2009 Impact Evaluation of Ontario Power Authority’s Commercial & Industrial Demand Response Programs

September 9, 2010

Freeman, Sullivan & Co.
101 Montgomery St., 15th Floor
San Francisco, CA  94104

Prepared for:
Nik Schruder
Ontario Power Authority
120 Adelaide Street West, Ste. 1600
Toronto, Ontario

Prepared by:
Josh Bode, M.P.P.
Josh Schellenberg, M.A.
Paul Mangasarian, B.S.
Freeman, Sullivan & Co.
“Copyright of the Ontario Power Authority. This report may not be copied or reproduced in whole or in part without the prior written permission of the Ontario Power Authority.” [This report may be used solely for informational purposes, and should not be relied on by third parties. The OPA makes no representation, warranty or guarantee as to the accuracy or correctness of this report.]
Table of Contents

1 Executive Summary ........................................................................................................... 1
  1.1 DR-1 Summary of Results .......................................................................................... 3
  1.2 DR-2 Summary of Results .......................................................................................... 6
  1.3 DR-3 Summary of Results .......................................................................................... 11

2 Introduction 16
  2.1 Evaluation Objectives and Key Research Questions .................................................. 18
  2.2 How to Use Load Reduction Results ........................................................................... 19
  2.3 Reliability of Demand Response as a Resource ......................................................... 20
  2.4 Report Structure ......................................................................................................... 22

3 Evaluation Methodology .................................................................................................... 23
  3.1 Regression Model Development ................................................................................. 24
  3.2 DR-1 Analysis Approach Overview ............................................................................... 26
  3.3 DR-2 Analysis Approach Overview ............................................................................... 27
  3.4 DR-3 Analysis Approach Overview ............................................................................... 28
  3.5 Assessment of Accuracy and Precision of Evaluation Models .................................... 28

4 DR-1 Demand Buy Back Program ....................................................................................... 30
  4.1 Program Background ................................................................................................... 30
  4.2 Program Participation .................................................................................................. 31
  4.3 2009 Events and Event Condition .............................................................................. 32
  4.4 Participation Bidding Behavior .................................................................................... 33
  4.5 2009 Ex post Load Reduction Results ........................................................................... 36
  4.6 Load Reductions for Long-Term Planning and Cost-Effectiveness ................................ 41
  4.7 Conclusions and Recommendations ........................................................................... 42
1 EXECUTIVE SUMMARY

This report documents the 2009 evaluation of the Ontario Power Authority’s DR-1, DR-2, and DR-3 programs. DR-1 is a voluntary demand response (DR) program offered by the OPA to pilot the concept of DR among large industrial and commercial customers. It is a demand-bidding program in which participants are paid for reducing load when Independent Electricity System Operator (IESO) market prices exceed a strike price proposed by participants. DR-2 is a contractual load shifting program in which participants specify the load shift window and amount of load shifting. The program does not compensate participants for changes in the overall level of electricity consumption that may be associated with economic fluctuations. Load reduction during peak periods must be accompanied by shifting to off-peak periods. DR-3 is a contractual DR program that offers large commercial and industrial consumers reserve payments for being available to provide load reductions when called by the OPA. Participants are required to meet a 95% reliability requirement when they are called, or else financial penalties apply.

The load reduction evaluations presented here were designed to meet the following objectives:

- Develop ex post load reduction estimates for each program for 2009 and 2010;
- Provide estimates of the load reduction capability (i.e., ex ante load reductions) for each program for planning purposes;
- Provide estimates of the operational reliability of each DR program;
- Develop inputs for conducting cost-effectiveness analysis of the programs;
- Provide recommendations for improvement in program design and operations; and
- Comply with OPA Load Impact Protocols1 that guide program evaluations for DR resources in Ontario.

To the extent possible, the three program evaluations comply with OPA’s load impact protocols for DR, which require a rigorous analysis of DR programs based on historical data and program performance. The protocols are designed to produce consistent, transparent, unbiased, and relatively precise load reduction estimates for DR programs. The required outputs also facilitate comparisons of load reductions across DR programs and support cost-effectiveness analysis. The evaluation methods differ from those used for financial settlement with participating customers. They are more robust and methodical since they do not have to be produced as

quickly, nor do they require the same degree of simplicity, as settlement baselines. This evaluation includes an assessment of the accuracy and precision of current settlement methods and provides recommendations for improvement.

The joint evaluations of the DR-1, DR-2, and DR-3 programs highlight the different role of the programs in OPA’s DR portfolio. DR-1 is designed to provide commercial and industrial facilities with an opportunity to test and develop their demand response capability without the risk associated with a firm commitment. Because a firm commitment is not required for DR-1, participants can decide whether or not to bid in load reductions on an event-by-event basis. As such, the bid amounts and realized load reductions vary significantly across events. DR-1 participants are expected to eventually move to other DR programs that require more firm, contractual commitments for load reductions (DR-3) or load shifting (DR-2). Several large customers that were enrolled in DR-1 in 2007 and 2008 have since moved to these other programs.

DR-2 addresses industrial customers that can engage in daily load shifting. It is a contractual commitment designed to produce consistent load reductions.

DR-3 requires a contractual commitment and, as such, provides more consistent load reductions across events. It provides a dispatchable resource to be used by the system operator and requires a dynamic response from participants. In addition, as long as settlement baselines are reasonably accurate, the DR-3 program is self-correcting, since payments are reduced if participants underperform relative to their contracted load reductions.

The graph below provides a look at the demand response capability of the three programs over time. Overall, the load reduction capability has grown steadily since 2008 and the program mix has changed. The largest participants in DR-1 migrated to DR-2 in March, 2010.
1.1 DR-1 Summary of Results

DR-1 program load reductions significantly decreased in 2009 in comparison to prior years. Two main factors contributed to this trend. First, fewer participants were enrolled in the program. The program saw its three largest participants (representing 95% of the contracted load in the program) migrate to a contracted DR program. At the same time, Ontario experienced one of the coolest summers on record, which, along with economic recession, contributed to depressed energy prices in the IESO market. DR-1 curtailment events are triggered by IESO market prices.

Ex ante load reductions are most appropriate for demand response (DR) programs since they are designed to provide an insurance or option-based estimate of load reduction resources under normal and extreme conditions (typically weather or electricity price related). DR programs are designed to help address electricity system capacity issues when the system is under the most stress. Table 1-1 describes remaining DR-1 resources across market price

---

2 For the graph, load reduction capability is defined as follows. For DR-1, the load reduction capability in the top 50 hours is presented. For the DR-2 transition period actual load reductions realized are presented. For all other DR-2 periods and for DR-3, load reduction capability is based on contracted load reduction or shifting and adjustments for scheduled non-performance and unscheduled under or over performance based on historical data – i.e., ex-ante load impacts.
thresholds. As program eligibility is determined by these prices, this is a key metric over which to view available program resources.

Table 1-1:

<table>
<thead>
<tr>
<th>IESO 3-Hour Price Threshold ($/MWh)</th>
<th>Average Reference Load (MW)</th>
<th>Average Bid Amount (MW)</th>
<th>Average Settlement Load Reduction (MW)</th>
<th>Average Regression Based Load Reduction (MW)</th>
<th>Average Number of Participating Customers</th>
<th>Percent Hours Eligible by Price Conditions (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$50</td>
<td>22.34</td>
<td>0.11</td>
<td>0.06</td>
<td>0.01</td>
<td>0.05</td>
<td>3.98</td>
</tr>
<tr>
<td>$60</td>
<td>24.31</td>
<td>0.39</td>
<td>0.17</td>
<td>0.03</td>
<td>0.10</td>
<td>4.69</td>
</tr>
<tr>
<td>$70</td>
<td>24.11</td>
<td>0.72</td>
<td>0.23</td>
<td>0.01</td>
<td>0.18</td>
<td>2.06</td>
</tr>
<tr>
<td>$80</td>
<td>25.07</td>
<td>3.20</td>
<td>1.98</td>
<td>0.22</td>
<td>0.67</td>
<td>100.00</td>
</tr>
<tr>
<td>$90</td>
<td>24.39</td>
<td>2.65</td>
<td>1.15</td>
<td>0.22</td>
<td>0.64</td>
<td>100.00</td>
</tr>
<tr>
<td>$100</td>
<td>25.85</td>
<td>2.50</td>
<td>0.85</td>
<td>0.62</td>
<td>0.71</td>
<td>100.00</td>
</tr>
<tr>
<td>$110</td>
<td>21.72</td>
<td>5.90</td>
<td>5.84</td>
<td>0.29</td>
<td>1.20</td>
<td>100.00</td>
</tr>
<tr>
<td>$120</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>$130</td>
<td>26.35</td>
<td>5.25</td>
<td>4.09</td>
<td>1.61</td>
<td>1.50</td>
<td>100.00</td>
</tr>
</tbody>
</table>

It is critical to link DR-1 resources to system demand and periods when resources are needed most. Historically, high market prices have not correlated strongly with system demand. Many monthly system peak hours have not been eligible for DR-1 participation. Instead of reporting reductions at monthly system peaks, reductions are reported as averages over top system hours for the year. These are the hours in which DR-1 resources are most likely to be needed. The statistics in Table 1-2 paint a more accurate picture of expected load reductions for the DR-1 program on an ex ante basis. They factor in the likelihood that the hour would be eligible for a DR-1 event under the current trigger mechanism and bid patterns of remaining participants. The reductions may be different if the trigger mechanism is modified. For DR-1, program reductions are highest during the top 50 system hours and decrease as additional hours are included. The ex ante outlook for DR-1 given current participants is under 1 MW of load reduction capability. Based on the top 50 system load hours, DR-1 ex ante load reduction capability is 0.23 MW. Though the program reductions are currently small, the program has served its intended purpose: to help participants better understand load reduction capabilities and eventually transfer to more contractual DR programs.
Table 1-2:
DR-1 *Ex ante* Reductions over Top System Load Hours (Remaining Participants)

<table>
<thead>
<tr>
<th>IESO Top System Load Hours</th>
<th>Average Reference Load (MW)</th>
<th>Average Bid Amount (MW)</th>
<th>Average Settlement Load Reduction (MW)</th>
<th>Average Regression Based Load Reduction (MW)</th>
<th>Average Number of Participating Customers</th>
<th>Percent Hours Eligible by Price Conditions (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>26.24</td>
<td>0.92</td>
<td>0.44</td>
<td>0.23</td>
<td>0.22</td>
<td>0.26</td>
</tr>
<tr>
<td>100</td>
<td>23.12</td>
<td>0.88</td>
<td>0.55</td>
<td>0.19</td>
<td>0.23</td>
<td>0.26</td>
</tr>
<tr>
<td>200</td>
<td>21.70</td>
<td>1.04</td>
<td>0.67</td>
<td>0.14</td>
<td>0.25</td>
<td>0.26</td>
</tr>
<tr>
<td>500</td>
<td>21.08</td>
<td>0.80</td>
<td>0.42</td>
<td>0.06</td>
<td>0.18</td>
<td>0.18</td>
</tr>
</tbody>
</table>

It is important to keep in mind the role of DR-1 in OPA's demand response portfolio. Load reductions from demand buy-back programs are volatile, though not unpredictable. In almost all jurisdictions where such programs are offered, customer bid amounts vary substantially and the load reduction delivered is substantially less than the bid amounts. Despite their flaws, bidding programs can have a place in a portfolio of DR programs by providing a no or low-risk option that allows participants to test their load response capabilities or learn the extent to which they can provide response. The goal, however, should be to move those customers to more committed options.

DR-1 has been successful at helping commercial and industrial facilities become familiar with DR and develop load response strategies. A number of the customers enrolled in DR-1 have since migrated to more committed DR programs (i.e., DR-2 and DR-3), which are now being offered by OPA. Those customers currently provide predictable and reliable load reductions for operation. Although the transition of large participants has substantially reduced the DR-1 load reduction capacity, DR-1 will continue to operate through the end of 2010. There is significant value in its ability to attract customers new to DR who might otherwise be too risk-averse to try DR-2 or DR-3 initially, which entail penalties for non-performance, without having DR-1 as a no-risk training ground.

However, there are multiple opportunities to improve the program and its operations, including:

- The program’s main goal should be to provide an opportunity for facilities new to DR to test their load reductions capability and eventually move to a more committed option or stop participation. Hap-hazard participation does not benefit OPA or participants. As a result, we recommend limiting the amount of time participants can remain on a demand buy-back program to a year, or at most, two years. Doing so forces participants to decide whether to progress to more committed types of DR or not. In light of that role, technical
assistance to help customers identify discretionary load and ways to reduce it quickly is recommended;

- Custom baselines have a place in the program, particularly for very large participants, but it is critical to conduct a robust and detailed assessment of baseline accuracy and select the baseline that minimizes error;

- We recommend a requirement of load reductions larger than 0.5 MW and greater than 15% of the average summer peak period load. OPA currently only has a megawatt load reduction threshold. The percent load reduction component is critical in that settlement baseline errors are magnified in the baseline load reductions and settlement payments. The smaller the percent load reduction, the more the baseline error affects load reductions estimated for settlement and payments;

- We recommend introducing a requirement that participant’s reductions exceed 50% of their bid amount to be eligible for payment. Currently, settlement risk is asymmetrical. Even when error for a baseline is, on average, zero, baseline errors occur for specific customers and for specific event days. OPA may potentially pay participants that do not engage in load reductions but benefit from settlement error. The threshold reduces the potential for such errors;

- The program can and should revise the trigger mechanism to restrict the number of event hours to less than 300. Doing so signals to participants that their participation is indeed critical to the system and avoids exhausting participants; and

- DR-1 needs to improve the alignment between the dispatch of the program, system needs, and the dispatch of other DR programs. Several viable options include the use of generator heat rates or the IESO’s supply cushion. Heat rates provide an advantage in that they are not linked to natural gas prices; rather, they are linked to how much of the generation supply stack is in operation. The IESO’s day ahead supply cushion estimates also provide a clear indication of the need for additional resources. When the supply cushion is small, relative to demand, there is a higher likelihood of system disturbances that could endanger the operations of the electricity grid.

1.2 DR-2 Summary of Results

DR-2 is a contractual load shifting program in which participants specify the load shift window and amount of load shifting. The program does not provide compensation for changes in the overall level of electricity consumption that may be associated with economic fluctuations or other factors. Load reduction during peak periods must be accompanied by shifting to off-peak periods to receive payment.

DR-2 resources are not dispatchable since the program is not event based. The program initially started in 2009. It was in a transition phase from March through October. The transition allowed OPA and participants to determine whether the initial program rules were effective or required modification. They also allowed participants to test their load shifting capabilities and adjust the amount of load shifted and the peak participation hours to reflect their capabilities.
Throughout the transition phase payment offsets were not applied. The program rules were modified in October 2009 and the program officially launched in November 2009.

Importantly, the experience with DR-2 program performance in 2009 is very limited. Any conclusions should be drawn with caution until a full year of experience with the program has been analyzed. FSC believes that the results for the transition period are not representative of future performance. There are several reasons for this:

- Payment offsets and penalties were not enforced during the transition period;
- Participants did not have the final baseline data;
- Participants obtained a better understanding of their load shifting capability during the transition period; and
- Participation and settlement rules were modified for the final contract period.

As a result, the *ex ante* load reductions are entirely based on the contract period, even though it was limited to November and December of 2009. Given their preliminary nature, the findings and recommendation should be used with caution.

*Ex ante* load reductions are most appropriate for demand response (DR) programs since they are designed to provide an insurance or option-based estimate of load reductions that can be expected under normal and extreme conditions (typically weather or electricity price related). DR programs are designed to help address electricity system capacity issues when the system is under the most stress. *Ex ante* load reductions are based on the amount of contractual load response and the degree to which participants have historically delivered those load reductions. For the *ex ante* load reductions, we factored in the scheduled non-performance rate derived from historical facility shut down periods in 2006–2009 and the extent to which load reductions deviated from the contracted load reduction on an hourly basis.\(^3\)

Table 1-3 shows the hourly *ex ante* load reduction estimates for the DR-2 program.

\(^3\) During DR-2 transition and enrollment periods, participants shut down the main processes at their facilities in over 90% of scheduled non-performance days. In addition, almost no shutdowns occurred outside of scheduled non-performance. As a result, facility shut downs are used as a proxy for non-performance during the 2006–2008 time period allowing *ex ante* estimates to be de-rated in a way that incorporates appropriate historical seasonal variation.
<table>
<thead>
<tr>
<th>Season</th>
<th>Hour Ending</th>
<th>Contracted MW</th>
<th>Day Ahead Contracted MW (Scheduled)</th>
<th>Ex ante Load Reduction Estimate</th>
<th>90% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summer Months</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8:00 AM</td>
<td>106</td>
<td>102.8</td>
<td>118.6</td>
<td>116.9</td>
<td>120.4</td>
</tr>
<tr>
<td>9:00 AM</td>
<td>106</td>
<td>91.6</td>
<td>92.5</td>
<td>91.3</td>
<td>93.7</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>106</td>
<td>91.6</td>
<td>87.3</td>
<td>86.3</td>
<td>88.2</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>106</td>
<td>91.6</td>
<td>82.3</td>
<td>81.5</td>
<td>83.1</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>106</td>
<td>91.6</td>
<td>87.7</td>
<td>87.1</td>
<td>88.3</td>
</tr>
<tr>
<td>1:00 PM</td>
<td>106</td>
<td>91.6</td>
<td>77.1</td>
<td>76.7</td>
<td>77.5</td>
</tr>
<tr>
<td>2:00 PM</td>
<td>119</td>
<td>104.0</td>
<td>91.8</td>
<td>91.4</td>
<td>92.2</td>
</tr>
<tr>
<td>3:00 PM</td>
<td>119</td>
<td>104.0</td>
<td>94.3</td>
<td>93.9</td>
<td>94.6</td>
</tr>
<tr>
<td>4:00 PM</td>
<td>119</td>
<td>94.8</td>
<td>94.1</td>
<td>90.1</td>
<td>98.2</td>
</tr>
<tr>
<td>5:00 PM</td>
<td>65</td>
<td>51.8</td>
<td>57.9</td>
<td>45.7</td>
<td>70.2</td>
</tr>
<tr>
<td>6:00 PM</td>
<td>65</td>
<td>49.9</td>
<td>49.6</td>
<td>39.1</td>
<td>60.1</td>
</tr>
<tr>
<td>7:00 PM</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>AVERAGE</strong></td>
<td><strong>102.1</strong></td>
<td><strong>87.8</strong></td>
<td><strong>84.8</strong></td>
<td><strong>81.4</strong></td>
<td><strong>88.1</strong></td>
</tr>
<tr>
<td>Winter and Shoulder Months</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8:00 AM</td>
<td>106</td>
<td>102.5</td>
<td>118.2</td>
<td>116.5</td>
<td>120.0</td>
</tr>
<tr>
<td>9:00 AM</td>
<td>106</td>
<td>92.7</td>
<td>93.6</td>
<td>92.3</td>
<td>94.8</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>106</td>
<td>92.7</td>
<td>88.3</td>
<td>87.4</td>
<td>89.3</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>106</td>
<td>92.7</td>
<td>83.3</td>
<td>82.5</td>
<td>84.0</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>106</td>
<td>92.7</td>
<td>88.8</td>
<td>88.1</td>
<td>89.4</td>
</tr>
<tr>
<td>1:00 PM</td>
<td>119</td>
<td>105.2</td>
<td>88.6</td>
<td>88.1</td>
<td>89.0</td>
</tr>
<tr>
<td>2:00 PM</td>
<td>119</td>
<td>105.2</td>
<td>92.9</td>
<td>92.5</td>
<td>93.3</td>
</tr>
<tr>
<td>3:00 PM</td>
<td>119</td>
<td>105.2</td>
<td>95.4</td>
<td>95.0</td>
<td>95.8</td>
</tr>
<tr>
<td>4:00 PM</td>
<td>65</td>
<td>52.8</td>
<td>52.4</td>
<td>50.2</td>
<td>54.7</td>
</tr>
<tr>
<td>5:00 PM</td>
<td>65</td>
<td>50.7</td>
<td>56.7</td>
<td>44.7</td>
<td>68.7</td>
</tr>
<tr>
<td>6:00 PM</td>
<td>52</td>
<td>40.5</td>
<td>40.2</td>
<td>31.7</td>
<td>48.8</td>
</tr>
<tr>
<td>7:00 PM</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>AVERAGE</strong></td>
<td><strong>97.2</strong></td>
<td><strong>84.8</strong></td>
<td><strong>81.9</strong></td>
<td><strong>78.6</strong></td>
<td><strong>85.2</strong></td>
</tr>
</tbody>
</table>

Several observations from the *ex ante* load reductions are noteworthy:

- For the average hour, the *ex ante* load reduction estimate is 84.8 MW in the summer and 81.9 MW in the winter and shoulder months;
- During the summer months, on average, 83.1% of the contractual resources (84.8 MW out of 102.1 MW) can be expected for planning purposes;
- During winter and shoulder months, 84.3% of the contractual resources (81.9 MW out of 97.2 MW) can be expected for planning purposes;
- On average, the DR-2 program is expected to deliver on 96.6% of the day-ahead contracted load reductions (81.9 MW out 84.8 MW), which factor in non-performance notifications;
- Reductions are not differentiated by month because the program is contractual and there is no evidence as of yet to assess if seasonal effects exist;
- Load reduction estimates are the same for extreme and normal conditions, reflecting the contractual nature of the program and the fact that participants are not weather sensitive; and
- The above results assume that a 52 MW participant who shut down its facility during the November and December 2010 resumes operations and performs in similar fashion as the other participants.

