Appendix C
Statistical Considerations

The monthly volumes and prices 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 are supplemented with data from the Form EIA-910, *Monthly Natural Gas Marketer Survey*. (See Appendix B, "Data Sources," for a description of these Forms.)

A description of the sample design and the estimation procedures is given below.

Sample Design

The Form EIA-857 implements a monthly sample survey of companies delivering natural gas to consumers. It includes inter- and intrastate pipeline companies and 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-923, *Power Plant Operations Report*.

Sample Universe. The sample of 323 respondents in use for 2019 and 2020 was selected from a universe of 2,083 companies. These companies were respondents to the Form EIA-176, *Annual Report of Natural and Supplemental Gas Supply and Disposition*, for reporting year 2018, 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. Beginning in 2012, EIA changed its sample design from single stage probability-proportional-to-size ("PPS") to a cut off sample in order to reduce respondent burden and decrease processing costs. The skewness of the volume of reported natural gas within the primary traditional sectors of interest for the EIA-857, residential, commercial, and industrial, is sufficiently large to enable a cut-off design that will maintain estimates of similar quality as under the previous PPS sample design.

One of the purposes of the sample is to provide a ratio useful for scaling the monthly state sector volumes in the sample up to precise estimates of the actual state sector monthly volumes. In cases where at most two suppliers in a state service a particular sector, all suppliers are included in the cut off sample. In states with more than two suppliers in a sector, the cut off sample is constructed by adding suppliers in descending order of volume until the change in the ratio of interest is considered statistically insignificant.

The single stage PPS sample design typically yielded approximately 390 respondents for the EIA-857 during 2009-2011. The cutoff sample approach as described above reduces this figure to approximately 300 respondents.

Estimation Procedures

Estimates of Volumes. Total gas sales and deliveries is estimated as the sum of estimated volume of natural gas sales by EIA-857 sample companies and estimated volume of natural gas transported on behalf of others by EIA-857 sample companies**.** To estimate each component volume for a state, a ratio estimator is calculated separately for the volume of natural gas sales byEIA-857 sample companies and the volume of natural gas transported on behalf of others by EIA-857 sample companies. These ratios are then applied to the reported sales volumes and transportation volumes respectively. This is done in each state by sector for the sampled companies.

Ratio estimators are typically calculated for each consumer sector — residential, commercial, and industrial — in each state where companies are sampled using annual data from the most recent submission of Form EIA-176.

The classical ratio estimator (CRE) for the volume of gas in a state for consumer sector j is:

$$V_j = \sum_{i \in s} y_{ij} + \frac{\sum_{i \in s} y_{ij}}{\sum_{i \in s} x_{ij}} \sum_{i \notin s} x_{ij}$$

or (1)

$$V_j = \frac{\sum_{i \in s} y_{ij}}{\sum_{i \in s} x_{ij}} \sum_{i \in s \cup s^c} x_{ij}$$

where for either the natural gas sales by **EIA-857 sample companies or the** natural gas transported on behalf of others by EIA-857 sample companies:

V_j = the estimate of monthly gas volumes in a state for consumer sector *j*,

y_{ij} = the monthly volume within a state for sector j and company i,

x_{ij} = the annual volume within a state for sector j and company i,

Ideally, the ratio $\sum_{i \in s} y_{ij} \Big/ \sum_{i \in s} x_{ij}$ is calculated separately for each state by sector cross classification. However, because certain states have extremely small sample sizes, we combine some states for estimation purposes only, in order to obtain more stable estimates. The state groupings are referred to as estimation groups. For most states, especially the large ones, the ratio will continue to be estimated separately as described above. The final volume estimates, V_j, will continue to be published separately for all states.

When an estimation group involves more than one state, such states are geographical neighbors, and the states in each group fall within a temperature and precipitation region, as defined by the National Climatic Data Center (NCDC).

Grouping states this way is an application of small area estimation, where more reliable estimates are obtained by “borrowing strength” from neighboring states with similar climate patterns. For more details see http://interstat.statjournals.net/YEAR/2014/articles/1405001.pdf

For most respondents, the estimation of residential and commercial consumption requires one additional step than the industrial sector. In this extra step the company level reported volumes are adjusted using sendout.

Sendout, which was added to Form EIA-857 for the August 2010 report month, is the total volume of gas dispatched for delivery during a calendar month. It should be noted that the adjusted volumes for the residential and commercial sector are used in equation 2.

