac{X_j}{2n} \quad (1)$$

where:

$C_{.j}$ = cutoff value for consumer sector j,

n = target sample size to be selected for the State, 25 percent of the companies in the State,

X_{ij} = the annual volume of natural gas deliveries by company i to customers in consumer sector j,

$X_{i.}$ = the sum within State of annual gas volumes for company i,

$X_{.j}$ = the sum within State of annual gas volumes in consumer sector j,

$X_{..}$ = the sum within State of annual gas volumes in all consumer sectors.

Noncertainty Stratum. All other companies formed the noncertainty stratum. They were systematically sampled with probability proportional to size. The measure of size for each company was the total volume of gas sales to all consumer sectors ($X_{i.}$). The number of companies to be selected from the noncertainty stratum was calculated for each State, with a minimum of 2.

The formula for selecting the number of noncertainty stratum companies was:

$$m = n \frac{X_2}{X_{..}} \quad (2)$$

where:

m = the sample size for the noncertainty stratum within a State,

X_2 = the sum within State of the $X_{i.}$ for all companies in the noncertainty stratum.

Companies were listed in ascending order according to their measure of size and then a cumulative measure of size in the stratum was calculated for each company. The cumulative measure of size was the sum of the measures of size for that company and all preceding companies on the list. An interval of width I for selecting the companies systematically was calculated using.

A uniform random number R was selected between zero and $\left(I = \frac{X_2}{m} \right) I$. The first sampled company was

the first company on the list to have a cumulative measure of size greater than R. The second company selected was the first company on the list to have a cumulative measure of size greater than R + I. R + I

was increased again by I to determine the third company to be selected. This procedure was repeated until the entire sample was drawn.

Subgroups. In four States, the noncertainty stratum was divided into subgroups to ensure that gas in each consumer sector could be estimated. The systematic sample with probability proportional to size design described above was applied independently in each subgroup. The methods for determining the subgroup sample size and calculating the subgroup interval for sample selection were the same as the methods described above for the noncertainty stratum, except that X_2 was the sum within State of the $X_{i.}$ for only those companies in the subgroup.

These subgroups were defined only for the purpose of sample selection. They are:

Kansas, Louisiana, Texas: companies delivering gas only to industrial consumers and those delivering to any other sector.

South Carolina: companies delivering more than 3 Bcf to consumers and those below that level.

Estimation Procedures

Estimates of Volumes. A ratio estimator is applied to the volumes reported in each State by the sampled companies to estimate the total gas sales and deliveries for the State. Ratio estimators are calculated for each consumer sector – residential, commercial, and industrial – in each State where companies are sampled. The following annual data are taken from the most recent submissions of Form EIA-176:

The formula for calculating the ratio estimator (E_{vj}) for the volume of gas in consumer sector j is:

$$E_{vj} = \frac{\gamma_{.j}}{\gamma'_{.j}} \quad (3)$$

where:

γ_j = the sum within State of annual gas volumes in consumer sector j for all companies,

γ'_j = the sum within State of annual gas volumes in consumer sector j for those companies in the sample.

The ratio estimator is applied as follows:

$$V_{vj} = y_{.j} \times E_{vj} \quad (4)$$

where:

V_j = the State estimate of monthly gas volumes in consumer sector j ,

y_j = the sum within State of reported monthly gas volumes in consumer sector j .

Computation of Natural Gas Prices. The natural gas volumes that are included in the computation of prices represent only those volumes associated with natural gas sales by natural gas companies except as explained below.

The price of natural gas for a State within a sector is calculated as follows:

$$P_j = \frac{R_j}{V_j} \quad (5)$$

where:

P_j = the average price for gas sales within the State in consumer sector j ,

R_j = the reported revenue from natural gas sales within the State in consumer sector j ,

V_j = the reported volume of natural gas sales within the State in consumer sector j .

All average prices are weighted by their corresponding sales volume estimates when national average prices are computed.

The monthly average prices of natural gas to residential and commercial consumers in Georgia, Maryland, New York, Ohio and Pennsylvania are monthly average prices of natural gas are based on total sales (sales by local distribution companies and natural gas marketers). Volumes of gas delivered for the account of others to these consumer sectors are not included in the State or national average prices except in these states.