As noted earlier, the *ex ante* load reductions are based in part on the degree to which participants have historically delivered day-ahead scheduled resources. In order to calculate 2009 load reductions for DR-2, the participants load in the absence of load shifting—the counterfactual—was estimated using regression analysis. For most DR programs, this is accomplished by using a combination of the following:

- Pre-enrollment data;
- Behavior is repeatedly observed for similar days under both event and non-event conditions (a within customer control) ; or
- Through the use of a control group.

DR-2 is a unique program and presents several challenges from an evaluation perspective. Given the limited number of facilities of similar size and industry, it is also not possible to develop an adequate control group. It is also not possible to observe the naturally occurring facility behavior while customers are participating in DR-2 because the program is designed to produce load shifting on a daily basis. Unlike an event based program, behavior under event and non-event conditions cannot be repeatedly observed. Load shifting is in effect after a participant is enrolled. As a result, the evaluation relied on differences in pre- and post-enrollment behavior after controlling for differences in weather, market prices, and economic conditions.

The load usage patterns and load reductions were estimated using 2006 through 2009 hourly demand data for pre-enrollment, transition, and enrollment periods. Plant shut downs were excluded from the analysis and are factored through de-rating for the frequency of non-performance days.
Table 1-4 summarizes the hourly load reductions and the 90% confidence intervals for the two months when the program operated under final rules in 2009. The uncertainty associated with the load reductions estimates is relatively narrow.

Table 1-4: 2009 DR-2 *Ex post* Hourly Load Reductions and Confidence Intervals for Contractual Period

<table>
<thead>
<tr>
<th>Hour Ending</th>
<th>Contracted Load Shift</th>
<th>Day-Ahead Contracted Resources (Scheduled)</th>
<th>Load Reductions</th>
<th>90% Confidence Interval</th>
<th>% of Scheduled</th>
<th>90% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
<td>Upper</td>
<td>Lower</td>
</tr>
<tr>
<td>8:00 AM</td>
<td>106.0</td>
<td>54.0</td>
<td>64.4</td>
<td>63.4</td>
<td>65.3</td>
<td>119%</td>
</tr>
<tr>
<td>9:00 AM</td>
<td>106.0</td>
<td>54.0</td>
<td>58.7</td>
<td>57.9</td>
<td>59.4</td>
<td>109%</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>106.0</td>
<td>54.0</td>
<td>54.6</td>
<td>54.0</td>
<td>55.2</td>
<td>101%</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>106.0</td>
<td>54.0</td>
<td>51.6</td>
<td>51.1</td>
<td>52.1</td>
<td>96%</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>106.0</td>
<td>54.0</td>
<td>52.6</td>
<td>52.2</td>
<td>53.0</td>
<td>97%</td>
</tr>
<tr>
<td>1:00 PM</td>
<td>119.0</td>
<td>67.0</td>
<td>56.6</td>
<td>56.3</td>
<td>56.9</td>
<td>84%</td>
</tr>
<tr>
<td>2:00 PM</td>
<td>119.0</td>
<td>67.0</td>
<td>58.7</td>
<td>58.5</td>
<td>58.9</td>
<td>88%</td>
</tr>
<tr>
<td>3:00 PM</td>
<td>119.0</td>
<td>67.0</td>
<td>61.0</td>
<td>60.7</td>
<td>61.2</td>
<td>91%</td>
</tr>
<tr>
<td>4:00 PM</td>
<td>65.0</td>
<td>13.0</td>
<td>34.3</td>
<td>32.8</td>
<td>35.7</td>
<td>263%</td>
</tr>
<tr>
<td>5:00 PM</td>
<td>65.0</td>
<td>13.0</td>
<td>14.7</td>
<td>11.6</td>
<td>17.8</td>
<td>113%</td>
</tr>
<tr>
<td>6:00 PM</td>
<td>52.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7:00 PM</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>97.2</strong></td>
<td><strong>49.7</strong></td>
<td><strong>50.7</strong></td>
<td><strong>48.7</strong></td>
<td><strong>52.7</strong></td>
<td><strong>102%</strong></td>
</tr>
</tbody>
</table>

Given the limited history of the program, the findings and recommendations for DR-2 are limited by necessity. A full year of program reductions under the contract rules is necessary before drawing stronger conclusion. Despite those limitations, the evaluation of DR-2 provides several insights:

- During the contract period, the DR-2 program delivered an average of 102% of the day-ahead contracted load reductions, which factor in non-performance notifications;
- In general, participants reduced more load than they shifted. The program reductions were not restricted to shifting;
- There was a substantial amount of scheduled non-performance. On average, over the transition period, 62.7% of the contracted load was scheduled for non-performance. Over the contractual period, 53.7% of the contracted load was scheduled for non-performance. Most of the scheduled non-performance is from one participant, but it reflects the concentration of the program and the lack of diversity of the current participants across industries. While the extent of non-performance is likely related to economic conditions.
and leads to conservative estimates for planning, it is the only empirical data available for this unique and new program;

- The load reductions of customers counteracted each other during the transition period because their event windows are not coordinated. One participant may be shifting load to an hour, while another may shifting load away from the same hour;

- Contractual load shifting amounts are lower for the critical summer mid to late afternoon hours. The highest amount of load shifting occurs between 9 AM and 12 PM;

- The program can better coordinate load shifting by increasing the incentives for the more critical hours (and decreasing them for less critical hours), narrowing the eligible hours, or requiring that load increases occur outside of the program window, not simply outside of the load shift window elected by the participant; and

- Participants delivered a higher share of the expected load reduction during the contract period. Their actions suggest the transition helped them better understand their load shifting capabilities. All participants adjusted both the load shift amount and window for the contract period.

1.3 DR-3 Summary of Results

DR-3 allows participants and aggregators to enter into contractual agreements for load reductions with OPA. Participants can choose to enroll directly with OPA, provided they meet minimum load reduction criteria, or can participate through an aggregator. In exchange for load reductions, the DR-3 program makes both availability (capacity) and energy payments. Unlike OPA’s demand buy-back program (DR-1), OPA can reduce payments if participants fail to provide the contractual load reduction or are unavailable to provide load reductions (the equivalent of a scheduled outage). For settlement purposes, compliance with the contracted load reductions is determined through day matching baseline methods.

DR-3 participants notify OPA and the IESO of any short-term fluctuations in load reduction capability due to facility maintenance or down time. These days are classified as non-performance days and are analogous to scheduled generator outages. They enable the IESO to better operate the system and schedule alternate resources for those days. Non-performance days lead to reductions in the participant payments. The payment reductions are higher if an event is called during a participant non-performance day. Unscheduled non-performance—failure to meet contractual obligations during events—leads to even larger payment reductions.

There are several reasons why the 2009 evaluation of the program is significant. The load reduction capability of the program and the participant mix has changed substantially. First, since December 2008, the contracted load reduction has nearly doubled from 86 MW to 170 MW. Second, the program has diversified and the role of aggregators has grown. While a few
large direct participants still have a substantial effect on the overall program results, the
program better reflects the benefits of aggregating across multiple loads, which generally
increases predictability. Third, all of the 2009 events took place during the summer. Since the
program is largely designed to provide capacity, it adds to the body of knowledge of whether the
contracted load reductions are realized. Fourth, the unique economic conditions of 2009 likely
reflect the upper bound of scheduled non-performance days.

*Ex ante* load reductions are most appropriate for demand response (DR) programs since they
are designed to provide an insurance- or option-based estimate of load reductions that can be
expected under normal and extreme conditions. *Ex ante* estimates for the DR-3 program factor
in both scheduled non-performance and the extent to which participants reliably delivered the
day-ahead scheduled resources.

Table 1-5 summarizes the contractual load reductions, the expected day ahead contracted
resources, and the *ex ante* load reductions (e.g., the load reductions to be employed for
program and system planning purposes).

<table>
<thead>
<tr>
<th>Month</th>
<th>Contracted MW</th>
<th>Day Ahead Contracted MW</th>
<th>Ex-ante Load Reductions (For Planning)</th>
<th>90% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower bound</td>
</tr>
<tr>
<td>January</td>
<td>167</td>
<td>153.1</td>
<td>128.2</td>
<td>93.1</td>
</tr>
<tr>
<td>February</td>
<td>167</td>
<td>153.1</td>
<td>128.2</td>
<td>93.1</td>
</tr>
<tr>
<td>March</td>
<td>167</td>
<td>153.1</td>
<td>128.1</td>
<td>93.1</td>
</tr>
<tr>
<td>April</td>
<td>167.2</td>
<td>153.2</td>
<td>128.3</td>
<td>93.2</td>
</tr>
<tr>
<td>May</td>
<td>167.4</td>
<td>153.4</td>
<td>128.4</td>
<td>93.3</td>
</tr>
<tr>
<td>June</td>
<td>168.2</td>
<td>154.2</td>
<td>129.1</td>
<td>93.8</td>
</tr>
<tr>
<td>July</td>
<td>168.2</td>
<td>154.2</td>
<td>129.1</td>
<td>93.8</td>
</tr>
<tr>
<td>August</td>
<td>168.2</td>
<td>154.2</td>
<td>129.1</td>
<td>93.8</td>
</tr>
<tr>
<td>September</td>
<td>168.1</td>
<td>154.1</td>
<td>129</td>
<td>93.7</td>
</tr>
<tr>
<td>October</td>
<td>167.5</td>
<td>153.5</td>
<td>128.5</td>
<td>93.4</td>
</tr>
<tr>
<td>November</td>
<td>167.4</td>
<td>153.5</td>
<td>128.5</td>
<td>93.3</td>
</tr>
<tr>
<td>December</td>
<td>167.1</td>
<td>153.1</td>
<td>128.2</td>
<td>93.1</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>167.5</strong></td>
<td><strong>153.6</strong></td>
<td><strong>128.6</strong></td>
<td><strong>93.4</strong></td>
</tr>
</tbody>
</table>
Highlights from the *ex ante* findings include:

- The contracted resources are relatively constant and reflect aggregator and direct participant estimates of the extent to which their load reduction capability varies. For the average month, the *ex ante* load reduction estimate is 128.6 MW. The lowest *ex ante* load reduction estimate is 128.1 MW (March) and the highest is 129.1 MW (August);

- The de-ratings associated with the long-term reductions for planning factor in scheduled non-performance, adjustments for settlement baseline error, and participant performance;

- Overall, 76.7% of the contractual resources can be expected for program planning purposes. There is 90% confidence that 55.7 to 97.7% of the DR-3 contractual resources are available for long-term system planning;

- On average, the DR-3 program is expected to deliver 83.7% of the day-ahead contracted load reductions, factoring in non-performance notifications; and

- The *ex ante* estimates assume scheduled non-performance does not have a seasonal pattern. This was determined after reviewing the historical data for the direct customers (who accounted for the bulk of scheduled non-performance and are more sensitive to economic cycles) in 2006–2008.

While the *ex ante* load reduction estimates factor in the scheduled non-performance and departures from the day-ahead resources committed for operations, they do not directly account for the timing and magnitude of load reductions on effective load carrying capacity. From a capacity perspective, resources available in the summer mid-afternoon hours, the hours of system peak, generally have higher value than resources in non-peaking hours. It is necessary to incorporate the extent to which DR-3 resources can provide capacity without increasing the likelihood of unserved electricity. This is typically done for cost-effectiveness analysis and long-term planning.

Table 1-6 shows the average hourly load reductions for each of the event days called in 2009 and the 90% confidence bands for the estimates.
Table 1-6
DR-3 Program 2009 Load Reductions by Event Day

<table>
<thead>
<tr>
<th>Event Date</th>
<th>Group</th>
<th>Contracted MW</th>
<th>Day-Ahead Contracted MW</th>
<th>Estimated Load Reduction (MW)</th>
<th>90% Confidence Interval</th>
<th>% of Day-Ahead Contracted MW</th>
<th>90% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
<td>Upper</td>
<td>Lower</td>
</tr>
<tr>
<td>6/25/2009</td>
<td>200 only</td>
<td>9.8</td>
<td>9.8</td>
<td>8.6</td>
<td>7.6</td>
<td>9.6</td>
<td>87.7%</td>
</tr>
<tr>
<td>8/14/2009</td>
<td>100 and 200</td>
<td>145.8</td>
<td>50.6</td>
<td>46.2</td>
<td>31.8</td>
<td>60.6</td>
<td>91.2%</td>
</tr>
<tr>
<td>8/17/2009</td>
<td>100 and 200</td>
<td>145.8</td>
<td>84.3</td>
<td>77.1</td>
<td>60.3</td>
<td>93.9</td>
<td>91.5%</td>
</tr>
<tr>
<td>9/9/2009</td>
<td>200 only</td>
<td>20.1</td>
<td>20.1</td>
<td>13.6</td>
<td>11.8</td>
<td>15.4</td>
<td>67.6%</td>
</tr>
<tr>
<td>9/10/2009</td>
<td>100 and 200</td>
<td>141.0</td>
<td>129.9</td>
<td>94.9</td>
<td>71.0</td>
<td>118.9</td>
<td>73.1%</td>
</tr>
<tr>
<td>9/14/2009</td>
<td>100 and 200</td>
<td>141.0</td>
<td>129.9</td>
<td>107.2</td>
<td>83.6</td>
<td>130.8</td>
<td>82.5%</td>
</tr>
<tr>
<td>Average Event (2009)</td>
<td></td>
<td>100.6</td>
<td>70.8</td>
<td>57.9</td>
<td>51.2</td>
<td>64.6</td>
<td>81.8%</td>
</tr>
</tbody>
</table>

For the average 2009 event, the load reduction estimates are roughly 58% of contracted MW. The capacity contracted for the average event was 100.6 MW while the estimated load reductions were 57.9 MW.

The three main factors that explain the difference between the average contracted MW and estimated load reductions (42.7 MW) are as follows:

- **Scheduled non-performance.** Scheduled non-performance was relatively high in comparison to historical patterns in 2009 in part due to the effect of the economic downturn on direct participants. Overall, the scheduled non-performance explains 70% of the gap between contracted load reduction and load reduction estimates (29.8 MW out of 42.7 MW). Importantly, the scheduled non-performance was factored into operations;

- **Settlement baseline bias.** From the 2008 evaluation, FSC found a 6.9% program level bias in the settlement baseline, which translated into a 20% bias in settlement baseline load reduction estimates. Assuming the upward bias is similar in 2009, the baseline method error explains 27% of the gap between contracted load reduction and load reduction estimates (11.6 MW out 42.7 MW); and

- **Participant non-performance.** The remaining unexplained portion of the gap is less than 2% of the contracted load reduction and is smaller than the uncertainty of the load reduction estimates. It is potentially due to participant non-performance, though this cannot be fully determined.
The 2009 DR-3 evaluation adds a substantial amount of information about participant performance in terms of system operations and long-term planning. Despite the wealth of additional information, any conclusions drawn need to factor in the unique 2009 economic conditions and the fact that the customer mix will likely continue to diversify.

While DR-3 resources have diversified, they remain highly concentrated, with five participants accounting for approximately 42% of the contracted load reduction. There was a substantial amount of scheduled non-performance days in 2009, and the direct participants accounted for the bulk of it. For the average weekday, 9.6% of the contracted MW was scheduled for non-performance. That being said, participants consistently delivered the load reduction expected by the system operator. Once scheduled non-performance is accounted for, on average, participants delivered 81.2% of the day ahead contracted resources in 2009. Factoring in 2008, participants delivered 84%. In 2008, the gap between the actual load reductions and day ahead contracted resources was entirely explained by settlement baseline error and the fact that many participants manage load based on the baseline. We are confident this remains the case in 2009.

Key recommendations resulting from the evaluation include:

- Update the analysis of settlement baseline accuracy and extend it to model the effect of baseline errors on settlement payments;
- Until the baseline error is corrected, adjust day-ahead commitments to account for the settlement baseline error;
- Ensure direct participants fully understand their load reduction capabilities and usage patterns, particularly on-peak load, prior to engaging in contractual load reduction commitments;
- Update estimates used for operation and planning with additional data on DR-3 events. Specifically, refine estimates of the share of expected load reduction participants deliver and also refine estimates of scheduled non-performance days;
- Further diversify the program; doing so will not only will make the program more robust to economic downturns, it will also help reduce the statistical uncertainty and provide more precise reduction estimates; and
- Consider transitioning large direct participants to another program that relies less so or not at all on settlement baselines. Throughout North America, most programs and contracts that provide availability payment either operate as an interruptible program wherein participants reduce loads to a pre-specified level or, alternatively, rely substantially on aggregation to reduce volatility in operational reliability.
2 INTRODUCTION

This report contains *ex post* and *ex ante* load reduction estimates for the Ontario Power Authority’s DR-1, DR-2, and DR-3 programs. All three programs target commercial and industrial customers; combined, they represent the Demand Response C&I portfolio. The three programs have substantial differences in their design.