The adjustment is implemented by taking reported sendout and removing gas that is consumed in company operations and distribution use as well as gas delivered to electric power, industrial, and vehicle use sectors. The remaining gas is allocated to the residential and commercial sectors proportionally based on reported totals for the report month. For a small number of companies, the adjustment to sendout is applied to all sectors in order to better represent their billing practices. In these cases, sendout is applied to all end-use sectors proportionally instead of just to residential and commercial.

The formulas for adjusting the reported residential and commercial volumes at the company level are:

$$Radj = (SO * (1 - Oth) - EP - IN) * \frac{R}{R + C} \qquad (2)$$

$$Cadj = (SO * (1 - Oth) - EP - IN) * \frac{C}{R + C} \qquad (3)$$

where:

R_{adj} = sendout adjusted residential volume,

C_{adj} = sendout adjusted commercial volume,

SO = sendout volume,

Oth = ratio of annual distribution use and vehicle fuel to total deliveries,

EP = reported electric power volume,

IN = reported industrial volume,

R = reported residential volume,

C = reported commercial volume.

Estimates of Revenue. Revenue is calculated similarly to volumes. Sales revenue is estimated as the sum of estimated revenue from natural gas sales by EIA-857 sample companies for each sector by state and applying the ratio estimator in equation 1. Similar to volumes, states are combined if the sample size (by sector) is too small in order to estimate the ratio.

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 within a state for sector j is calculated as follows:

$$P_j = \frac{R_j}{V_j} \quad (4)$$

where:

P_j = the average price for gas sales within a state for consumer sector *j*,

R_j = the reported revenue from natural gas sales within a state for consumer sector *j*,

V_j = the reported volume of natural gas sales within a state for consumer sector *j*.

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

For select states (Georgia, New York and Ohio), the monthly average prices of natural gas to residential and commercial consumers are based on total sales (sales by local distribution companies and natural gas marketers). See the Form EIA-910 description in Appendix B, "Data Sources," for details.

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 where EIA-910 data are used is calculated as follows:

$$P_c = \left[\left(\frac{R_s}{V_s}\right) * \left(\frac{V_s}{V_s + V_t}\right)\right] + \left[\left(\frac{Rm_s}{Vm_s}\right) * \left(\frac{V_t}{V_s + V_t}\right)\right] = \frac{R_s + \left(\frac{Rm_s}{Vm_s}\right) * V_t}{V_s + V_t} \quad (5a)$$

for Georgia, and,

$$P_c = \left[\left(\frac{R_s}{V_s}\right) * \left(\frac{V_s}{V_s + V_t}\right)\right] + \left[\left(\frac{Rm_s}{Vm_s} + \frac{R_t}{V_t}\right) * \left(\frac{V_t}{V_s + V_t}\right)\right] = \frac{R_s + \left(\frac{Rm_s}{Vm_s}\right) * V_t + R_t}{V_s + V_t} \quad (5b)$$

for New York and Ohio, where:

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 estimated revenue from natural gas sales by local distribution companies within the state in sector *s* (residential or commercial),

V_s = the estimated volume of natural gas sales by local distribution companies within the state in sector *s* (residential or commercial),

V_t = the estimated volume of natural gas transported by local distribution companies for marketers within the state in sector *s* (residential or commercial),

Rm_s = the reported revenue from natural gas sales by marketers within the state in sector *s* (residential or commercial),

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

R_t = the estimated revenue from natural gas transported by local distribution companies for marketers within the state in sector *s* (residential or commercial).

Table 24 shows the percent of the total state volume that represents volumes from natural gas sales to the residential, commercial, and industrial sectors. This table may be helpful in evaluating residential, commercial, and industrial price data.

Estimation for Nonrespondents and Edit Failures.

A volume for each delivered and transported consumer category is imputed for companies that fail to respond in time for inclusion in the published estimates (unit nonresponse) or for which reported volumes have failed the edit and not been confirmed or corrected (item nonresponse).

In the case of unit nonresponse and edit failures, the imputed volumes and revenues for the residential, commercial, and industrial sectors are predicted using CRE. In other words, they are predicted the same way as out of sample cases. In particular, for company i, which can be a nonrespondent, edit failure or not sampled, the predicted value is given as:

$$y_{ij}^{*} = \frac{\sum_{i \in s, j} y_{ij}}{\sum_{i \in s, j} x_{ij}} x^{*}{}_{ij}$$

where:

y_{ij}^{*} = the estimated value (volume or revenue) of the nonrespondent company or for a company value that fails an edit,

x_{ij}^{*} = the reported value (volume or revenue) of the nonrespondent company from the latest Form EIA 176 survey.