The price of natural gas in the residential and commercial sectors in Georgia, Maryland, New York, Ohio and Pennsylvania is calculated as follows:

$$P_c = \left[\left(\frac{R_s}{V_s} \right) * \left(\frac{V_s}{V_s + V_t} \right) \right] + \left[\left(\frac{Rm_s}{Vm_s} \right) * \left(\frac{V_t}{V_s + V_t} \right) \right] \quad (6)$$

P_c = the combined average price for gas sales by local distribution companies and marketers within the State in sector s (residential or commercial)

R_s = the reported revenue from natural gas sales by local distribution companies within the State in s (residential or commercial)

V_s = the reported volume of natural gas sales by local distribution companies within the State in s (residential or commercial)

V_t = the reported volume of natural gas transported by local distribution companies for marketers within the State in s (residential or commercial)

Rm_s = the reported revenue from natural gas sales by marketers within the State in s (residential or commercial)

Vm_s = the reported volume of natural gas sales by a marketer within the State in s (residential or commercial)

Table 25 shows the percent of the total State volume that represents volumes from natural gas sales to the commercial and industrial sectors. This table may be helpful in evaluating commercial and industrial price data. All natural gas prices to the residential sector represent onsystem sales volumes only except in Georgia, Maryland, New York, Ohio and Pennsylvania.

See the section on consumer price calculations in this Appendix for further price information.

Estimation for Nonrespondents. A volume for each consumer category is imputed for companies that fail to respond. The imputation is based on the previous month's value reported by the non-responding company and the change from the previous month to the current month in volumes reported by other companies in the State. The imputed volumes are included in the State totals. To estimate prices for non-respondents, the unit price (dollars per thousand cubic feet) reported by the company in the previous month is used.

The formula for imputing volumes of gas volumes for nonrespondents was:

$$F_t = F_{t-1} \times \frac{y_{jt}}{y_{j,t-1}} \quad (7)$$

where:

F_t = imputed gas volume for current month t ,

F_{t-1} = gas volume for the company for the previous month,

y_{jt} = gas volume reported by companies in the State stratum for report month t ,

y_{jt-1} = gas volume in the previous month for companies in the State stratum that reported in month t.

Final Revisions

Adjusting Monthly Data to Annual Data. After the annual data reported on the Form EIA-176 have been submitted, edited, and prepared for publication in the *Natural Gas Annual*, revisions are made to monthly data. The revisions are made to the volumes and prices of natural gas delivered to consumers that have appeared in the *Natural Gas Monthly (NGM)* to match them to the annual values appearing in the *Natural Gas Annual*. The revised monthly estimates allocate the difference between the sum of monthly estimates and the annual reports according to the distribution of the estimated values across the months.

Before the final revisions are made, changes or additions to submitted data received after publication of the monthly estimate and not sufficiently large to require a revision to be published in the *NGM*, are used to derive an updated estimate of monthly consumption and revenues for each State's residential, commercial, or industrial natural gas consumption.

For each State, two numbers are revised, the estimated consumption and the estimated price per thousand cubic feet.

The formula for revising the estimated consumption is:

$$V_{jm}^* = V_{jm} + \left[(V_{ja} - V'_{jm}) \left(\frac{V_{jm}}{V'_{jm}} \right) \right] \quad (8)$$

where:

V_{jm}^* = the final volume estimate for month m in consumer sector j,

V_{jm} = the estimated volume for month m in consumer sector j,

V_{ja} = the volume for the year reported on Form EIA-176,

V'_{jm} = the annual sum of estimated monthly volumes

The price is calculated as described above in the Estimation Procedures section, using the final revised consumption estimate and a revised revenue estimate.

The formula for revising the estimated revenue is:

$$R_{jm}^* = R_{jm} + \left[(R_{ja} - R'_{jm}) \left(\frac{R_{jm}}{R'_{jm}} \right) \right] \quad (9)$$

where:

R_{jm}^* = the final revenue estimate for month m in consumer sector j,

R_{jm} = the estimated revenue for month m in consumer sector j,

R_{ja} = the revenue for the year reported on Form EIA-176,

R'_{jm} = The annual sum of estimated monthly revenues.

Revision of Volumes and Prices for Deliveries to Electric Power Sector. Revisions to monthly deliveries to the electric power sector are published throughout the year as they become available.

Reliability of Monthly Data

The monthly data published in this report are subject to two sources of error - nonsampling error and sampling error. Nonsampling errors occur in the collection and processing of the data. See the discussion of the Form EIA-857 in Appendix B for a description of nonsampling errors for monthly data.

Sampling error may be defined as the difference between the results obtained from a sample and the results that a complete enumeration would provide. The standard error statistic is a measurement of sampling error.

Standard Errors. A standard error of an estimate is a statistical measure that indicates how the estimate from the sample compares to the result from a complete enumeration. Standard errors are calculated based on statistical theory that refers to all possible samples of the same size and design.