DR-1 is a voluntary demand response (DR) program offered by the OPA to pilot the concept of DR among large industrial and commercial customers. It is a demand-bidding program in which participants are paid for reducing load when Independent Electricity System Operator (IESO) market prices exceed a strike price proposed by participants.

DR-2 is a load shifting program in which participants contract to shift specific amount of load reduction from peak hours. Participants have the ability to specify the amount of load and hours of load shifting on a monthly basis. The program only compensates customers for shifting load from peak hours and does not provide payment for overall reductions in energy consumption that may be associated with economic conditions.

DR-3 is a contractual demand response (DR) program that offers large commercial and industrial consumers reserve payments for being available to provide load reductions when called by the OPA. Participants are required to meet a 95% reliability requirement when they are called, or else financial penalties apply. Table 2-1 (next page) compares some of the key program design features for each program.
<table>
<thead>
<tr>
<th>Program Feature</th>
<th>DR-1</th>
<th>DR-2</th>
<th>DR-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Program Description Summary</td>
<td>A voluntary event based demand buy-back program triggered by market prices. Enrollment in the program is voluntary, as is the decision of whether and how much to bid for a specific eligible period. The program is designed to allow commercial and industrial facilities test their load reduction capabilities without risk.</td>
<td>A load shifting program in which participant contract to shift specific amount of load reduction from peak hours. Participants have the ability to specify the amount of load and hours of load shifting on a monthly basis.</td>
<td>A contractual DR program that offers aggregators or direct participants availability payments for being available to provide load reductions when called by the OPA.</td>
</tr>
<tr>
<td>Amount of Contracted Resources</td>
<td>Contracts do not apply to DR-1. In 2009, average load reduction during eligible hours was 37 MW.</td>
<td>119 MW of load reduction.</td>
<td>170.5 contracted MW in 2009.</td>
</tr>
<tr>
<td>Number and concentration of customers</td>
<td>20 enrolled participants, of which 7 were active participants in 2010. Participants are smaller relative to DR-1 and DR-2.</td>
<td>3 very large participants, with 119 MW of contracted load shifting (153 MW during transition).</td>
<td>121 contributors. 5 accounts for 54% of the committed load reduction.</td>
</tr>
<tr>
<td>Event Frequency</td>
<td>132 Event hours in 2009. Most in the winter. Events were fewer in 2009 due to lower market prices.</td>
<td>DR-2 is a permanent load shift program operating on all non-holiday weekdays. A transition phase lasted from March to October. November and December operated under final event rules.</td>
<td>A total of 6 events were called in 2009. The 100-hour commitment group was called 4 times and the 200-hour commitment group was called 6 times. All 2009 events occurred during the months of June through September.</td>
</tr>
<tr>
<td>2009 Program Growth</td>
<td>Significant participant attrition, with enrolled participants dropping from 29 in 2008 to 7 in 2009.</td>
<td>DR-2 is a new program and saw its first 3 participants join in 2009—first for a transition period, then fully enrolled by November.</td>
<td>DR-3 experienced significant growth during 2009.</td>
</tr>
<tr>
<td>Non-performance rules</td>
<td>None—DR-1 is a voluntary program with no penalties for non-performance.</td>
<td>Participants subject to set-offs for non-performance at the hourly, daily, and monthly levels. Set-off charges are typically equivalent to payments for performance. Set-offs may be mitigated through advanced warning by participants.</td>
<td>Participants incur penalties if they provide less than 85% of their contracted MW amount relative to the baseline.</td>
</tr>
<tr>
<td>Length of program experience</td>
<td>Program has been active since 2007.</td>
<td>Program is new in 2009.</td>
<td>Program has been active since 2008.</td>
</tr>
<tr>
<td>Settlement baselines</td>
<td>Same hour average of past 10 non-event weekdays.</td>
<td>Baselines based on pre-contract period and adjusts for change in consumption.</td>
<td>Top 15 out of 20 past days, excluding non-performance days.</td>
</tr>
</tbody>
</table>
2.1 Evaluation Objectives and Key Research Questions

This report focuses on the evaluation of 2009 *ex post* and *ex ante* load reductions. The 2009 program evaluation of these programs is designed to meet multiple objectives, including:

- Develop *ex post* load reduction estimates for each program for 2009;
- Provide estimates of the load reduction capability (i.e., *ex ante* load reductions) for each program for planning purposes;
- Provide estimates of the operational reliability of each DR program;
- Develop inputs for conducting cost-effectiveness analysis of the programs;
- Provide recommendations for improvement in program design and operations; and
- Comply with OPA Load Reduction Protocols\(^4\) that guide program evaluations for DR resources in Ontario.

Due to time constraints, the 2009 evaluation does not include a detailed analysis of baseline settlement method accuracy. This will be provided in the more comprehensive 2009–2010 evaluation. Since it is an integral part of the programs—particularly DR-1 and DR-3—understanding baseline errors and their implications is critical for improving program operations. The 2009 time frame was marked by substantive changes in the programs and unique economic conditions. The 2009 economic downturn affected several industries, their production level—and by connection—their overall energy demand. Logically, the ability to provide demand response resources is tied to the amount of electricity demand by participants.

The program themselves underwent significant changes. The amount of resources in DR-1 decreased by more than 95% as 3 large customers transitioned to DR-2. In addition, market price conditions changed substantially in 2009 relative to prior years and the program is undergoing a full redesign. The DR-2 program was launched in 2009, first undergoing a transitional or test period from March to November 2009, before settling into the final program design and contractual level starting November 2009. DR-3 experienced substantial growth in the contracted load reduction, doubling from 86 to 170 MW during the year.

The substantial changes in the DR portfolio make the interim evaluation critical for understanding the DR resources, their load reduction capability, and their operational reliability. At the same time, any conclusions drawn need to factor in the unique 2009 economic conditions and the fact that some programs still have limited experience, or are undergoing redesign.

\(^4\) Ibid., p.4
2.2 How to Use Load Reduction Results

There are various ways in which how to use DR load reduction estimates. The estimation methods and output requirements may vary depending on how the estimates are used. The estimates presented here include ex post analysis for events called in 2009 and ex ante estimates of the load reduction capability. The analysis also provides useful input for improvements in program design and operation.

Figure 2-1, taken from OPA’s Load Reduction Protocols, provides an overview of the various ways in which load reduction estimates can be used and the interrelationship ex post and ex ante analysis. Ex post load reductions determine what happened over a historical period, based on the conditions that were in effect during that time. Because historical performance is tied to past conditions such as weather, price levels, and dispatch strategy, ex post load reductions may not reflect the full option value of a DR resource. For example, during several event days in 2009, the full load reduction capability of the DR-3 program was not utilized because it was not needed. As such, it would be inappropriate to use ex post reductions to determine DR program cost-effectiveness, because cost effectiveness should be based on the option value of the resource and its impact on overall resource needs.

Ex ante load reductions are forward-looking and are designed to reflect the load reduction capability of a DR resource under a standard set of conditions that match the market and system conditions that drive the need for additional capacity. Typically, ex ante load reductions are based on regression models developed from DR program data and customer behavior (that is, ex post analysis), but the models allow for adjustments to be made in the estimates to reflect the appropriate, ex ante event conditions. After a program matures and the participant mix becomes stable, ex ante load reductions can also be used to estimate future expected load reductions. Ex ante load reductions are an important input to DR cost-effectiveness analysis, both for program planning (comparing different DR program designs) and resource planning (comparing archetypal DR options against other conservation and supply resources).

In the analysis presented in this report, several constraints limited the robustness of the ex ante load reductions. As indicated in the draft load reduction protocols, ex ante estimates are to be provided for a common set of weather conditions, as many DR resources are weather sensitive and such resources have their greatest value under extreme weather conditions. However, in the case of DR-1, program events, participant bid behavior, and load reductions were largely driven by market prices and event frequency—not by weather—largely due to the significant

role of wholesale market participants. Thus, incorporating weather conditions without the corresponding market prices, system conditions, and event frequency patterns, would provide false precision.

The regression model estimates can also contribute to program operation. They can form the basis for developing dispatch models to predict, in near real time, the impact of DR programs as they are activated, provided the participant mix does not change substantially.

Although not required by the load impact protocols, this evaluation also provides useful analysis of the reliability of DR as a resource, both in the context of operations and long-term planning.

2.3 Reliability of Demand Response as a Resource

In reporting results, we have attempted to provide information that enables readers to assess the reliability of demand response resources. The term ‘reliability’ is often used loosely in discussing system operations and long-term planning. For clarity, throughout the report, we distinguish between operational reliability and capacity realization. The extent to which these
concepts apply is linked to whether or not a specific amount of DR resources is contracted. The
distinction is highly analogous to generator operations since a key goal is to enable
comparisons between generation and DR resources.

Operational reliability affects the ability to balance the grid on a short-term basis. The primary
determinant is whether the expected, scheduled resource is delivered when called upon. In the
context of generators, notification of scheduled outages or prolonged forced outages allows the
system operator to schedule additional resources for operations. In addition, the system
operators will have operating reserves available in case of unforeseen generator or
transmission outages. In context of contractual DR resources, operational reliability factors out
planned non-performance notification to the system operator. These are analogous to
scheduled outages in that participants are notifying the system operator not to expect or rely on
them as a resource for a specific period. The operational reliability is the share of the scheduled
resources that is indeed delivered. Variable DR resources such as DR-1 or air conditioner
cycling programs are similar to wind and solar generation in that the expected available
resources depend primarily on the conditions of a given day. The operational reliability is based
on the accuracy of the resource predictive models, how well the resource is incorporated into
operations, and whether the expected resources are indeed delivered.

In contrast, the capacity realization rate is designed to assess the extent in which the resources
can be factored into mid-term and long-term planning. For generation, the standard practice in
planning is to de-rate nameplate capacity for both scheduled and unscheduled outages. More
sophisticated approaches factor in the timing and likelihood of scheduled and forced outages.
For contractual DR resources, both planned and unplanned non-performance are factored into
the capacity realization rate and are used to de-rate the contracted load reduction resource.

In using *ex ante* load reduction estimates for planning purposes, it is critical to link DR resources
to system demand and other resources. While the *ex ante* load reduction estimates factor in the
scheduled non-performance and departures from the expected load reductions for operations,
they do not directly account for the timing and magnitude of load reductions. From a capacity
perspective, resources available in the summer mid-afternoon hours, which are the hours of
system peak, generally have higher value than resources in non-peaking hours. It is necessary
to estimate and incorporate the extent to which DR resources serve as capacity without
increasing the likelihood of unserved electricity—a value more commonly known as effective
load carrying capacity. This is typically done for cost-effectiveness analysis and
long-term planning.
2.4 Report Structure

This report is divided into four main sections that separately summarize the evaluation methodology and present results from the DR-1, DR-2, and DR-3 evaluations. The methodology section describes the regression model development, summarizes the evaluation approach employed for each program, and describes the check we conducted to ensure results are accurate and valid. More detailed discussions about the evaluation methodology for each program are presented as appendices. Each program section provides background information about the program, participant mix, and events called. This is followed by a discussion of the 2009 ex post load reduction results, followed by load reductions for operations. We then present the ex ante load impacts for long term planning and cost-effectiveness which incorporate adjustments for both historical scheduled non-performance and historical deviations from day-ahead committed resources. Each section concludes with a summary of the findings and recommendations.
3 EVALUATION METHODOLOGY

To calculate load reductions for demand response programs, the participant's load patterns in the absence of program participation—the counterfactual—must be estimated. For most DR programs, this is accomplished by using pre-enrollment data, observing behavior during event and non-event days (or within customer control), use of a control group, or a mixture of the aforementioned. The most rigorous method for impact evaluations is a well-executed experiment with random assignment to either a control or treatment condition. While randomized experiments are rarely feasible for actual programs, particularly when several participants are unique, there are multiple available methods for assessing reductions that approach the rigor of experiments. The best available method is a function of the program characteristics, available data, the ability to incorporate research design elements, and statistical methods.

In all instances, the evaluation used unperturbed load data prior to any program enrollment and post enrollment data. None of the evaluations used an outside control group. Instead, the customers were used as their own control group. DR-1, DR-2, and DR-3 each include large commercial customers and, for each, there are a limited number of facilities of similar size and industry, impairing the ability to draw an adequate control group. Because DR-1 and DR-3 are event based, it is possible to observe the naturally occurring facility behavior during non-event days even after customers enrolled in the program. DR-2 is not event based and instead is designed to produce load shifting on a daily basis. The approach employed for DR-2 is referred to as an interrupted time series—a quasi-experimental method. The key to an interrupted time series analysis is knowledge of the exact point at which an intervention occurred. An intervention can be an event day or a payment stimulating a change in behavior—in this case load shifting—after a specific point.

Regression methods were employed to analyze DR-1, DR-2, and DR-3 customer load patterns with or without event participation for load shifting. Regression techniques have several distinct advantages over the day-matching or baseline methods typically used for DR program settlement. First, they can help identify the key drivers and predictors of load patterns and load reductions. Second, regression results provide more robust estimates of load reductions and are not as sensitive to biases in the reference load. Third, they can be used to predict load reductions for operations or for long-term planning by factoring in expected weather, system, and market conditions, if available. In other words, they can better predict load reductions on a day-of or day-ahead basis in order to assist operations as well as predict load reduction resources under system peaking conditions in order to assist long-term planning.

With baseline methods, load reductions are calculated as the difference between the baseline method prediction and the metered load for an event day. As a result, bias in the settlement
baseline translates into larger biases in the estimated reductions. The smaller the percent load reduction the more the bias is magnified. For example, if the true event period load and load reductions are 100 MW and 20 MW, a 5% upward bias in the settlement baseline will lead to an estimated reference load of 105 MW and a calculated load reduction of 25 MW (105 MW minus the metered load of 80 MW). While the baseline upward bias is 5%, the baseline estimated reductions are biased upwards by 25%, 25 MW of load reduction was estimated rather than the actual 20 MW. If instead the true load reduction were lower, say 10 MW (10%), the settlement method would estimate a load reduction of 15 MW (105 MW minus the metered load of 90 MW). With the smaller load reduction, the 5% upward bias in the settlement baseline leads to a 50% upward bias in the estimated reductions, therefore 15 MW of load reduction was estimated rather than the actual 10 MW. Regression methods are not as sensitive to baseline error because reductions are determined by reduction coefficients. If the residuals from the regression are not correlated to the treatment, the results are accurate. For both regression and baseline methods, it is critical to conduct a number of validity tests to determine the ability of the evaluation approach to produce accurate results.

The remainder of this section details the process employed in regression model development, provides an overview of the analysis approach for DR-1, DR-2, and DR-3, and discusses the tests and checks undertaken to ensure reduction estimates are accurate. Appendices A, B, and C provide more detailed discussion of the methodology for each evaluation and present the results from the comparisons of actual and regression predicted loads.

3.1 Regression Model Development
Load reductions for OPA's demand response programs were analyzed using regression analysis. The load usage patterns and load reductions were estimated using 2006 through 2009 hourly demand. Regression methods vary depending on the requirements for ex ante load reduction estimation. Figure 3-1 illustrates various options with associated advantages and disadvantages.

Given that the typical customer enrolled in OPA's DR programs is a large process-driven C&I customer, individual customer regressions were employed rather than aggregate or panel models. The high concentration of load, and participation in the programs, was a significant factor in the selection of individual participant regressions. Changes in the number and mix of participants, as well as differences in the available data, made it difficult to analyze aggregated data. Moreover, individual participant regressions can better address variation in the size, production schedules, and processes of participants.
Given that the typical customer enrolled in OPA’s DR programs is a large process-driven C&I customer, individual customer regressions were employed rather than aggregate or panel models. The high concentration of load, and participation in the programs, was a significant factor in the selection of individual participant regressions. Changes in the number and mix of participants, as well as differences in the available data, made it difficult to analyze aggregated data. Moreover, individual participant regressions can better address variation in the size, production schedules, and processes of participants.

The regressions were developed with the primary goal of accurately estimating ex post load reductions and energy use patterns. The focus was primarily on prediction accuracy and on the robustness of the variables reflecting load reductions.

Specific regression techniques varied by program, and are discussed in detail in the associated appendices. Regardless of the technique chosen, the regression models can be thought of in terms of three main components:

- **Variables that reflect the average load shape of customers (load shape variables):** These are typically accounted by various shape indicator variables by hour of day and day of week.

- **Variables that explain deviation in hourly usage from the average load shape:** These include factors such as economic conditions, wholesale market prices, temperature, and seasonality.
- **Variables that estimate the hourly load reductions**: These variables are designed to estimate the hourly effects of the program participation during events. Event-based variables vary depending on the program type. For example, DR-1 load reductions are modeled as a function of participant bid amounts, DR-3 load reductions use variables based on contractual load reduction amounts, and DR-2 load shifts are estimated for each month and hour during weekdays.

### 3.2 DR-1 Analysis Approach Overview

The load reduction analysis estimated the incremental load reductions provided by participants in 2009 that were attributable to the DR-1 program. All participants already faced wholesale market prices for their electric commodity and, in general, were in the habit of decreasing their load during high priced periods in the absence of DR-1 incentives. While real, these price-induced load reductions should not be attributed to the DR-1 program.

Although the load reduction analysis explored the drivers of participant behavior and variation in load reductions, the nature of the program, and its participant mix posed several challenges to interpreting the results. The voluntary nature of the program required an analysis of load response patterns rather than a focus on a specific set of events. As a result, the regressions accurately reflect the average load response across events and capture systematic variation across different load response drivers such as market prices and hour of day. However, they do not and cannot accurately estimate the load reduction for each event or each event hour.

For DR-1, it is also difficult to develop estimates for day types representing *ex ante* event conditions. DR-1 is called when prices exceed a certain threshold. Historically, the frequency of eligible hours has varied from close to 2,000 hours per year to 132 hours in 2009. For such a program, defining normal and extreme year conditions solely based on weather would be inaccurate due to strong correlations between weather, market prices, event frequency, and participant load response. In short, there is insufficient market and event history to fully define *ex ante* load conditions for DR-1. Lastly, because load response for DR-1 is tied to event frequency, duration and avoided wholesale market costs, modifying the strike price would not only reduce the number of event hours but could fundamentally influence participation and overall load response. This is a particularly perplexing analysis problem for which there are no easy solutions.

Because of the variation in bids submitted by individual customers, the bid amounts were used as variables to predict the expected load reduction for a given event period. This method allowed the analysis to incorporate participant projections of load reduction and assess the extent to which participants exceeded or failed to meet those short-term estimates.
3.3 DR-2 Analysis Approach Overview

DR-2 is a unique program and presents several challenges from an evaluation perspective. In order to calculate load reductions for DR-2, the participants load in the absence of load shifting—the counterfactual—must be estimated. It is not possible to observe the naturally occurring facility behavior while customers are participating in DR-2 because the program is designed to produce load shifting on a daily basis. Given the limited number of facilities of similar size and industry, it is also not possible to develop an adequate control group.