Again the ratio $\sum_{i \in s} y_{ij} \Big/ \sum_{i \in s} x_{ij}$ is calculated separately for most states, but several states can be combined as described in the "Estimates of Volume" section above when sample size is too small for a state by sector classification.

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* (NGA), 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:

$$V^*_{jm} = V_{jm}\left(\frac{V_{ja}}{V'_{jm}}\right) \qquad (6)$$

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:

$$R^*_{jm} = R_{jm}\left(\frac{R_{ja}}{R'_{jm}}\right) \qquad (7)$$

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.

The standard errors for monthly natural gas volume and price estimates by state are given in Table C1. They are constructed so that over the long run the true value is within two standard errors of its estimate 95% of the time. For volume estimates, the relative standard error (RSE) is published as a percent. Standard errors of prices are expressed in dollars.

The relative standard error of a natural gas volume estimate is the square root of the variance of the estimate divided by the published volume. The formula for calculating the variance of the volume estimate is:

$$Var(V_{.j}) = \sigma_\epsilon^2 \sum_{i=n+1}^{N} x_{ij} + \left(\sigma_\epsilon^2 \Big/ \sum_{i=1}^{n} x_{ij} \right) \left(\sum_{i=n+1}^{N} x_{ij} \right)^2 \qquad (8)$$

where:

$V_{.j}$ = the estimate of monthly gas volumes in a state for sector j,

$\sigma_e^2 = \sum_{i=1}^{n} \sigma_{0i}^2 \Big/ (n-1)$, the sum of squared residuals,

$\sigma_{0i} = \frac{(y_{ij} - T_{.j} x_{ij})}{\sqrt{x_{ij}}}$, the weighted residual for company i,

$T_{.j}$ = the ratio estimator defined in equation 1,

x_{ij} = the reported annual volume for company i in sector j,

y_{ij} = the reported monthly volume for company i in sector j,

n = sample size (i > n implies company i not in sample),

N = frame size.

As the price of natural gas is calculated as a quotient of revenue to volume, the standard error of natural gas prices is derived using the variance properties of a quotient. The variance of revenue is calculated in the same manner as the volumes described in equation 8, with revenue substituted for volume. The formula for calculating the standard error of the price in equation 4 is:

$$Var(P_j) = Var\left(\frac{R_j}{V_j}\right) = \left(\frac{R_j}{V_j}\right)^2 * \left(\frac{Var(R_j)}{R_j^2} + \frac{Var(V_j)}{V_j^2} - 2\frac{Cov(R_j,V_j)}{R_j * V_j}\right) \quad (9)$$

where:

P_j = natural gas price in sector j,

V_j = volume in sector j, which is estimated in equation 1,

R_j = revenue in sector j, which is estimated in same manner as equation 1,

$Var(V_j)$ = the variance of volume in sector j calculated using equation 8,

$Var(R_j)$ = the variance of revenue in sector j calculated in the same manner as equation 8,

$Cov(V_j,R_j)$ = the covariance between revenue and volume in sector j. The formula for estimating the covariance is:

$$\left[\left(\sum_{i=n+1}^{N} x_{Rij}\right)\left(\sum_{i=n+1}^{N} x_{Vij}\right)\frac{\sum_{i=1}^{n}\sqrt{x_{Rij}x_{Vij}}}{\left(\sum_{i=1}^{n} x_{Rij}\right)\left(\sum_{i=1}^{n} x_{Vij}\right)} + \sum_{i=n+1}^{N}\sqrt{x_{Rij}x_{Vij}}\right] * \left[\left(\frac{1}{n-1}\right)\sum_{i=1}^{n}\sqrt{x_{Rij}x_{Vij}}\left(y_{Rij} - T_{Rj}\right)\left(y_{Vij} - T_{vj}\right)\right]$$

and,

x_{Rij} = the reported annual revenue for company i in sector j,

x_{Vij} = the reported annual volume for company i in sector j,

y_{Rij} = the reported monthly revenue for company i in sector j,

y_{Vij} = the reported monthly volume for company i in sector j,

$$T_{Rj} = \frac{\sum_{i\in s} y_{Rij}}{\sum_{i\in s} x_{Rij}},$$

$$T_{Vj} = \frac{\sum_{i\in s} y_{Vij}}{\sum_{i\in s} x_{Vij}}.$$

The formula for calculating the standard error of price in equations 5a and 5b is derived after a suitable expansion that involves repeated applications of equations 8 and 9.