The standard errors for monthly natural gas volume estimates by State are given in Table C1. Ninety-five percent of the time, the volume that would have been obtained from a complete enumeration will lie in the range between the estimated volume minus two standard errors and the estimated volume plus two standard errors.

The standard error of the natural gas volume estimate is the square root of the variance of the estimate. The formula for calculating the variance of the volume estimate is:

$$V(\hat{\gamma}) = \sum_{h=1}^H \left[N_h^2 \frac{\left(1 - \frac{n_h}{N_h}\right)}{n_h(n_h - 1)} \left(\sum_{i=1}^{n_h} (y_i - Tx_j)^2 \right) \right] \quad (10)$$

where:

H = the total number of strata

N_h = the total number of companies in stratum h

n_h = the sample size in stratum h

y_i = the reported monthly volume for company i

x_i = the reported annual volume for company i

T = the ratio of the sum of the reported monthly volumes for sample companies to the sum of the reported annual volumes for the sample companies.

Table C-1. Standard Error for Natural Gas Deliveries and Price to Consumers by State, August 2004

State	Volume Million Cubic Feet				Price Dollars per Thousand Cubic Feet		
	Residential	Commercial	Industrial	Total	Residential	Commercial	Industrial
Alabama	111	263	2,402	2,418	0.54	NA	NA
Alaska	0	NA	0	NA	—	—	—
Arizona	2	46	0	46	0.14	0.09	—
Arkansas	1	6	4	8	0.06	0.01	0.03
California	258	94	628	685	0.08	0.18	0.25
Colorado	141	197	573	622	1.01	0.51	0.24
Connecticut	0	0	0	0	—	—	—
Delaware	0	0	0	0	—	—	—
District of Columbia	0	0	0	0	—	—	—
Florida	83	101	278	307	NA	NA	NA
Georgia	227	131	1,184	1,213	NA	NA	NA
Hawaii	0	0	0	0	—	—	—
Idaho	0	0	0	0	—	—	—
Illinois	376	82	394	550	0.19	NA	NA
Indiana	249	102	1,146	1,177	1.53	0.13	0.48
Iowa	93	44	NA	NA	NA	NA	NA
Kansas	26	47	511	514	0.06	0.26	NA
Kentucky	58	416	826	927	0.41	NA	NA
Louisiana	651	199	2,498	2,590	NA	0.12	0.02
Maine	0	0	0	0	—	—	—
Maryland	8	20	22	31	0.07	0.33	0.22
Massachusetts	2	19	NA	NA	NA	NA	NA
Michigan	17	25	76	82	0.02	0.01	0.23
Minnesota	311	20	523	609	0.32	0.18	0.38
Mississippi	171	152	343	412	0.72	0.44	NA
Missouri	19	124	181	220	0.91	0.47	NA
Montana	1	2	0	2	0.11	0.29	—
Nebraska	6	146	431	455	NA	NA	NA
Nevada	0	0	0	0	—	—	—
New Hampshire	0	0	0	0	—	—	—
New Jersey	0	0	0	0	—	—	—
New Mexico	31	28	264	268	0.21	0.49	NA
New York	57	NA	163	NA	0.24	0.21	0.98
North Carolina	15	NA	283	NA	0.05	0.10	0.99
North Dakota	0	0	0	0	—	—	—
Ohio	148	572	966	1,132	0.46	NA	0.25
Oklahoma	10	50	1,807	1,808	0.23	0.36	NA
Oregon	0	0	0	0	—	—	—
Pennsylvania	20	169	180	248	0.54	0.15	NA
Rhode Island	0	0	0	0	—	—	—
South Carolina	17	67	154	168	NA	0.31	0.16
South Dakota	0	0	0	0	—	—	—
Tennessee	72	149	300	343	0.20	0.88	0.49
Texas	841	529	0	993	0.64	NA	—
Utah	0	0	NA	NA	—	—	—
Vermont	0	0	0	0	—	—	—
Virginia	47	34	156	166	0.91	0.98	0.66
Washington	NA	NA	NA	NA	NA	NA	NA
West Virginia	19	72	1	74	0.46	0.18	0.02
Wisconsin	105	445	489	670	NA	NA	0.61
Wyoming	7	60	164	175	0.93	0.36	NA
Total	1,295	1,258	12,766	12,893	0.13	0.26	0.25

NA Not Available.
— Not Applicable.

Source: Energy Information Administration, Form EIA-857, "Monthly Report of Natural Gas Purchases and Deliveries to Consumers."