The approach employed for DR-2 is referred to as an interrupted time series, which is a quasi-experimental method. The key to an interrupted time series analysis is knowing the exact point at which an intervention occurred. An intervention can be an event day or a payment stimulating a change in behavior—in this case load shifting—after a specific point. Part of the challenge for DR-2 is the multiple interventions to participant load that occurred since 2007. In 2007 and 2008, they participated in DR-1, which altered their behavior during event days. From March to October 2009, they were engaging in load shift during the transition period. For November and December of 2009, there was a different intervention in place; namely, the DR-2 full program rules were in effect and the settlement baseline calculations were updated to incorporate the DR-1 evaluation data. A strong point of this evaluation is that the timing of each intervention (i.e., shift in DR participation) is well documented. The downside is that disentangling the effect of each of the programs requires us to rely on historical participants that are distant from the evaluation period.

A careful analysis of the load data showed that a key difference between participant loads prior to and during DR-2 participation was the amount of days facilities were shut down for the day. While enrolled in DR-2, participants provided the IESO notice of non-performance days. These were typically, though not always, days in which the facility was shut down. For the pre-enrollment time frame, days when the facility was shut down can be identified by their load levels. In order to provide comparability across the changing economic conditions, days in which the facilities were shut down were excluded from DR-2 transition and enrollment periods as well as from the pre-enrollment data. Facility shut downs are factored into the long-term planning reduction estimates by de-rating, which is based on the historical frequency of non-performance days.

Variation in the demand for the three participants was step-like, with load levels clustered around specific values that reflect different industrial processes in operation. This creates a unique data pattern as the variation in demand is not continuous. The expected load level is best reflected by the probability that one, two, or multiple processes are operating jointly. As a result, a specific regression technique—ordinal probit regression—was applied for these customers. The approach essentially predicts the likelihood that the facility is at specific
demand levels. It also allows us to identify the most likely facility load level and produce the expected load by factoring the likelihood of each demand level.

The regressions provide the hourly reductions of the program, but the load shifting to off-peak periods does not always match the reduction during the peak period. This reflects the fact that some customers cannot fully shift their load reductions to off-peak periods due to the physical capacity of the plant. To obtain the amount of load shifting, we had to normalize the pre-enrollment, transition, and enrollment results. This process is simple and, in essence, subtracts out differences in overall consumption and allows us to focus solely on changes in the load shape. By construction, both the reference load shape and the transition period load shapes are directly comparable in terms of consumption.

3.4 DR-3 Analysis Approach Overview
The DR-3 load reductions were estimated through regression methods using data from event and non-event days and the available pre-enrollment data. Individual participant regressions were developed for each of the 121 participants in 2008 and 2009 and aggregated to the settlement account level and to the program as a whole.

Based on the program dispatch pattern, the DR-3 program naturally produces an alternating or repeated treatment design. The interventions—event notices—are introduced in some days and not in others, making it possible to observe behavior with and without events under similar conditions. A repeated treatment design enabled us to assess whether the outcome—electricity consumption—rises or falls with the presence or absence of the treatment, which is a program event. This approach works only if the effect of the event dissipates after it is removed. For evaluation, the entire day is evaluated to estimate both load reductions during event hours and load shifting to non-event hours. In addition, we conducted analysis to assess if event day effects spilled over into non-event days and determine if program stand-by notices led to changes in participant behavior.

A primary focus of the analysis was the actual load reductions relative to the contractual load reductions—the realization rate. The contractual load reductions did not vary for most participants. Comparing the event coefficients to contractual load reductions allowed the regressions to assess the extent to which participants exceeded or failed to meet contractual load reduction.

3.5 Assessment of Accuracy and Precision of Evaluation Models
In order to ensure that the regression-based load reductions were not an artifact of the final regression model, multiple regression specifications and techniques were tested and their effect on the reduction variables was carefully analyzed. For most participants, the coefficients of the
load reduction variables did not vary much in magnitude or statistical significance with the inclusion or exclusion of explanatory variables. In other words, the load reduction estimates were robust for each program.

In addition, regression-based estimates are rigorously validated to ensure their accuracy:

- The regression based predictions were compared to actual load under both event and non-event conditions in order to assess the ability of the models to accurately predict load patterns and reductions;

- Predicted and actual load are compared across temperatures, system conditions, and market prices to ensure the models predict accurately at all levels of the above. Specific attention was paid to performance under event-like conditions such as when system load or temperatures were high; and

- A commonly used metric for assessing precision of regressions is the R-squared statistic, which measures the degree to which the regression predictions improve on predictions based on a simple average. The R-squared metric measures the amount of variation around the mean explained by the regression.

Importantly, small differences in predicted values versus actual values will not affect the load reduction estimates, which are determined by the regression load reduction variable coefficients. The detailed analysis methodology and validation by specific program are referenced in the appendices.
4 DR-1 DEMAND BUY BACK PROGRAM

The DR-1 program is a voluntary demand buy-back program that allows industrial customers to bid in load reductions during hours when strike price conditions are met. Enrollment in the program is voluntary, as is the decision of whether and how much to bid for a specific eligible period. The program is designed to allow commercial and industrial facilities to test their load reduction capabilities without firm commitments or participant risk. The goal is to eventually move participants to other DR programs that require a firm commitment.

DR-1 concludes at the end of 2010 and a new program will replace it in 2011. The program is being redesigned to better align program events with system needs. In addition, the load reduction capability for the program has decreased by over 95% with the successful migration of the very large participants to DR-2, which is OPA's load shifting program.

Given the circumstances, 2010 will see marginal reductions from the few remaining customers who are active in DR-1. The evaluations are focused on answering the following key questions:

- What are the key drivers of participant bid behavior?
- What load reductions were realized during 2009?
- What are the load reductions that can be expected for high price and high system load hours?

4.1 Program Background

With the DR-1 program, strike price conditions are assessed on an hourly basis by comparing the three-hour ahead electricity wholesale market price forecasts to a predetermined monthly strike price. If participants elect to bid, they are eligible for payments at a minimum of three hours or during the time period in which market prices exceed the strike price. DR-1 is a no-risk program, meaning that participants do not face penalties for not participating during an event or for providing less load reduction than was bid into the market. By design, DR-1 allows participants to test their demand response capabilities and eventually move to programs that require permanent load shifting or firm load reduction commitments.

When participants enroll, they are asked to estimate their maximum load reduction capability. For 2009, the sum of maximum load reduction capabilities estimated by participants totaled 267.8 MW in January, 2009. By end of the year, the sum of the participant reported maximum load reduction capabilities was 175 MW. The aggregate load reduction bids never exceeded 200 MW for an eligible hour and varied substantially across eligible hours. On average, 111 MW of load reduction per eligible hour was bid into the program in 2009. The gap between the sum of participants' estimated load reduction capabilities and their bid amounts may be partly due to inaccurate estimates of load reduction capability by participants inexperienced with
demand response, but more importantly, it is also because participants do not jointly bid in their maximum load reduction capability. In other words, not all participants provide load reductions during all events and, for many participants, the load reduction amounts bid into DR-1 varied on an event-by-event basis.

4.2 Program Participation

Participation in DR-1 is only possible during eligible hours, defined as hours during which market prices exceed predetermined strike prices. 2009 saw a significant decrease in market prices across the entire year. Strike prices were adjusted downwards toward the end of the year. However, the number of eligible hours still fell substantially in 2009 as compared with previous years. 132, or roughly 1.5%, of hours in 2009 were eligible for DR-1 participation. Additionally, this price-based trigger does not align well with the province’s need for capacity. These circumstances have motivated OPA to redesign its voluntary demand response program. DR-1 concludes at the end of 2010 and a new program will replace it in 2011.

At its peak, the DR-1 program had 29 participants in 2008, with 21 actively bidding. In contrast, 20 customers remained enrolled in 2009 and 7 actively bid into the program. In March, the largest 3 of the 7 transitioned to DR-2, a load shift program with a contractual commitment and penalties for non-performance. As a result, the current DR-1 load reduction capability and performance is substantially different from when these customers were enrolled. These large customers are included in this ex post analysis to provide an accurate summary of 2009 DR-1 performance, but are not factored into ex ante estimates. After March of 2009, the program had a total of 4 active customers with a maximum load reduction capability of 22.5 MW and a maximum aggregated actual bid amount of 8 MW.

In 2009, bid amounts and frequency were highly concentrated among the 3 largest participants. Customers with the highest loads bid more often and bid a larger share of their load than other participants. In 2009, the 3 largest participants bid in load reductions for over 85% of eligible hours. In contrast, the remaining 4 active participants submitted load reduction bids for roughly 64% of the eligible hours. The 3 largest participants accounted for 98% of the load reduction bid into the program. Figure 4-1 shows the distribution among active participants of total load reductions bid in 2009.
4.3 2009 Events and Event Condition

Participation in DR-1 is only possible during eligible hours; defined as hours during which market prices exceed predetermined strike prices. 2009 saw a significant decrease in market prices across the entire year as compared to previous years; in part due to the slowdown in economic activity and more moderate climatic conditions. The 3-hour ahead pre-dispatch price, which is the metric that triggers DR-1, averaged roughly $63 from 2002–2008. In 2009, it fell by over half to an average of $31. Although strike prices were adjusted downwards towards the end of the year, the number of eligible event hours fell in 2009 as compared with previous years. In total, 132 or roughly 1.5% of hours in 2009 were eligible for DR-1 participation. Additionally, this price-based trigger does not align well with the province’s need for capacity, which is based on system load levels.

The sharpest decline in market prices came towards the end of winter 2009 and strike prices were not adjusted downwards until near the end of the year. This resulted in a distribution of eligible hours across the months as follows, and is a main driver of the monthly load reduction pattern.

![Figure 4-1: Distribution of Total Load Reduction Bid in 2009 (MWh)](image)
Eligible hours fell largely during the morning and evening hours. This reflects the wholesale market price patterns from January through March when the program was most active.

### 4.4 Participation Bidding Behavior

DR-1 participants are not required to reduce load for each event nor are they required to consistently deliver the same amount of load reduction. Those decisions are made by each participant on an event-by-event basis. As a result, there is substantial variation in both the level of active participation and the aggregate load reduction bid for an event. Whether or not participants deliver the load reduction bid in for a given event is addressed in the load reduction...
Table 4-1 shows actual bid amounts, average propensity to bid, and the product of the two, which is the average bid amount per eligible hour. The latter statistic is a useful way to view bidding behavior by participant and can be used as input into the load reduction model to specify explanatory variables relating to load reductions.

**Table 4-1:**
**DR-1 Participant Average Bidding Behavior During 2009**

<table>
<thead>
<tr>
<th>Participant</th>
<th>Actual Bid Amount (MW)</th>
<th>Average Propensity to Bid (percent)</th>
<th>Average Bid per Eligible Period (MW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large Participant 1</td>
<td>79.29</td>
<td>80.8%</td>
<td>64.04</td>
</tr>
<tr>
<td>Large Participant 2</td>
<td>10.14</td>
<td>97.4%</td>
<td>9.87</td>
</tr>
<tr>
<td>Large Participant 3</td>
<td>96.67</td>
<td>92.9%</td>
<td>89.76</td>
</tr>
<tr>
<td>Remaining Participant 1</td>
<td>5.00</td>
<td>80.8%</td>
<td>4.04</td>
</tr>
<tr>
<td>Remaining Participant 2</td>
<td>0.50</td>
<td>50.0%</td>
<td>0.25</td>
</tr>
<tr>
<td>Remaining Participant 3</td>
<td>1.29</td>
<td>50.0%</td>
<td>0.64</td>
</tr>
<tr>
<td>Remaining Participant 4</td>
<td>3.50</td>
<td>45.5%</td>
<td>1.59</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>196.37</strong></td>
<td><strong>86.7%</strong></td>
<td><strong>170.19</strong></td>
</tr>
</tbody>
</table>

Figure 4-4 shows the distribution of load reduction bids for all eligible hours in 2009. The analysis in this section explains some of the variation in event participation levels and is designed to help predict bid behavior to better operate the program. Importantly, some of the variability in DR-1 bid behavior was unique to 2009 and driven by the infrequency of eligible hours, uneven distribution across the months, and the loss of the program’s 3 largest participants.
Average bids per eligible hour were significantly higher during the first 3 months of the year, when the 3 large participants were still enrolled in DR-1. There is no bidding during the summer months due to minimal eligible hours during this time period. Bidding by the remaining active participants resumed in October through December, albeit at markedly lower amounts.

Table 4-2:
Average Program Bid per Eligible Hour by Month in 2009

<table>
<thead>
<tr>
<th>Month</th>
<th>Average Bid per Eligible Period (MW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>43.12</td>
</tr>
<tr>
<td>February</td>
<td>36.75</td>
</tr>
<tr>
<td>March</td>
<td>67.67</td>
</tr>
<tr>
<td>April</td>
<td>0.00</td>
</tr>
<tr>
<td>May</td>
<td>0.00</td>
</tr>
<tr>
<td>June</td>
<td>0.00</td>
</tr>
<tr>
<td>July</td>
<td>0.00</td>
</tr>
<tr>
<td>August</td>
<td>0.00</td>
</tr>
<tr>
<td>September</td>
<td>0.00</td>
</tr>
<tr>
<td>October</td>
<td>0.19</td>
</tr>
<tr>
<td>November</td>
<td>0.69</td>
</tr>
<tr>
<td>December</td>
<td>0.60</td>
</tr>
</tbody>
</table>
We analyzed customer bid decisions using regression methods. The high concentration of bids among the three largest participants was a significant factor in selecting the particular regression techniques used to evaluate bid behavior. These participants not only had larger loads, but were more likely to bid than smaller customers. Almost all of the large participants submitted bid amounts that varied by eligible period, with the bid amounts generally clustered around specific values that reflected industrial processes they were planning to shut off. Key factors that explained bid decisions and behavior include:

- Event characteristics such as duration, onset time, payment for load reductions, and avoided wholesale energy market costs;
- Variables reflecting market conditions such as hourly prices, price patterns by season, system load levels, and frequency of event conditions in prior days and weeks;
- Weather conditions, as captured by cooling and heating degree hours;
- Variables reflecting customer operating schedules, including day type and time-of-day characteristics; and
- Participant characteristics and participant-specific effects.

Appendix B provides further detail about the bid regression models developed along with validation information.

4.5 2009 Ex post Load Reduction Results

In 2009, the DR-1 program provided approximately 37 MW of load reduction per eligible hour and a total of approximately 7,778 MWh of energy savings. The average load reduction and energy savings are driven primarily by activity in the first quarter of 2009 prior to the migration of large participants from DR-1 to DR-2 and DR-3. They do not reflect DR-1 load reduction capability at the end of 2009. Load reductions per hour were slightly lower in 2009 than in 2008, when the program produced a load reduction of 39 MW per eligible hour. The program was operated far less frequently than in previous years due to falling market prices, on which the trigger mechanism to determine eligibility is based. Eligibility and participation in DR-1 were concentrated in the first and last three months of the year. For summer months, there were less than 10 eligible hours all together and there was no program participation.

Based on minimum strike prices, participants were eligible for load reductions during approximately 132 hours in 2009. When a participant submits a bid during an eligible hour, they are eligible to participate and receive compensation for load reductions for at least three consecutive hours, regardless of price eligibility in the two last hours. After the additional hours are factored in, they could participate in a total of 208 event hours.
Participants generally bid more load reduction than they reduced, both based on the settlement amounts and the regression analysis. Also, the regression based load reductions are lower than the estimates developed for settlement purposes using the baseline methodologies. Last year, FSC found significant bias in baseline calculations for participants on custom baselines, which influenced the settlement-based load reductions. The 3 largest participants used the same custom baselines for the first 3 months of 2009, before their migration to DR-2. Because these participants provided 95% of load reductions for the program, this is likely the cause for the discrepancy between settlement load reductions and regression-based load reductions.

After these participants left the program, there was no program participation through the summer months. Participation resumed on a small scale in October through the balance of the year. Figure 4-5 shows the distribution of load reductions across eligible hours in 2009.

Figure 4-5: Distribution of Load Reductions per Eligible Hour

There was substantial variation in the load reductions across eligible hours, but such variation does not mean that the reductions are unpredictable or unreliable. The bids and load reductions for individual hours vary systematically based on several factors, including forecasted market prices, event onset time, hour of day, and event duration. As a result, the regression models predict the expected load reduction for specific system and event characteristics with a high degree of accuracy.
As applicable according to the Load Impact Protocols⁶, DR-1 load reductions are reported for the following conditions:

- Reductions at key hours defined by IESO system load—to establish the coincidence of DR-1 reductions with times when they have the most system value;
- Maximum monthly reductions observed during 2009 DR-1 events—to provide an indication of the program potential—without regard for timing; and
- Reductions at various price thresholds—to demonstrate the amount of DR that is likely to occur at various market prices.

Table 4-3 presents load reductions for each IESO monthly system peak hour. It also shows the maximum load reductions achieved in each month. The results indicate that DR-1 reductions have historically not been consistent with IESO system peaks, which typically occur in the afternoon. Rather, the largest load reductions occurred in the early morning hours when the largest participants were more likely to be operating at full capacity.

⁶ Ibid., p. 4
### Table 4-3:
IESO Monthly System Peak Hour and Monthly Maximum Load Reductions For All 2007-2009 Participants

<table>
<thead>
<tr>
<th>Metric</th>
<th>Month</th>
<th>Average Reference Load (MW)</th>
<th>Average Expected MW (Bid Amount)</th>
<th>Average Regression Based Load Reduction (MW)</th>
<th>Load Reduction as % of Expected MW</th>
<th>Average Number of Participating Customers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>January</td>
<td>168.91</td>
<td>105.00</td>
<td>39.64</td>
<td>37.8%</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>February</td>
<td>115.75</td>
<td>110.00</td>
<td>37.15</td>
<td>33.8%</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>March</td>
<td>78.18</td>
<td>108.00</td>
<td>26.33</td>
<td>24.4%</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>April</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>May</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>June</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>July</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>August</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>September</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>October</td>
<td>20.49</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>November</td>
<td>27.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>December</td>
<td>2.45</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2009 IESO Monthly System Peak Days</td>
<td>January</td>
<td>252.84</td>
<td>195.00</td>
<td>97.58</td>
<td>50.0%</td>
<td>4.00</td>
</tr>
<tr>
<td></td>
<td>February</td>
<td>161.40</td>
<td>110.00</td>
<td>67.09</td>
<td>61.0%</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>March</td>
<td>115.07</td>
<td>100.00</td>
<td>59.40</td>
<td>59.4%</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>April</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>May</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>June</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>July</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>August</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>September</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>October</td>
<td>22.42</td>
<td>1.50</td>
<td>0.32</td>
<td>21.1%</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>November</td>
<td>24.12</td>
<td>2.00</td>
<td>1.93</td>
<td>96.3%</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>December</td>
<td>2.44</td>
<td>1.00</td>
<td>0.39</td>
<td>38.6%</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Comparing reductions during the first 3 months of the year to the last 3 months, the migration of the largest 3 participants to DR-2 has a profound effect on the load reduction potential of DR-1. The remaining participants are few and relatively small; with a limited amount of load that can be reduced.