Table C1. Standard error for natural gas deliveries and price by consumers, by state, July 2020

State	Volume Relative Standard Error (Percent)				Price (Dollars per Thousand Cubic Feet)		
	Residential	Commercial	Industrial	Total	Residential	Commercial	Industrial
Alabama	3.0	1.9	1.9	1.8	0.40	0.10	0.04
Alaska	1.4	2.8	1.8	1.2	0.16	0.28	0.38
Arizona	2.9	4.8	3.2	2.4	0.54	0.28	0.03
Arkansas	1.6	0.1	5.9	4.0	0.18	0.02	0.12
California	1.2	0.6	1.5	1.0	0.23	0.11	0.15
Colorado	1.4	1.4	2.5	1.4	0.17	0.10	0.48
Connecticut	0.3	2.7	0.2	1.2	0.07	0.30	0.01
Delaware	0.0	0.0	0.0	0.0	0.00	0.00	0.00
District of Columbia	0.0	0.0	0.0	0.0	0.00	0.00	0.00
Florida	4.0	1.2	1.2	0.9	0.71	0.43	NA
Georgia	0.0	0.0	1.2	0.0	0.09	0.06	0.18
Hawaii	0.0	0.0	0.0	0.0	0.00	0.00	0.00
Idaho	0.9	0.7	0.1	0.2	0.03	0.03	0.00
Illinois	0.5	1.2	1.7	1.0	0.17	0.11	0.16
Indiana	1.3	1.4	3.0	2.6	0.46	0.42	0.18
Iowa	7.1	2.2	0.8	0.8	NA	0.34	0.27
Kansas	3.2	0.3	3.4	2.7	0.10	0.22	0.13
Kentucky	2.6	2.5	2.7	2.2	0.26	0.28	0.09
Louisiana	4.6	5.5	1.0	1.0	0.56	0.58	0.11
Maine	0.0	0.0	0.0	0.0	0.00	0.00	0.00
Maryland	0.7	3.1	3.9	1.7	0.02	0.03	0.00
Massachusetts	1.4	0.5	1.2	0.6	0.13	0.17	0.12
Michigan	1.4	1.5	1.8	1.0	0.33	0.24	0.05
Minnesota	1.6	1.6	9.0	5.9	0.19	0.03	0.14
Mississippi	0.8	2.5	2.3	2.1	0.14	0.12	0.12
Missouri	1.5	1.5	1.8	1.1	0.20	0.38	0.42
Montana	1.2	0.5	0.0	0.2	0.07	0.15	0.41
Nebraska	1.2	1.7	1.8	1.4	0.11	0.14	0.09
Nevada	0.0	5.2	0.1	1.8	0.00	0.00	0.06
New Hampshire	0.0	0.0	0.0	0.0	0.00	0.00	0.00
New Jersey	0.0	0.0	4.1	1.4	0.00	0.00	1.40
New Mexico	3.2	2.0	2.7	1.6	0.16	0.17	0.02
New York	0.0	0.0	5.2	0.0	0.14	0.06	0.12
North Carolina	8.6	1.0	1.0	1.1	0.52	0.24	0.36
North Dakota	0.5	0.2	1.2	1.0	0.07	0.02	0.02
Ohio	0.0	0.0	3.1	0.0	0.16	0.06	0.42
Oklahoma	NA	NA	6.8	5.9	NA	0.59	0.09
Oregon	0.0	0.0	2.3	1.5	0.00	0.00	0.00
Pennsylvania	1.5	1.2	1.7	1.2	0.09	0.14	0.45
Rhode Island	0.0	0.0	0.0	0.0	0.00	0.00	0.00
South Carolina	4.5	3.4	1.7	1.5	1.06	0.19	0.21
South Dakota	1.6	0.3	2.7	2.4	0.06	0.03	0.00
Tennessee	2.8	2.2	1.4	1.2	0.41	0.20	0.13
Texas	2.3	2.1	2.0	1.8	0.58	0.18	0.04
Utah	0.5	0.2	0.2	0.2	0.00	0.00	0.04
Vermont	0.0	0.0	0.0	0.0	0.00	0.00	0.00
Virginia	1.0	1.9	1.1	0.8	0.29	0.11	0.26
Washington	1.5	1.4	2.6	1.5	0.19	0.09	0.02
West Virginia	3.7	0.6	1.7	1.3	0.53	0.14	0.00
Wisconsin	2.6	1.6	1.7	1.2	0.40	0.08	0.09
Wyoming	0.8	0.8	4.4	3.9	0.25	0.07	0.29
Total	**0.4**	**0.4**	**0.6**	**0.5**	**0.08**	**0.04**	**0.03**

NA Not available.

Source: U.S. Energy Information Administration (EIA): Form EIA-857, *Monthly Report of Natural Gas Purchases and Deliveries to Consumers*.