Table 4-4 summarizes the load reductions attained at various market price thresholds for all active participants in 2009. Low load reductions at the lower price thresholds is a consequence of the fact that participation was only eligible at these prices during the last months of the year, when the 3 large participants had already migrated to DR-2. In general, participant load is reduced along with higher market prices. Part of the load reduction is due to participants' inherent price responsiveness and part of the load reduction is attributable to DR-1. As market prices rise, participants have added incentive to bid into the program, but they also have an incentive to reduce load independent of DR-1 to avoid the higher costs of electricity. Though load reductions in response to higher prices are very real, only the portion of this response that is incremental—i.e., over and above customers’ natural price response—can be attributed to the DR-1 program.

<table>
<thead>
<tr>
<th>IESO 3-Hour Ahead Price Threshold ($/MWh)</th>
<th>Average Reference Load (MW)</th>
<th>Average Expected Load (MW) (Bid Amount)</th>
<th>Average Settlement Load Reduction (MW)</th>
<th>Average Regression Based Load Reduction (MW)</th>
<th>Average Number of Participating Customers</th>
<th>Percent Hours Eligible by Price Conditions (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$50-$60</td>
<td>161.63</td>
<td>4.41</td>
<td>2.94</td>
<td>1.20</td>
<td>0.13</td>
<td>3.98</td>
</tr>
<tr>
<td>$60-$70</td>
<td>170.42</td>
<td>16.95</td>
<td>10.62</td>
<td>5.42</td>
<td>0.38</td>
<td>4.69</td>
</tr>
<tr>
<td>$70-$80</td>
<td>168.54</td>
<td>31.08</td>
<td>21.99</td>
<td>9.77</td>
<td>0.71</td>
<td>2.06</td>
</tr>
<tr>
<td>$80-$90</td>
<td>176.30</td>
<td>151.91</td>
<td>140.05</td>
<td>53.18</td>
<td>3.22</td>
<td>100.00</td>
</tr>
<tr>
<td>$90-$100</td>
<td>167.94</td>
<td>145.51</td>
<td>138.31</td>
<td>52.16</td>
<td>3.05</td>
<td>100.00</td>
</tr>
<tr>
<td>$100-$110</td>
<td>142.44</td>
<td>116.07</td>
<td>108.89</td>
<td>34.73</td>
<td>2.71</td>
<td>100.00</td>
</tr>
<tr>
<td>$110-$120</td>
<td>107.03</td>
<td>103.90</td>
<td>103.49</td>
<td>28.37</td>
<td>2.80</td>
<td>100.00</td>
</tr>
<tr>
<td>$120-$130</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$130</td>
<td>181.07</td>
<td>192.75</td>
<td>200.22</td>
<td>65.08</td>
<td>4.50</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Finally, an electronic file containing estimates of the hourly load reductions for each of the highest 10 system load days in which events occurred in 2009, along with load reductions for each of the monthly system peak days in which events occurred, will be provided to OPA.
hourly load reduction tables distinguish between reductions from all active participants and those active in the DR-1 program after 2009.

4.6 Load Reductions for Long-Term Planning and Cost-Effectiveness

The \textit{ex ante} load reductions are designed to provide an overview of load reductions that can be used for planning and cost-effectiveness analysis. Characteristics of the DR-1 program and its participants determine the format in which load reductions are reported. The participants are not weather sensitive and the voluntary nature of the program leads to a variable demand response resource.

As a result, the \textit{ex ante} load reductions, which only include participants still enrolled in the program at the year's end, are reported in two tables designed to reflect the capacity available from the DR-1 program. First, reductions are reported across market price thresholds. Table 4-5 shows the variation in DR-1 \textit{ex ante} reductions over the spectrum of market prices encountered in 2009. As program eligibility is determined by these prices, this is a key metric over which to view program reductions.

<table>
<thead>
<tr>
<th>IESO 3-Hour Market Threshold ($/MWh)</th>
<th>Average Reference Load (MW)</th>
<th>Average Bid Amount (MW)</th>
<th>Average Settlement Load Reduction (MW)</th>
<th>Average Regression Based Load Reduction (MW)</th>
<th>Average Number of Participating Customers</th>
<th>Percent Hours Eligible by Price Conditions (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$50</td>
<td>22.34</td>
<td>0.11</td>
<td>0.06</td>
<td>0.01</td>
<td>0.05</td>
<td>3.98</td>
</tr>
<tr>
<td>$60</td>
<td>24.31</td>
<td>0.39</td>
<td>0.17</td>
<td>0.03</td>
<td>0.10</td>
<td>4.69</td>
</tr>
<tr>
<td>$70</td>
<td>24.11</td>
<td>0.72</td>
<td>0.23</td>
<td>0.01</td>
<td>0.18</td>
<td>2.06</td>
</tr>
<tr>
<td>$80</td>
<td>25.07</td>
<td>3.20</td>
<td>1.98</td>
<td>0.22</td>
<td>0.67</td>
<td>100.00</td>
</tr>
<tr>
<td>$90</td>
<td>24.39</td>
<td>2.65</td>
<td>1.15</td>
<td>0.22</td>
<td>0.64</td>
<td>100.00</td>
</tr>
<tr>
<td>$100</td>
<td>25.85</td>
<td>2.50</td>
<td>0.85</td>
<td>0.62</td>
<td>0.71</td>
<td>100.00</td>
</tr>
<tr>
<td>$110</td>
<td>21.72</td>
<td>5.90</td>
<td>5.84</td>
<td>0.29</td>
<td>1.20</td>
<td>100.00</td>
</tr>
<tr>
<td>$120</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>$130</td>
<td>26.35</td>
<td>5.25</td>
<td>4.09</td>
<td>1.61</td>
<td>1.50</td>
<td>100.00</td>
</tr>
</tbody>
</table>

\textit{Ex ante} load reductions increase as market prices rise up until a certain point. Hours higher than roughly $100 are scarce, and as such, load reductions become more volatile. While Table 4-5 sheds light on the DR-1 resources available at different market prices, it is critical to link resources to system demand and other resources when resources are needed most.
Historically, high-market prices have not correlated strongly with system demand. For this reason, many monthly system peak hours have not been eligible for DR-1 participation. Instead of reporting reductions at monthly system peaks, reductions are reported as averages over top system hours for the year. These are the hours in which DR-1 resources are most likely to be needed. The statistics in Table 4-6 paint a more accurate picture of expected load reductions for the DR-1 program on an *ex ante* basis. They factor in the likelihood that the hour would be eligible for a DR-1 event under the current trigger mechanism and bid patterns of remaining active participants. The reductions may be different if the trigger mechanism is modified.

Regardless, we recommend using the load reductions during the top 50 hours as indicator of the program load reduction capability for planning and cost-effectiveness applications.

<table>
<thead>
<tr>
<th>IESO Top System Load Hours</th>
<th>Average Reference Load (MW)</th>
<th>Average Bid Amount (MW)</th>
<th>Average Settlement Load Reduction (MW)</th>
<th>Average Regression Based Load Reduction (MW)</th>
<th>Average Number of Participating Customers</th>
<th>Percent Hours Eligible by Price Conditions (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>26.24</td>
<td>0.92</td>
<td>0.44</td>
<td>0.23</td>
<td>0.22</td>
<td>0.26</td>
</tr>
<tr>
<td>100</td>
<td>23.12</td>
<td>0.88</td>
<td>0.55</td>
<td>0.19</td>
<td>0.23</td>
<td>0.26</td>
</tr>
<tr>
<td>200</td>
<td>21.70</td>
<td>1.04</td>
<td>0.67</td>
<td>0.14</td>
<td>0.25</td>
<td>0.26</td>
</tr>
<tr>
<td>500</td>
<td>21.08</td>
<td>0.80</td>
<td>0.42</td>
<td>0.06</td>
<td>0.18</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Reductions are minimal because the remaining active participants in DR-1 are relatively small, and furthermore, they bid a small percentage of their load and do so infrequently. DR-1 is designed as a feeder program and all 2009 attrition was due to participants transferring to DR-2. Though the program reductions are currently small, the program has served its intended purpose: to help participants better understand load reduction capabilities and eventually transfer to more contractual DR programs. For DR-1, program reductions are highest during the top 50 system hours and decrease as additional hours are included. The *ex ante* outlook for DR-1, based on the top 50 system load hours and given current active participants, is for 0.23 MW of load reduction capability, however, any conclusions should be drawn with caution because a program re-design is scheduled for 2011.

4.7 Conclusions and Recommendations

It is important to keep in mind the role of DR-1 in OPA's demand response portfolio. Load reductions from demand buy-back programs are volatile, though not unpredictable. In almost all jurisdictions where such programs are offered, customer bid amounts vary substantially and the
load reduction delivered is substantially less than the bid amounts. Despite their flaws, bidding programs can have a place in a portfolio of DR programs by providing a no or low-risk option that allows participants to test their load response capabilities or learn the extent to which they can provide response. The goal, however, should be to move those customers to more committed options. DR-1 has been successful at helping commercial and industrial facilities become familiar with DR and develop load response strategies. A number of the customers enrolled in DR-1 have since migrated to more committed DR programs (i.e., DR-2 and DR-3), which are now being offered by OPA. Those customers currently provide predictable and reliable load reductions for operation. Although the transition of large participants has substantially reduced the DR-1 load reduction capacity, DR-1 will continue to operate through the end of 2010. There is significant value in its ability to attract customers new to DR who might otherwise be too risk-averse to try DR-2 or DR-3 initially, which entail penalties for non-performance, without having DR-1 as a no-risk training ground. That being said, there are multiple opportunities to improve the program and its operations.

FSC recommends that the program’s main goal should be to provide an opportunity for facilities new to DR to test their load reductions capability and eventually move to a more committed option or stop participation. Hap-hazard participation does not benefit OPA or participants. As a result, we recommend limiting the amount of time participants can remain on a demand buy-back program to a year, or at most, two years. Doing so would force participants to decide whether to progress to more committed types of DR or not. In light of that role, technical assistance to help customers identify discretionary load and ways to reduce it quickly and or reliably without affecting key facility operations is recommended.

The program can and should reduce the amount of error in settlement payments. Improving accuracy benefits both participants and the OPA in that it ensures adequate compensation for participants. Much of the gap between resources bid into the program and load reductions was due to inaccurate custom baselines for participants that have since left the program. There are several insights related to baselines that can improve program operations:

- Custom baselines have a place in the program, particularly for very large participants, but it is critical to conduct a robust and detailed assessment of baseline accuracy and select the baseline that minimizes error;
- Load reductions should be larger than 0.5 MW or 15% of the average summer peak period load. OPA currently only has a megawatt load reduction threshold. The percent load reduction component is critical in that settlement baseline errors are magnified in the baseline load reductions and settlement payments. The smaller the percent load reduction, the more the baseline error affects load reductions estimated for settlement and payments;
- Require that participants exceed a threshold of reductions relative to their bid amounts. We recommend that participants should only be paid for reductions if they exceed 50% of the bid amount. This is critical in that the settlement risk is asymmetrical. Even when average error for a baseline is, on average, zero, baseline errors occur for specific customers and for specific event days. Without such a requirement, OPA can potentially pay customers that do not engage in load reductions but benefit from settlement error. The threshold reduces those errors.

The program can and should revise the trigger mechanism to improve when and how often operations are called. First, participants bid on more load reduction and participate more often when events are limited to instances of true-system needs. Most of the value from DR-1 and similar programs is related to capacity value. The need for additional resources is typically driven by the substantially higher levels of demand in the top 100 hours of extreme weather years. DR programs have the capability to target those high-need hours without having to supply energy for several hundred in order for the investment to remain viable. The current DR-1 trigger mechanism is linked to Ontario’s electricity market prices and is not highly correlated with system need for capacity. Market prices are higher during winter because of the dual use of natural gas for heating and generation; price spikes usually occur when the system is recovering from shocks such as forced generation or transmission outages. As a result, customer bids and verified load reductions exhibit low coincidence with system peaking conditions that drive the need for additional capacity. Under the current rules, many of the most critical hours do not pass the DR-1 program eligibility threshold. As a result, FSC recommends two main changes:

- Restrict the number of event hours to less than 300. Doing so signals to participants that their participation is indeed critical to the system and avoids exhausting participants; and

- Improve alignment between the dispatch of the program, system needs, and the dispatch of other DR programs. A couple of very viable options include the use of generator heat rates or the IESO’s supply cushion. Heat rates provide an advantage in that they are not linked to natural gas prices; rather, they are linked to how much of the generation supply stack is in operation. The IESO’s day ahead supply cushion estimates also provide a clear indication of the need for additional resources. When the supply cushion is small, relative to demand, there is a higher likelihood of system disturbances that could endanger the operations of the electricity grid.
5 DR-2 PROGRAM

DR-2 is a contractual load shifting program in which participants specify the load shift window and amount of load shifting. The program does not compensation or for changes in the overall level of electricity consumption that may be associated with economic fluctuations or other factors. Load reduction during peak periods must be accompanied by shifting to off-peak periods.

DR-2 resources are not dispatchable since the program is not event based. The program initially started in 2009. It was in a transition phase from March through October. The transition allowed OPA and participants to determine whether the initial program rules were effective or required modification. They also allowed participants to test their load shifting capabilities and adjust the amount of load shifted and the peak participation hours to reflect their capabilities. Throughout the transition phase payment offsets were not applied. The program rules were modified in October, 2009 and the program officially launched in November, 2009.

This section focuses on the insights that can be drawn for both program design and operations. It provides information on the program background and rules. This is followed by an overview of the participant mix and the contracted load reduction resources. We then include a discussion of program reductions versus load shifting and the subtle difference between the two. This is followed by a summary of program reductions observed in 2009. We then present load reductions for operation, for *ex ante* load reductions for cost-effectiveness and long term planning. The section concludes with recommendations to improve DR-2 program operations and performance.

*Ex ante* load reductions are most appropriate for demand response (DR) programs since they are designed to provide an insurance or option-based estimate of load reductions that can be expected under extreme conditions (typically weather or electricity price related). DR programs are designed to help address electricity system capacity issues when the system is under the most stress. *Ex ante* load reductions are based on the amount of contractual load response and the degree to which participants have historically delivered those load reductions. For the *ex ante* load reductions, FSC factored in the scheduled non-performance rate derived from historical facility shut down periods in 2006-2009, and the extent to which load reductions deviated from the contracted load reduction on an hourly basis.

Importantly, the experience with DR-2 program performance in 2009 is very limited. Any conclusions should be drawn with caution until a full year of experience with the program has been analyzed. FSC believes that the results for the transition period are not representative of future performance. There are several reasons for this:

- Payment offsets and penalties were not enforced during the transition period;
Participants did not have the final baseline data;

Participants obtained a better understanding of their load shifting capability during the transition period; and

Participation and settlement rules were modified for the final contract period.

As a result, the *ex ante* load reductions are entirely based on the contract period, even though it is limited to November and December of 2009. Given their preliminary nature, the findings and recommendation from the FSC report should be used with caution.

## 5.1 Program Background

DR-2 is a contractual load shifting program in which participants specify the load shift window and amount of load shifting. An example of load shift would be to reduce production below normal levels during the on-peak period and to undertake that production during the off-peak hours. The peak period for the program is extensive, lasting from 7 AM to 7 PM and providing participants a wide degree of flexibility for load shift hours. Participants of the program can contract to reduce a predetermined amount of load for a minimum period of 4 consecutive hours up to a maximum of 12 consecutive hours. There are three options for participation; summer months, winter and summer months, or all year.

DR-2 participants notify OPA and the IESO of non-performance days, which are analogous to scheduled generator outages. They enable the IESO to better operate the system and schedule alternate resources to meet the higher demand for those days. Non-performance days lead to reductions in the participant payments. Unscheduled non-performance—failure to meet contractual obligations during events—leads to even larger payment reductions.

Participants receive compensation for both availability and utilization. The availability payments are designed to reflect the capacity value from the resource. The utilization payments reflect changes in the cost of electricity associated with shifting usage from higher priced period to lower price periods. Payments differ for summer and winter months.

For settlement purposes, compliance with the contracted load reductions is determined through a monthly baseline method. Failure to comply with the contract requirements of the DR-2 Program can result in set-offs against potential revenue. The payment calculations are complex but can be reduced to a few basic components:

- Pre-DR-2 hourly consumption data is used to determine the customer load shape;
- The customer must change its load shape—load reduction during peak periods must be accompanied by shifting to off-peak periods;
- The load shift estimates are calibrated for changes in overall monthly consumption;
- Participants must provide consistent reductions during the selected peak hours during each day; and
- Participants must reduce volatility over selected peak hours—spikes in demand above specific thresholds lead to payment setbacks.

5.2 Program Participation

As mentioned earlier, the program is relatively new and was in a transition phase for most of 2009. Importantly, all the direct participants transitioned from DR-1 to DR-2. From 2007 through February 2009, these customers experienced a number of events wherein customers perturbed their naturally occurring load patterns by reducing load. 1,800 DR-1 event hours were called in 2007 and 1,200 event hours were called in 2008.

In total, DR-2 has enrolled three large participants which jointly constitute all of the program load and load reduction. Figure 5-1 shows the distribution of load and contracted load reduction post the transition phase among the direct participants as of December 2009.

![Figure 5-1: Distribution of Peak Load and Contracted Load Shift by Participant As of December 2009](image)

5.3 Contracted Shift Amounts and Scheduled Non-performance Patterns

Table 5-1 summarizes the contracted load shift amount and window for the transition period and official contracts. Although the transition period lasted from March through October, the contracted load shift and window remained constant. For the contractual period, participants lowered the contracted load shift amount and narrow the window. Overall, the contracted load reduction lowered from 153 MW to 119 MW.
### Table 5-1: DR-2 Contractual Load Shifting by Period

<table>
<thead>
<tr>
<th>Period</th>
<th>Participant Name</th>
<th>Contracted MW</th>
<th>Contracted Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transition</td>
<td>Participant 1</td>
<td>63</td>
<td>8 AM to 6 PM</td>
</tr>
<tr>
<td></td>
<td>Participant 2</td>
<td>20</td>
<td>12 PM to 4 PM</td>
</tr>
<tr>
<td></td>
<td>Participant 3</td>
<td>70</td>
<td>8 AM to 4 PM</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>153</td>
<td>Varies by hour</td>
</tr>
<tr>
<td>Summer</td>
<td>Participant 1</td>
<td>54</td>
<td>7 AM to 3 PM</td>
</tr>
<tr>
<td></td>
<td>Participant 2</td>
<td>13</td>
<td>1 PM to 5 PM</td>
</tr>
<tr>
<td></td>
<td>Participant 3</td>
<td>52</td>
<td>8 AM to 6 PM</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>119</td>
<td>Varies by hour</td>
</tr>
<tr>
<td>Winter and Shoulder</td>
<td>Participant 1</td>
<td>54</td>
<td>7 AM to 3 PM</td>
</tr>
<tr>
<td></td>
<td>Participant 2</td>
<td>13</td>
<td>12 PM to 4 PM</td>
</tr>
<tr>
<td></td>
<td>Participant 3</td>
<td>52</td>
<td>8 AM to 6 PM</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>119</td>
<td>Varies by hour</td>
</tr>
</tbody>
</table>

The changes reflect both better understanding of the program by participants and the substitution of DR-1 evaluation reductions for settlement baseline reductions in the program baseline. Due to their prior participation in DR-1, to estimate the customer’s naturally occurring load shape for settlement baselines, both DR-1 related load reductions and shifting were added back to participant loads. For the transition period, this was done using reductions estimated through DR-1 settlement baselines. For the final enrollment period, the add-ins were based on the 2007-2008 impact evaluation of the program. There were significant differences between the evaluation and settlement reductions due mostly to the use of custom baselines for some of those customers. Those differences are detailed in the 2007-2008 DR-1 evaluation. The transition period load reduction commitments were not based on the final baseline estimates, while the contract baseline estimates were. Since many participant manage load to meet settlement baselines, this may have affected load shifting during the transition period.

In 2009, the contracted load shifting from peak hours was impacted by the economic conditions, leading to substantial amount of scheduled non-performance days. The frequency of scheduled non-performance in 2009 is not likely representative of future patterns. It does, however,

---

highlight the need to diversify the program. All of the participants are in the same industry, further concentrating the effect of economic downturns.

Figure 5-2 shows the scheduled weekday performance as the percent of the contracted MW for each of the months in the transition and contract periods. Because concentration of the program scheduled non-performance by a single participant severely affected the amount of expected load reductions.

![Figure 5-2: DR-2 Average Scheduled Performance by Month](image)

Figure 5-3 shows the load patterns for each of the participants for the non-performance days and other days. Non-performance days correlate highly to facility shut downs. During DR-2 transition and enrollment periods, participants shut down the main processes at their facilities in over 90% of scheduled non-performance days. Participant 1 and Participant 3 frequently scheduled non-performance days. In contrast, Participant 2 scheduled non-performance days in less than 10% of weekdays. Almost no shutdowns occurred outside of scheduled non-performance.
While full program participation is limited to two months, significantly more data is available to determine non-performance rates if facility shut downs are used as a proxy for non-performance during the 2006-2008 time period. While less than half of the contracted load shifting was unavailable in 2009 due to facility shut downs, for 2006 and 2008, those same facilities were shut down less than 5% of the days in 2006 to 2008.

5.4 Differences Between Program Reductions and Load Shifting Calculated for Settlement

A key challenge for DR-2 is to distinguish between changes in consumption levels unrelated to the program and changes in hourly usage patterns. To avoid compensating participants from changes in overall output, the DR-2 rules make it clear that reductions in usage do not constitute load shifting, they must be accompanied by increases in use during off-peak. In contrast, the regression models can control for changes in economic conditions and estimate the change in electricity use patterns, including reductions, attributable to the program. The primary risk of doing so is confounding the program reductions with economic conditions if they are highly correlated.
Overall, electricity consumption levels of DR-2 participants changed relative to their pre-enrollment level consumptions that are used to calculate settlement baselines. The key difference between pre and post DR-2 enrollment is the number of days in which operations were shutdown. For almost all days when operations shut down, participants scheduled it as non-performance day with the IESO. Once those days are factored out average daily consumption level for the pre and post DR-2 period are similar.

Figure 5-4 shows the average weekday hourly load for DR-2 participants excluding non-performance days and days when the facility was shut down. The graph excludes scheduled non-performance, which affects the DR-2 official contract period load levels since Participant 1 shut down their facility throughout the period. The graph compares the time span prior to any participation in DR programs, the period when they participated in DR-1 (with DR-1 reductions added back in), the DR-2 transition period, and the DR-2 contractual period. It highlights several key issues. First, although the graph shows unfiltered results, it is clear that customers reduced load during peak hours and engaged in some load shifting during the transition period. This also true of the contract period and is also observed at the individual participant level. Second, customers have a limited amount of load to shift and some of their response is in fact reductions in consumption. However, DR-2 only provides compensation for shifting. Given the industrial processes in place, the maximum aggregate load for DR-2 participants is approximately 260 MW.
Importantly, the meter data presented above does not control for differences in electricity market prices, economy (aside from shutdowns), and other factors that vary across the different time periods. To estimate the load reductions the historical pre-enrollment load patterns from 2006-2008, which include the DR-1 reductions added back in, were employed to predict the counterfactual – that is, the expected load levels for DR-2 participation months, given underlying market and economic conditions, if the participants had not signed up for the program. Based on the regression results, the load shifting to off-peak periods does not always match load reduction during the peak period.

As a cross-check to the regression based load reduction estimates, we estimated the amount shifting without the influence of changes in consumption. To do this, the load shapes were normalized by dividing the hourly load for each participant by the average hourly load for all qualifying days in the month. This process essentially subtracts out differences in overall consumption—the average value for all shapes is equal to one—and allows us to focus solely on changes in the load shape. The load shapes can be directly compared in their normalized form or scaled for the actual consumption in the month. In this instance, the reference or counterfactual load shape was scaled to reflect the actual consumption observed during each DR-2 month. By construction, the reductions over the peak period window match the increases...
in load outside of the peak window. Figure 5-5 compares the normalized loads with and without DR and reflect the load shifting under the strict rule that load reductions, even if attributable, to the program do not count as load shifting.

Figure 5-5 shows the hourly load shifting for the transition and contract period. The regression predicted counterfactual and load with DR have been rescaled so the overall consumption of is equal. The average demand throughout the day is exactly the same and reductions in consumption over the peak window match the increases outside of the selected peak window.

**Figure 5-5:**
**Hourly Load Shifting for DR-2 Participants (Rescaled Counter-Factual) For Transition and Contract Period, Excluding Planned Non-performance Days**

Table 5-2 compares the total contracted and day ahead scheduled resources to the actual load reductions and shift attribution for the contract period. The load reductions across all program shift hours were, on average 99.3% of the expected drop in load over that time frame. In contrast, the shift attribution is lower, 88.4%. The program reductions from the regression analysis were employed as the basis for the *ex post* and *ex ante* load reduction estimates rather than rely on shift attribution. It is the load reductions that directly affect the system load. The shift attribution is primarily an administrative measure designed to ensure OPA does not erroneous compensate participants for changes in production.
Table 5-2:
2009 Hourly Comparison of Load Reduction to Shift Attribution

<table>
<thead>
<tr>
<th>Hour Ending</th>
<th>Contracted MW</th>
<th>Day Ahead Contracted MW (scheduled)</th>
<th>Load Reduction MW</th>
<th>% of Scheduled MW</th>
<th>Load Shift MW</th>
<th>% of Scheduled MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00 AM</td>
<td>106</td>
<td>54</td>
<td>62.3</td>
<td>115.40%</td>
<td>52.7</td>
<td>97.50%</td>
</tr>
<tr>
<td>9:00 AM</td>
<td>106</td>
<td>54</td>
<td>54.5</td>
<td>100.90%</td>
<td>46.8</td>
<td>86.60%</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>106</td>
<td>54</td>
<td>51.5</td>
<td>95.30%</td>
<td>43.8</td>
<td>81.10%</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>106</td>
<td>54</td>
<td>48.5</td>
<td>89.80%</td>
<td>44.4</td>
<td>82.20%</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>106</td>
<td>54</td>
<td>51.7</td>
<td>95.70%</td>
<td>47.4</td>
<td>87.70%</td>
</tr>
<tr>
<td>1:00 PM</td>
<td>119</td>
<td>67</td>
<td>56.4</td>
<td>84.20%</td>
<td>56.6</td>
<td>84.60%</td>
</tr>
<tr>
<td>2:00 PM</td>
<td>119</td>
<td>67</td>
<td>59.2</td>
<td>88.30%</td>
<td>59.1</td>
<td>88.20%</td>
</tr>
<tr>
<td>3:00 PM</td>
<td>119</td>
<td>67</td>
<td>60.8</td>
<td>90.70%</td>
<td>58.7</td>
<td>87.50%</td>
</tr>
<tr>
<td>4:00 PM</td>
<td>65</td>
<td>13</td>
<td>34.2</td>
<td>262.70%</td>
<td>28.8</td>
<td>221.20%</td>
</tr>
<tr>
<td>5:00 PM</td>
<td>65</td>
<td>13</td>
<td>14.6</td>
<td>111.90%</td>
<td>1.5</td>
<td>11.40%</td>
</tr>
<tr>
<td>6:00 PM</td>
<td>52</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7:00 PM</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AVERAGE</td>
<td>97.2</td>
<td>49.7</td>
<td>49.3</td>
<td>99.30%</td>
<td>44</td>
<td>88.40%</td>
</tr>
</tbody>
</table>

5.5 2009 *Ex post* Load Reduction Results
The participants delivered a higher share of the day-ahead scheduled resources during the contract period than during the transition months. Their actions suggest the transition helped them better understand their load shifting capabilities. All participants adjusted both the load shift amount and window for the contract period.

Table 5-3 summarizes the hourly load reductions and the 90% confidence intervals for the 2 months when the program was in its official version in 2009. The uncertainty associated with the load reductions estimates is relative narrow. For several hours, the load reductions for some participants are counter-balanced by load shifting to that hour by other participants. This is especially true for some of the more critical hours, 2 PM to 6 PM, when system peaks are more likely to occur. While flexibility is one key features of the program, it can produce more value by increasing the incentives for the more critical hours (and decreasing them for less critical hours), narrowing the eligible hours, or requiring that load increases occur outside of the program window, not simply outside of the window elected by the participant.
Table 5-3:
2009 DR-2 Ex post Hourly Load Reductions and Confidence Interval for Contractual Period

<table>
<thead>
<tr>
<th>Hour Ending</th>
<th>Contracted Load Shift</th>
<th>Day-Ahead Contracted Resources (Scheduled)</th>
<th>Load Reductions</th>
<th>90% Confidence Interval</th>
<th>% of Scheduled</th>
<th>90% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
<td>Upper</td>
<td>Lower</td>
</tr>
<tr>
<td>8:00 AM</td>
<td>106.0</td>
<td>54.0</td>
<td>64.4</td>
<td>63.4</td>
<td>65.3</td>
<td>119%</td>
</tr>
<tr>
<td>9:00 AM</td>
<td>106.0</td>
<td>54.0</td>
<td>58.7</td>
<td>57.9</td>
<td>59.4</td>
<td>109%</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>106.0</td>
<td>54.0</td>
<td>54.6</td>
<td>54.0</td>
<td>55.2</td>
<td>101%</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>106.0</td>
<td>54.0</td>
<td>51.6</td>
<td>51.1</td>
<td>52.1</td>
<td>96%</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>106.0</td>
<td>54.0</td>
<td>52.6</td>
<td>52.2</td>
<td>53.0</td>
<td>97%</td>
</tr>
<tr>
<td>1:00 PM</td>
<td>119.0</td>
<td>67.0</td>
<td>56.6</td>
<td>56.3</td>
<td>56.9</td>
<td>84%</td>
</tr>
<tr>
<td>2:00 PM</td>
<td>119.0</td>
<td>67.0</td>
<td>58.7</td>
<td>58.5</td>
<td>58.9</td>
<td>88%</td>
</tr>
<tr>
<td>3:00 PM</td>
<td>119.0</td>
<td>67.0</td>
<td>61.0</td>
<td>60.7</td>
<td>61.2</td>
<td>91%</td>
</tr>
<tr>
<td>4:00 PM</td>
<td>65.0</td>
<td>13.0</td>
<td>34.3</td>
<td>32.8</td>
<td>35.7</td>
<td>263%</td>
</tr>
<tr>
<td>5:00 PM</td>
<td>65.0</td>
<td>13.0</td>
<td>14.7</td>
<td>11.6</td>
<td>17.8</td>
<td>113%</td>
</tr>
<tr>
<td>6:00 PM</td>
<td>52.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7:00 PM</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Average 7am-7pm</td>
<td>97.2</td>
<td>49.7</td>
<td>50.7</td>
<td>48.7</td>
<td>52.7</td>
<td>102%</td>
</tr>
</tbody>
</table>

Table 5-4 and Table 5-5 summarize the ex post load reductions and the shift attribution, by hour, for the average weekday and the monthly system peak day. They highlight several issues. As mentioned earlier, DR-2 program load reductions are higher than the shift attribution and contract period reductions are higher than transition period reductions. The load reductions reflect the reductions attributable to the program. The shift attribution provides a more conservative estimate without the potential of confounding program reductions with economic conditions. Second, because the program is highly concentrated, coincident non-performance days can lead to substantially smaller load reductions than the contracted load shift. This is especially evident in month of July when both of the larger participants had scheduled non-performance days coincident with the monthly system peak. In order to improve reliability OPA will need to diversify the participant base. For the contract period, reductions are lower because
of the scheduled non-performance by Participant 1, but they are also more stable because the participant hours align better and their shifting does not counter-balance each other.

Table 5-4:
2009 Hourly DR-2 *Ex post* Load Reductions by Month
For the Average Weekdays and for Monthly System Peak Days

<table>
<thead>
<tr>
<th>Day Type</th>
<th>Hour Ending</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8:00 AM</td>
</tr>
<tr>
<td>--------------------</td>
<td>-------------</td>
</tr>
<tr>
<td><strong>Average Weekday</strong></td>
<td>Mar</td>
</tr>
<tr>
<td></td>
<td>Apr</td>
</tr>
<tr>
<td></td>
<td>May</td>
</tr>
<tr>
<td></td>
<td>Jun</td>
</tr>
<tr>
<td></td>
<td>Jul</td>
</tr>
<tr>
<td></td>
<td>Aug</td>
</tr>
<tr>
<td></td>
<td>Sep</td>
</tr>
<tr>
<td></td>
<td>Oct</td>
</tr>
<tr>
<td></td>
<td>Nov</td>
</tr>
<tr>
<td></td>
<td>Dec</td>
</tr>
<tr>
<td><strong>Transition Period</strong></td>
<td>-19.1</td>
</tr>
<tr>
<td><strong>Contract Period</strong></td>
<td>-64.4</td>
</tr>
<tr>
<td><strong>Monthly System Peak</strong></td>
<td>Mar</td>
</tr>
<tr>
<td></td>
<td>Apr</td>
</tr>
<tr>
<td></td>
<td>May</td>
</tr>
<tr>
<td></td>
<td>Jun</td>
</tr>
<tr>
<td></td>
<td>Jul</td>
</tr>
<tr>
<td></td>
<td>Aug</td>
</tr>
<tr>
<td></td>
<td>Sep</td>
</tr>
<tr>
<td></td>
<td>Oct</td>
</tr>
<tr>
<td></td>
<td>Nov</td>
</tr>
<tr>
<td></td>
<td>Dec</td>
</tr>
<tr>
<td><strong>Transition Period</strong></td>
<td>-18.7</td>
</tr>
<tr>
<td><strong>Contract Period</strong></td>
<td>-62.3</td>
</tr>
<tr>
<td>Day Type</td>
<td>Month</td>
</tr>
<tr>
<td>---------------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Weekday</td>
<td>Mar</td>
</tr>
<tr>
<td></td>
<td>Apr</td>
</tr>
<tr>
<td></td>
<td>May</td>
</tr>
<tr>
<td></td>
<td>Jun</td>
</tr>
<tr>
<td></td>
<td>Jul</td>
</tr>
<tr>
<td></td>
<td>Aug</td>
</tr>
<tr>
<td></td>
<td>Sep</td>
</tr>
<tr>
<td>Transition Period</td>
<td>Mar</td>
</tr>
<tr>
<td>Contract Period</td>
<td>-49.1</td>
</tr>
<tr>
<td>Monthly System Peak</td>
<td>Mar</td>
</tr>
<tr>
<td></td>
<td>Apr</td>
</tr>
<tr>
<td></td>
<td>May</td>
</tr>
<tr>
<td></td>
<td>Jun</td>
</tr>
<tr>
<td></td>
<td>Jul</td>
</tr>
<tr>
<td></td>
<td>Aug</td>
</tr>
<tr>
<td></td>
<td>Sep</td>
</tr>
<tr>
<td></td>
<td>Oct</td>
</tr>
<tr>
<td></td>
<td>Nov</td>
</tr>
<tr>
<td>Transition Period</td>
<td>-50.4</td>
</tr>
<tr>
<td>Contract Period</td>
<td>-47.4</td>
</tr>
</tbody>
</table>

5.6 Load Reductions for Operations
The reliability of the day-ahead scheduled resources is critical for operations. The load reductions must be reliable and sustained through the event period. There several limitations to producing reductions for operations. We were unable to evaluate the performance under final
rules for one participant that accounts for over 40% of the contracted shifting. Neither were we able to observe load reductions under final rules for summer or shoulder months. In addition, the regression produces program reductions by month and hour, but doesn’t produce them on a daily basis.

Despite those limitations, there is clear evidence that participants are monitoring their load on a daily basis and controlling the volatility of their load. The coefficient of variation is a standard measure of volatility (standard deviation divided by the mean). It is possible to compare the load volatility over the same time period for those customers prior to DR-2 enrollment and during their participation period. When we do so, we find that volatility over the load shift has decrease by roughly 75%.

5.7 Load Reductions for Long Term Planning and Cost-Effectiveness

The ex ante load reductions are designed to provide an overview of load reductions that can be used for long term planning and cost-effectiveness. The DR-2 participants are not weather sensitive and have a contractual amount of load shifting for a specific period. As a result, the ex ante load reductions factor in the amount of contractual load response and the degree to which these customers have historically delivered those load reductions. In other words, they are de-rated for non-performance days and for deviations from the load reduction expected by system operators (which factor in scheduled non-performance). Importantly, the experience with DR-2 program performance in 2009 is very limited. Any conclusions should be drawn with caution until a full year of experience with the program.

Although customers participated in the transition period, we believe the program participation for that time frame is not representative of future performance. There are several reasons for this. Payment offsets and penalties were not enforced during the transition period. Participants also did not have the final baseline data. It is also clear that participant obtained a better understanding of their load shifting capability during the transition period. Finally, participation and settlement rules were modified for the final contract period. As result, the ex ante load reductions are based on the contract period, even though it is limited to November and December of 2009. Ideally, seasonal patterns in expected load reductions would be analyzed and incorporated, but the program history is too short to do so.

Non performance days correlate highly to facility shut downs. During DR-2 transition and enrollment periods, participants shut down the main processes at their facilities in over 90% of scheduled non-performance days. In addition, almost no shutdowns occurred outside of scheduled non-performance. While program participation on which ex ante estimates are based is limited to two months, significantly more data is available to determine non-performance rates if facility shut downs are used as a proxy for non-performance during the 2006-2008 time
period. This allows ex ante estimates to be de-rated in a way that incorporates appropriate historical seasonal variation.

There is a trade off here. Using facility shutdowns as a proxy for non-performance produces small errors in classification. Facility shutdowns that are were not scheduled for non-performance are incorrectly classified as non-performance day. Likewise, scheduled non-performance days when facilities were not shut down may be incorrectly classified as active days. The incidence of misclassification can be directly observed during the DR-2 transition and enrollment periods and it is small—approximately less than 10%.

There are risks associated with using only actual non-performance data as well. During the time period for which actual non-performance data is available, the participants were severely affected by the economic conditions leading to less production and temporary facility shut downs that were typically scheduled as non-performance days. The facility of a participant with 52 MW of contracted load reduction was completely shut down for the entire contract period. This facility has since reopened, which would not be captured if only using actual non-performance data. Any seasonal variation in non-performance periods is also lost, since actual data is confined to two months of the year.

For the ex ante load reductions, we factor in the scheduled non-performance rate derived from historical facility shut down periods in 2006-2009, and the extent to which load reductions deviated from the contracted load reduction on an hourly basis.

Another key issue is whether to use load reductions or shift attribution for planning purposes. As noted earlier, they are not one and the same. Load reductions are larger by roughly 14 MW (30%), with some variation by hour. We recommend employing the load reductions for long term planning because they directly affect the system load. The shift attribution is primarily and administrative measure designed to limit payment to load shifting. It helps avoid having to determine whether DR-2 changes in consumption are due to less production or due to DR-2.

Table 5-6 shows the hourly ex ante load reduction estimates. They are not differentiated by month since the program is contractual and we do not yet have evidence to assess whether or not seasonal effects exist. They are also the same for extreme and normal conditions. This again reflects the contractual nature of the program, but it also reflects the fact that the program participants are not weather sensitive.
<table>
<thead>
<tr>
<th>Season</th>
<th>Hour Ending</th>
<th>Contracted MW</th>
<th>Day Ahead Contracted MW (Scheduled)</th>
<th>Ex ante Load Reduction Estimates MW</th>
<th>90% Confidence Interval</th>
<th>Lower bound</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Summer Months</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8:00 AM</td>
<td>106</td>
<td>102.8</td>
<td>118.6</td>
<td></td>
<td></td>
<td>116.9</td>
<td>120.4</td>
</tr>
<tr>
<td>9:00 AM</td>
<td>106</td>
<td>91.6</td>
<td>92.5</td>
<td></td>
<td></td>
<td>91.3</td>
<td>93.7</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>106</td>
<td>91.6</td>
<td>87.3</td>
<td></td>
<td></td>
<td>86.3</td>
<td>88.2</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>106</td>
<td>91.6</td>
<td>82.3</td>
<td></td>
<td></td>
<td>81.5</td>
<td>83.1</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>106</td>
<td>91.6</td>
<td>87.7</td>
<td></td>
<td></td>
<td>87.1</td>
<td>88.3</td>
</tr>
<tr>
<td>1:00 PM</td>
<td>106</td>
<td>91.6</td>
<td>77.1</td>
<td></td>
<td></td>
<td>76.7</td>
<td>77.5</td>
</tr>
<tr>
<td>2:00 PM</td>
<td>119</td>
<td>104.0</td>
<td>91.8</td>
<td></td>
<td></td>
<td>91.4</td>
<td>92.2</td>
</tr>
<tr>
<td>3:00 PM</td>
<td>119</td>
<td>104.0</td>
<td>94.3</td>
<td></td>
<td></td>
<td>93.9</td>
<td>94.6</td>
</tr>
<tr>
<td>4:00 PM</td>
<td>119</td>
<td>94.8</td>
<td>94.1</td>
<td></td>
<td></td>
<td>90.1</td>
<td>98.2</td>
</tr>
<tr>
<td>5:00 PM</td>
<td>65</td>
<td>51.8</td>
<td>57.9</td>
<td></td>
<td></td>
<td>45.7</td>
<td>70.2</td>
</tr>
<tr>
<td>6:00 PM</td>
<td>65</td>
<td>49.9</td>
<td>49.6</td>
<td></td>
<td></td>
<td>39.1</td>
<td>60.1</td>
</tr>
<tr>
<td>7:00 PM</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>AVERAGE</strong></td>
<td></td>
<td><strong>102.1</strong></td>
<td><strong>87.8</strong></td>
<td></td>
<td></td>
<td><strong>81.4</strong></td>
<td><strong>88.1</strong></td>
</tr>
<tr>
<td><strong>Winter and Shoulder Months</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8:00 AM</td>
<td>106</td>
<td>102.5</td>
<td>118.2</td>
<td></td>
<td></td>
<td>116.5</td>
<td>120.0</td>
</tr>
<tr>
<td>9:00 AM</td>
<td>106</td>
<td>92.7</td>
<td>93.6</td>
<td></td>
<td></td>
<td>92.3</td>
<td>94.8</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>106</td>
<td>92.7</td>
<td>88.3</td>
<td></td>
<td></td>
<td>87.4</td>
<td>89.3</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>106</td>
<td>92.7</td>
<td>83.3</td>
<td></td>
<td></td>
<td>82.5</td>
<td>84.0</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>106</td>
<td>92.7</td>
<td>88.8</td>
<td></td>
<td></td>
<td>88.1</td>
<td>89.4</td>
</tr>
<tr>
<td>1:00 PM</td>
<td>119</td>
<td>105.2</td>
<td>88.6</td>
<td></td>
<td></td>
<td>88.1</td>
<td>89.0</td>
</tr>
<tr>
<td>2:00 PM</td>
<td>119</td>
<td>105.2</td>
<td>92.9</td>
<td></td>
<td></td>
<td>92.5</td>
<td>93.3</td>
</tr>
<tr>
<td>3:00 PM</td>
<td>119</td>
<td>105.2</td>
<td>95.4</td>
<td></td>
<td></td>
<td>95.0</td>
<td>95.8</td>
</tr>
<tr>
<td>4:00 PM</td>
<td>65</td>
<td>52.8</td>
<td>52.4</td>
<td></td>
<td></td>
<td>50.2</td>
<td>54.7</td>
</tr>
<tr>
<td>5:00 PM</td>
<td>65</td>
<td>50.7</td>
<td>56.7</td>
<td></td>
<td></td>
<td>44.7</td>
<td>68.7</td>
</tr>
<tr>
<td>6:00 PM</td>
<td>52</td>
<td>40.5</td>
<td>40.2</td>
<td></td>
<td></td>
<td>31.7</td>
<td>48.8</td>
</tr>
<tr>
<td>7:00 PM</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>AVERAGE</strong></td>
<td></td>
<td><strong>97.2</strong></td>
<td><strong>84.8</strong></td>
<td></td>
<td></td>
<td><strong>81.9</strong></td>
<td><strong>78.6</strong></td>
</tr>
</tbody>
</table>
In using *ex ante* load Reduction estimates for planning purposes, it is critical to link DR-2 resources to system demand and other resources. While the above estimates factor in the scheduled non-performance and departures from the expected load reductions for operations, they do not direct account for the timing and magnitude of load Reductions. From a capacity perspective, resources available in the summer mid-afternoon hours, the hours of system peak, generally have higher value than resources in non-peaking hours. It is necessary to estimate the incorporate the extent to which DR-2 resources serve as capacity without increasing the likelihood of unserved electricity – a value more commonly known as effective load carrying capacity. This is typically done for cost-effectiveness analysis and long-term planning.

### 5.8 Conclusions and Recommendations

Given the limited history of the program, the findings and recommendation from this report are by necessity limited. A full year of program Reductions under the contract rules is necessary before drawing stronger conclusion. Despite those evaluations, the evaluation of DR-2 provides several insights.

Some of the key finding and recommendations include:

- Load reductions were generally greater than the shift attribution. This is because in general participants reduced more load than they shifted.

- Overall, the average load reductions for the transition period were 50% of the contracted MW, after factoring out non-performance days. For the contractual period, however, the load reductions were 98.5% of the contracted MW.

- There was a substantial amount of scheduled non-performance. On average, over the transition period, 62.7% of the contracted load was scheduled for non-performance. Over the contractual period, 53.7% of the contracted load was scheduled for non-performance. Most of the scheduled non-performance is from one participant, but it reflects the concentration of the program and the lack of diversity of the current participants across industries. While the extent of non-performance is likely related to economic conditions and leads to conservative estimates for planning, it is the only empirical data available for this unique and new program.

- The load reductions customers can counteract each other because their event windows are not coordinated. One participant may be shifting load to an hour, while another may shifting load away from the same hour.

- Contractual load shifting amounts are lower for the critical summer mid to late afternoon hours. The highest amount of load shifting occurs between 9 am and 12 pm.

- The program better coordinate load shifting by increasing the incentives for the more critical hours (and decreasing them for less critical hours), narrowing the eligible hours, or requiring that load increases occur outside of the program window, not simply outside of the load shift window elected by the participant.
Participants delivered a higher share of the expected load reduction during the contract period. Their actions suggest the transition helped them better understand their load shifting capabilities. All participants adjusted both the load shift amount and window for the contract period.
6 DR-3 PROGRAMS

The DR-3 program was evaluated last year. The evaluation was conducted early in the development of the DR-3 program, which has several advantages but also limits the scope of the findings. The early evaluation provides OPA with detailed information on how well the existing program works, the realization rate of contractual load reductions at the time, and specific steps that could be taken to make the program more effective. However, the evaluation was conducted during a phase of substantial enrollment growth and almost all of the events occurred during non-summer months.

There are several reasons why the 2009 evaluation of the program is significant:

- The load reduction capability of the program and the participant mix has changed substantially. Since December 2008, the contracted load reduction has nearly doubled from 86 MW to 170 MW.

- The program has diversified and the role of aggregators has grown. While a few large direct participants still have a substantial effect on the overall program results, the program better reflects the benefits of aggregating across multiple loads, which generally increases predictability.

- All of the 2009 events took place during the summer. Since the program is largely designed to provide capacity, it adds to the body of knowledge of whether the contracted load reductions are realized.

The unique economic conditions of 2009 likely reflect the upper bound of non-performance days. DR-3 customers have the ability to notify the IESO if they are unable to perform for specific days. This is typically done in advance when facilities are shut down and is analogous to scheduled generator outages.

Despite the wealth of additional information, any conclusions need to factor in the unique 2009 economic conditions and the fact that the customer mix will likely continue to diversify. The direct participants and several less sizable individual contributors were affected by the economic conditions. Since the data is limited to 2008 and 2009, so is the data on the extent of non-performance days and the extent to which customers delivered the expected load reductions. Although 6 summer events were called in 2009, exogenous conditions in the general economy had a large impact on program performance. Therefore, differences in the extent to which customers delivered the contracted MW may be due to the impact of the general economy in 2009 as opposed to the summer season. Because of the limited history, the share of contractual capacity that can be expected is based on the average historical event in 2008 and 2009 and may very well be conservative.
6.1 Program Background
DR-3 allows participants and aggregators to enter into contractual agreements for load reductions with OPA. Participants can choose to enroll directly with OPA, provided they meet minimum load reduction criteria, or can participate through an aggregator. Under DR-3, an aggregator or direct participant must commit to a specific load reduction amount for either 100 or 200 hours per year. OPA has discretion regarding the timing of events and currently is determining event days based on IESO day-ahead supply cushion estimates. In exchange for load reductions, the DR-3 program makes both availability (capacity) and energy payments. Unlike OPA’s demand buy-back program (DR-1), OPA can reduce payments if participants fail to provide the contractual load reduction or are unavailable to provide load reductions (the equivalent of a scheduled outage). For settlement purposes, compliance with the contracted load reductions is determined through day matching baseline methods.

DR-3 participants notify OPA and the IESO of any short term fluctuations in load reduction capability due to facility maintenance or down time. These days are classified as non-performance days and are analogous to scheduled generator outages. They enable the IESO to better operate the system and schedule alternate resources for those days. Non-performance days lead to reductions in the participant payments. The payment reductions are higher if an event is called during a participant non-performance day. Unscheduled non-performance—failure to meet contractual obligations during events—leads to even larger payment reductions.

6.2 Program Participation
Since program launch, the DR-3 participant mix and load reduction capabilities have evolved substantially. The largest change in load reduction capability occurred with the enrollment of 3 large direct participants in September and October of 2008. In addition, aggregators have added participants since their contracts became effective and are expected to continue to ramp up during 2010. Based on program rules, aggregators need to attain 25 MW or more of load reduction within a ramp period. Figure 6-1 reflects the ramp up of aggregate contracted load reduction since program inception through December 2009.
Despite substantial growth of aggregator participation in the program in 2009, load reductions remain highly concentrated. As of December 31, 2009, the 3 direct program participants accounted for 38% of total participant load and 42% of the contracted load reduction. Figure 6-2 shows the distribution of load and contracted load reduction among aggregators and direct participants as of December 2009.
6.3 2009 Events and Event Conditions

Table 6-1 shows the events that were called in 2009. There were only 6 events called in 2009, compared to 14 in 2008. In 2009, the Ontario supply cushion was large compared to 2007 and 2008 due to historically low market-demand and new supply-side resources. The first event for 2009 was called on June 25 and the last one on September 14. As noted above, participants must choose between the 100-hour and 200-hour options, which represent the maximum number of event hours per year that can be called. Over the estimation period, the 200-hour commitment group was called for all 6 of the events and the 100-hour group was called for 4 events.
Table 6-1: 2009 Event Day Information

<table>
<thead>
<tr>
<th>Event Date</th>
<th>100 Hour Group</th>
<th>200 Hour Group</th>
<th>Total Contracted MW Dispatched</th>
</tr>
</thead>
<tbody>
<tr>
<td>6/25/2009</td>
<td>Activated</td>
<td></td>
<td>9.8</td>
</tr>
<tr>
<td>8/14/2009</td>
<td>Activated</td>
<td>Activated</td>
<td>145.8</td>
</tr>
<tr>
<td>8/17/2009</td>
<td>Activated</td>
<td>Activated</td>
<td>145.8</td>
</tr>
<tr>
<td>9/9/2009</td>
<td>Activated</td>
<td></td>
<td>20.1</td>
</tr>
<tr>
<td>9/10/2009</td>
<td>Activated</td>
<td>Activated</td>
<td>141.0</td>
</tr>
<tr>
<td>9/14/2009</td>
<td>Activated</td>
<td>Activated</td>
<td>141.0</td>
</tr>
</tbody>
</table>

6.4 2009 *Ex post* Load Reduction Results

For the 6 events called in 2009, the DR-3 Program delivered on average 81.8% of the day-ahead contracted load reductions, which factor in non-performance notifications. At its peak, the DR-3 Program delivered 107.2 MW of load reduction. Because of the changing mix of participants, the amount of contracted and delivered load reduction varied depending on the event day. Table 6-2 shows the average hourly load reductions for each of the event days called in 2009 and the 90% confidence bands for the estimates.

Table 6-2: DR-3 Program 2009 Load Reductions by Event Day

<table>
<thead>
<tr>
<th>Event Date</th>
<th>Group</th>
<th>Contracted MW</th>
<th>Day-Ahead Contracted MW</th>
<th>Estimated Load Reductions (MW)</th>
<th>90% Confidence Interval</th>
<th>% of Day-Ahead Contracted MW</th>
<th>90% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
<td>Upper</td>
<td>Lower</td>
</tr>
<tr>
<td>6/25/2009</td>
<td>200 only</td>
<td>9.8</td>
<td>9.8</td>
<td>8.6</td>
<td>7.6</td>
<td>9.6</td>
<td>87.7%</td>
</tr>
<tr>
<td>8/14/2009</td>
<td>100 and 200</td>
<td>145.8</td>
<td>50.6</td>
<td>46.2</td>
<td>31.8</td>
<td>60.6</td>
<td>91.2%</td>
</tr>
<tr>
<td>8/17/2009</td>
<td>100 and 200</td>
<td>145.8</td>
<td>84.3</td>
<td>77.1</td>
<td>60.3</td>
<td>93.9</td>
<td>91.5%</td>
</tr>
<tr>
<td>9/9/2009</td>
<td>200 only</td>
<td>20.1</td>
<td>20.1</td>
<td>13.6</td>
<td>11.8</td>
<td>15.4</td>
<td>67.6%</td>
</tr>
<tr>
<td>9/10/2009</td>
<td>100 and 200</td>
<td>141.0</td>
<td>129.9</td>
<td>94.9</td>
<td>71.0</td>
<td>118.9</td>
<td>73.1%</td>
</tr>
<tr>
<td>9/14/2009</td>
<td>100 and 200</td>
<td>141.0</td>
<td>129.9</td>
<td>107.2</td>
<td>83.6</td>
<td>130.8</td>
<td>82.5%</td>
</tr>
<tr>
<td>Average Event (2009)</td>
<td></td>
<td>100.6</td>
<td>70.8</td>
<td>57.9</td>
<td>51.2</td>
<td>64.6</td>
<td>81.8%</td>
</tr>
</tbody>
</table>
In aggregate, the load reductions ranged from 67.6%–91.5% of the day-ahead contracted load reduction. Several aggregators manage load based on baselines, and some of the variation between the load reduction expected and the load reduction delivered is related to differences in the accuracy of the baseline method at the participant level. In other words, inaccuracies in the settlement baseline affect the extent to which the expected load reduction is delivered.

The individual participant capacity realization rate matters due to the concentration of DR-3 contractual load reductions. Table 5-4 summarizes results by aggregator and direct participants and compares them to the contractual amounts and the load reduction estimates employed for settlement.

<table>
<thead>
<tr>
<th>Type of Participant</th>
<th>Estimated Load Reduction (MW)</th>
<th>Average Contracted MW</th>
<th>Load Reduction as a % of Contracted MW</th>
<th>Average Expected MW</th>
<th>Load Reduction as a % of Expected MW</th>
<th>Average Number of Customers per Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregator Participant</td>
<td>38.9</td>
<td>56.6</td>
<td>68.7%</td>
<td>45.8</td>
<td>85.0%</td>
<td>45.8</td>
</tr>
<tr>
<td>Direct Participant</td>
<td>28.6</td>
<td>66.0</td>
<td>43.3%</td>
<td>37.5</td>
<td>76.1%</td>
<td>2.5</td>
</tr>
<tr>
<td>Average Event (2009)</td>
<td>57.9</td>
<td>100.6</td>
<td>57.6%</td>
<td>70.8</td>
<td>81.8%</td>
<td>47.5</td>
</tr>
</tbody>
</table>

6.5 Understanding the Gap Between Contracted and Delivered Load Reductions

For the average 2009 event, the load reduction estimates are roughly 58% of contracted load reduction. The capacity contracted for the average event was 100.6 MW while the estimated load reductions were 57.9 MW. There are three main explanations for the difference: scheduled non-performance, settlement baseline bias, and participant non-performance.

DR-3 participants notify OPA and the IESO of any short-term fluctuations in load-reduction capability due to facility maintenance or down time. They schedule non-performance days. In essence, they provide the IESO a day-ahead commitment of resources available that enable the IESO to better operate the system and schedule alternate resources, if needed. As is discussed later, scheduled non-performance was relatively high in comparison to historical patterns in 2009 in part due to the effect of the economic downturn on direct participants. Overall, the scheduled non-performance explains 70% of the gap between contracted load reduction and the load reduction estimates (29.8 MW out of 42.7 MW). Importantly, the scheduled non-performance was factored into operations.
The remainder of the gap is explained almost entirely by settlement baseline bias rather than participant non-performance. In practice, aggregators and direct participants typically manage load based on the baseline in order to meet their contractual obligation. For settlement purposes, compliance is determined by a standard baseline method. However, when the baseline is biased, participants might be in full compliance with program rules while providing load reductions that are less than the contracted load reduction.

The 2008 DR-3 evaluation assessed the accuracy of the settlement baselines for participants with 98.9 MW of contracted load reduction, or 76% of the maximum load reduction dispatched in 2009, which was 129.9 MW. The analysis showed that DR-3 typically met or exceeded their day-ahead commitments as determined by settlement baseline. However, upward bias in the standard settlement method led to deviations between the load reductions estimated for the evaluation and day-ahead commitments.

With baseline methods, load reductions are calculated as the difference between the baseline method prediction and the metered load for an event day. As a result, bias in the settlement baseline translates into larger biases in the estimated reductions. The smaller the percent load reduction the more the bias is magnified. For example, if the true event period load and load reductions are 100 MW and 20 MW, a 5% upward bias in the settlement baseline will lead to an estimates reference load of 105 MW and a calculated load reduction of 25 MW (105 MW minus the metered load of 80 MW). While the baseline upward bias is 5%, the baseline estimated reductions are biased upwards by 25%, which means 25 MW of load reduction was estimated rather than the actual 20 MW. If instead the true load reduction were lower, say 10 MW (10%), the settlement method would estimate a load reduction of 15 MW (105 MW minus the metered load of 90 MW). With the smaller load reduction, the 5% upward bias in the settlement baseline leads to a 50% upward bias in the estimated reductions—15 MW of load reduction was estimated rather than the actual 10 MW.

The 2008 evaluation found a 6.9% program level bias in the settlement baseline, which translated into a 20% bias in settlement baseline load reduction estimates. The bias in the settlement baseline was higher for direct participants (10.0%) than for aggregators (4.5%). The standard baseline produces an upward bias for customers that are not weather sensitive, particularly if they are also sensitive to wholesale market prices. The 2008 evaluation load reductions were 86% of the day ahead committed load reductions because of baseline bias. The 2009 evaluation load reduction estimates are 82% of the day-ahead commitments.

---

8 Ibid., p. 52
We are confident that the gap between the 2009 evaluation load reductions and day-ahead scheduled load reduction is still primarily due to settlement baseline upward bias. Based on the 2008 baseline accuracy analysis, we know settlement baselines are upwardly biased for over 75% of the maximum DR-3 resourced dispatched in 2009. While the program has grown, a substantial share of the growth occurred after the 2009 events and the industry mix remains similar to 2008. The 2008 settlement accuracy analysis found a 20% upward bias in reductions estimates using settlement baselines. If we assume the upward bias is similar in 2009, the baseline method error explains 27% of the gap between contracted MW and load reduction estimates (11.6 MW out 42.7 MW).

The remaining unexplained portion of the gap is less than 2% of the contracted MW and is smaller than the uncertainty of the load reduction estimates. It is potentially due to participant non-performance, though this cannot be fully determined. DR-3 participants continue to typically meet or exceed their day-ahead contractual obligations which are measured by settlement baselines. However, the baseline estimate themselves are biased.

Figure 6-3 summarizes the gap between the contracted MW and estimated load reductions and the factors that explain it.
6.6 Load Reductions for Operations

For operations, a key aspect of DR-3 is the reliability of the day-ahead commitments. For a generator, once it is committed, whether it can deliver its resources is a function of whether it is in operation and power flow. In general, the generator is either operating or experiences an unforeseen forced outage. It either meets the day-ahead commitment in full or not at all. In contrast, DR-3 load reductions can deviate from the scheduled resources and the resource delivered is estimated rather than directly observed.

As noted above, a primary reason why DR-3 load reductions have deviated from the day-ahead scheduled load reductions in the past is an error in the settlement baselines—or, more specifically—the lack of adjustments in day-ahead commitments to account for upward bias in the baseline method. The contracted load reduction values should not be used for day-ahead commitments unless the upward baseline bias is eliminated. Until upward baseline bias is corrected, the day-ahead contracted load reductions should be de-rated for baseline error prior to day-ahead commitments. Doing so will ensure that DR-3 indeed meets its day-ahead resource commitments.

Completely eliminating baseline bias is difficult because no single baseline provides the best fit for all contributors. One alternative is to modify the current standard baseline in order to minimize bias given the current participant mix. Such an approach improves overall accuracy in the short term but may need recalibration as the program grows and the participant mix changes. Since the program remains highly concentrated among direct participants, another alternative is to standardize a baseline accuracy and selection process for direct participants rather than rely on the standard baseline.

In addition to adjustments for settlement baseline bias, day-ahead commitments need to take into account uncertainty due to variation in participant performance and statistical uncertainty. These three factors were combined based on historical data. Table 6-3 summarizes the amount of load reduction that can be expected for operations, along with the 90%–80% confidence intervals. It both adjusts for settlement baseline bias and participant performance and incorporates the event day uncertainty of reductions. The table allows OPA or the IESO to determine the amount of DR-3 resources for day-ahead commitments after factoring out scheduled non-performance. For example, if 120 MW of the contracted resources are scheduled to perform, de-rating for the bias inherent in the settlement baseline and average participant performance reduces the day-ahead commitment to 100 MW. If the statistical and

---

*The distribution that best described the historical variation in performance – weighted by the day-ahead MW commitment – was assessed through three distributional goodness of fit measured. Once that was accomplished the statistical uncertainty was incorporated via Monte Carlo simulation.*
event day performance uncertainty are incorporated, we can expect between 73.0 MW and 128.0 MW to be delivered with 90% confidence.

Figure 6-4 summarizes the relationship between the day-ahead contracted MW and uncertainty if baseline error is eliminated by adjusting the baseline method. As the day-ahead contracted MW scheduled to perform increases, so do the confidence bands.

**Figure 6-4:**
DR Load Reduction for Operations with Corrected Settlement Baseline

![Graph showing the relationship between Day-Ahead Contracted MW and Load Reduction Delivered (MW). The graph includes lines for 90% and 80% confidence intervals.](image)
Table 6-3:
Load Reduction Expected for Operations with Confidence Intervals
(Includes Adjustments for Baseline Error and Participant Performance)

<table>
<thead>
<tr>
<th>Day Ahead Contracted MW</th>
<th>Expected Delivered MW</th>
<th>90% Confidence Interval</th>
<th>80% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower</td>
<td>Upper</td>
<td>Lower</td>
</tr>
<tr>
<td>0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>10</td>
<td>8.4</td>
<td>6.1</td>
<td>10.7</td>
</tr>
<tr>
<td>20</td>
<td>16.7</td>
<td>12.2</td>
<td>21.3</td>
</tr>
<tr>
<td>30</td>
<td>25.1</td>
<td>18.2</td>
<td>32.0</td>
</tr>
<tr>
<td>40</td>
<td>33.5</td>
<td>24.3</td>
<td>42.7</td>
</tr>
<tr>
<td>50</td>
<td>41.9</td>
<td>30.4</td>
<td>53.3</td>
</tr>
<tr>
<td>60</td>
<td>50.2</td>
<td>36.5</td>
<td>64.0</td>
</tr>
<tr>
<td>70</td>
<td>58.6</td>
<td>42.6</td>
<td>74.6</td>
</tr>
<tr>
<td>80</td>
<td>67.0</td>
<td>48.6</td>
<td>85.3</td>
</tr>
<tr>
<td>90</td>
<td>75.3</td>
<td>54.7</td>
<td>96.0</td>
</tr>
<tr>
<td>100</td>
<td>83.7</td>
<td>60.8</td>
<td>106.6</td>
</tr>
<tr>
<td>110</td>
<td>92.1</td>
<td>66.9</td>
<td>117.3</td>
</tr>
<tr>
<td>120</td>
<td>100.5</td>
<td>73.0</td>
<td>128.0</td>
</tr>
<tr>
<td>130</td>
<td>108.8</td>
<td>79.0</td>
<td>138.6</td>
</tr>
<tr>
<td>140</td>
<td>117.2</td>
<td>85.1</td>
<td>149.3</td>
</tr>
<tr>
<td>150</td>
<td>125.6</td>
<td>91.2</td>
<td>159.9</td>
</tr>
<tr>
<td>160</td>
<td>133.9</td>
<td>97.3</td>
<td>170.6</td>
</tr>
<tr>
<td>170</td>
<td>142.3</td>
<td>103.4</td>
<td>181.3</td>
</tr>
<tr>
<td>180</td>
<td>150.7</td>
<td>109.5</td>
<td>191.9</td>
</tr>
<tr>
<td>190</td>
<td>159.1</td>
<td>115.5</td>
<td>202.6</td>
</tr>
<tr>
<td>200</td>
<td>167.4</td>
<td>121.6</td>
<td>213.3</td>
</tr>
<tr>
<td>210</td>
<td>175.8</td>
<td>127.7</td>
<td>223.9</td>
</tr>
<tr>
<td>220</td>
<td>184.2</td>
<td>133.8</td>
<td>234.6</td>
</tr>
<tr>
<td>230</td>
<td>192.5</td>
<td>139.9</td>
<td>245.2</td>
</tr>
<tr>
<td>240</td>
<td>200.9</td>
<td>145.9</td>
<td>255.9</td>
</tr>
<tr>
<td>250</td>
<td>209.3</td>
<td>152.0</td>
<td>266.6</td>
</tr>
<tr>
<td>260</td>
<td>217.7</td>
<td>158.1</td>
<td>277.2</td>
</tr>
<tr>
<td>270</td>
<td>226.0</td>
<td>164.2</td>
<td>287.9</td>
</tr>
<tr>
<td>280</td>
<td>234.4</td>
<td>170.3</td>
<td>298.6</td>
</tr>
<tr>
<td>290</td>
<td>242.8</td>
<td>176.3</td>
<td>309.2</td>
</tr>
<tr>
<td>300</td>
<td>251.2</td>
<td>182.4</td>
<td>319.9</td>
</tr>
</tbody>
</table>

Contracted MW less scheduled non-performance
6.7 Load Reductions for Long-term Planning and Cost-Effectiveness Analysis

The *ex ante* load reductions are designed to provide an overview of load reductions that can be used for planning purposes. They reflect the actual load reductions that can be expected, by month, under normal and extreme weather year system peaking conditions. These estimates factor in both scheduled non-performance and the extent to which participant load reductions deviate from day-ahead contracted resources. Since the data is limited to 2008 and 2009, so is the data on the extent of non-performance days and the extent to which customers delivered the expected load reductions. The *ex ante* load reductions are designed to provide an overview of load reductions that can be used for planning purposes. They reflect the actual load reductions that can be expected, by month, under normal and extreme weather year system peaking conditions. These estimates factor in both scheduled non-performance and the extent to which participants reliably deliver the load reduction expected by the system operator.

In the case of DR-3, the participants are not weather sensitive but the larger participants are exposed to IESO wholesale market prices. Several of them are price-responsive. OPA currently does not have a forecast of market prices under normal and extreme system conditions. To a large extent, the wholesale market prices are not necessarily relevant since the load reduction is contractual. As a result, the *ex ante* load reductions factor in the amount of contractual load response and the degree to which these customers have historically delivered those load reductions. In other words, they are de-rated for non-performance days and for deviations from the load reduction expected by system operators (which factor in scheduled non-performance).

Figure 6-5 shows the average weekday scheduled performance—contracted load reductions minus scheduled non-performance—for each month since the program started. Because the *ex ante* estimates factor in non-performance days during the economic downturn, they are likely conservative.
On average, 91.7% of the contracted MW was available for operations. However, the share of available load reduction had a distinctive drop in July and August of 2009, precisely the months with the highest likelihood of high system loads. Direct participants accounted for the bulk of scheduled non-performance. As a group, direct participants averaged 26.2 and 40.6 MW of scheduled non-performance in July and August of 2009, respectively. In over 80% of the scheduled non-performance days, these customers shut down their main industrial processes throughout the day. The large direct participants often schedule non-performance days because their expected load for the following day is not sufficient to provide their contracted load reduction. Because usage of the large direct participants decreased significantly in July and August of 2009, the amount of non-performance days increased. The historical data for these customers in 2006–2008 does not reflect similar shut down rates in July and August or, for that matter, any seasonal pattern. As a result, it may be better to assume that the percentage of load scheduled for performance is constant across months until more data is available. The effect of direct participant scheduled non-performance on the program highlights the need to continue to diversify the program and reflects the higher sensitivity of current direct participants to economic cycles.

While ideally, the expected amount of load reduction resources available for operations would reflect any seasonal patterns in performance under event conditions, the history of events is limited. In using ex ante load reduction estimates for planning purposes, it is critical to incorporate the time-varying information on DR-3 as a resource, including both scheduled non-

![Figure 6-5: Average Weekday Historical Scheduled Performance by Month](image)
performance and departures from the expected load reductions for operations. In particular, it is necessary to estimate the ability of DR-3 resources to effectively increase the capacity available without increasing the likelihood of unserved electricity.

Table 6-4 shows the *ex ante* load reduction estimates. It summarizes the contractual load reductions, the expected day ahead contracted resources, and the load reductions to be employed for planning purposes. The *ex ante* estimates assume scheduled non-performance does not have a seasonal pattern.

### Table 6-4: *Ex ante* Load Reduction Estimated by Month

<table>
<thead>
<tr>
<th>Month</th>
<th>Contracted MW</th>
<th>Day Ahead Contracted MW</th>
<th><em>Ex ante</em> Load Reductions MW</th>
<th>90% Confidence Interval</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>167.0</td>
<td>153.1</td>
<td>128.2</td>
<td>93.1</td>
<td>163.2</td>
<td></td>
</tr>
<tr>
<td>February</td>
<td>167.0</td>
<td>153.1</td>
<td>128.2</td>
<td>93.1</td>
<td>163.2</td>
<td></td>
</tr>
<tr>
<td>March</td>
<td>167.0</td>
<td>153.1</td>
<td>128.1</td>
<td>93.1</td>
<td>163.2</td>
<td></td>
</tr>
<tr>
<td>April</td>
<td>167.2</td>
<td>153.2</td>
<td>128.3</td>
<td>93.2</td>
<td>163.4</td>
<td></td>
</tr>
<tr>
<td>May</td>
<td>167.4</td>
<td>153.4</td>
<td>128.4</td>
<td>93.3</td>
<td>163.6</td>
<td></td>
</tr>
<tr>
<td>June</td>
<td>168.2</td>
<td>154.2</td>
<td>129.1</td>
<td>93.8</td>
<td>164.4</td>
<td></td>
</tr>
<tr>
<td>July</td>
<td>168.2</td>
<td>154.2</td>
<td>129.1</td>
<td>93.8</td>
<td>164.4</td>
<td></td>
</tr>
<tr>
<td>August</td>
<td>168.2</td>
<td>154.2</td>
<td>129.1</td>
<td>93.8</td>
<td>164.4</td>
<td></td>
</tr>
<tr>
<td>September</td>
<td>168.1</td>
<td>154.1</td>
<td>129.0</td>
<td>93.7</td>
<td>164.3</td>
<td></td>
</tr>
<tr>
<td>October</td>
<td>167.5</td>
<td>153.5</td>
<td>128.5</td>
<td>93.4</td>
<td>163.7</td>
<td></td>
</tr>
<tr>
<td>November</td>
<td>167.4</td>
<td>153.5</td>
<td>128.5</td>
<td>93.3</td>
<td>163.6</td>
<td></td>
</tr>
<tr>
<td>December</td>
<td>167.1</td>
<td>153.1</td>
<td>128.2</td>
<td>93.1</td>
<td>163.3</td>
<td></td>
</tr>
<tr>
<td>AVERAGE</td>
<td>167.5</td>
<td>153.6</td>
<td>128.6</td>
<td>93.4</td>
<td>163.7</td>
<td></td>
</tr>
</tbody>
</table>

The contracted resources are relatively constant and reflect aggregator and direct participant estimates of the extent to which their load reduction capability varies. The de-rating associated with the long-term reductions for planning factored in scheduled non-performance, adjustments for settlement baseline error, and participant performance. On average, contractual resources are de-rated 7.3% for historical non-performance, and an additional 16% for settlement baseline error. Overall, 76.7% of the contractual resources is expected for planning. There is 90% confidence that 55.7%–97.7% of the DR-3 contractual resources are available for long-term planning.
As noted in earlier, while the *ex ante* load reduction estimates factor in the scheduled non-performance and departures from the expected load reductions for operations, they do not directly account for the timing and magnitude of load reductions on effective load carrying capacity. From a capacity perspective, resources available in the summer mid-afternoon hours, the hours of system peak, generally have higher value than resources in non-peak hours. It is necessary to incorporate the extent to which DR-3 resources can provide capacity without increasing the likelihood of unserved electricity. This is typically done for cost-effectiveness analysis and long-term planning.

### 6.8 Conclusions and Recommendations

The 2009 DR-3 evaluation adds a substantial amount of information about participant performance in terms of system operations and long-term planning. Despite the wealth of additional information, any conclusions need to factor in the unique 2009 economic conditions and that the customer mix will likely continue to diversify.

While DR-3 resources have diversified, they remain highly concentrated, with five participants accounting for approximately 42% of the contracted MW. There was a substantial amount of scheduled non-performance days in 2009, and the direct participants accounted the bulk of it. For the average weekday, 9.6% of the contracted MW was scheduled for non-performance. That being said, participants consistently delivered the load reduction expected by the system operator. After scheduled non-performance is accounted for, on average, participants delivered 81.2% of the day ahead contracted resources in 2009. Factoring in 2008, participants delivered 84%. In 2008, the gap between the estimated load reductions and day ahead contracted resources was entirely explained by settlement baseline error and the fact that many participants manage load based on the baseline. We confident this remains the case in 2009.

Key recommendations resulting from the evaluation include:

- Update the analysis of settlement baseline accuracy and extend it to model the effect of baseline errors on settlement payments.
- Until the baseline error is corrected, adjust day-ahead commitments to account for the settlement baseline error.
- Ensure direct participants fully understand their load reduction capabilities and usage patterns, particularly on-peak load, prior to engaging in contractual load reduction commitments.
- Update estimates used for operation and planning with additional data on DR-3 events. Specifically, refine estimates of the share of expected load reduction participants deliver and scheduled non-performance days.
• Further diversify the program. Doing so will not only make the program more robust to economic downturns, it will also help reduce the statistical uncertainty and provide more precise reduction estimates.

• Consider transitioning large direct participants to another program that relies less so or not at all on settlement baselines. Throughout North America most programs and contracts that provide availability payments either operate as an interruptible program wherein participant reduce loads to a pre-specified level or, alternative, rely substantially on aggregation to reduce volatility in operational reliability.
APPENDIX A  DR-1 DETAILED METHODOLOGY AND VALIDATION

The load reduction analysis estimated the incremental load reductions provided by participants in 2009 that were attributable to the DR-1 program. All participants already faced wholesale market prices for their electric commodity and, in general, were in the habit of decreasing their load during high priced periods in the absence of DR-1 incentives. While real, these price-induced load reductions should not be attributed to the DR-1 program.

Although the load reduction analysis explored the drivers of participant behavior and variation in load reductions, the nature of the program and its participant mix posed several challenges to interpreting the results. The voluntary nature of the program required an analysis of load response patterns rather than a focus on a specific set of events. As a result, the regressions accurately reflect the average load response across events and capture systematic variation across different load response drivers such as market prices and hour of day. However, they do not and cannot accurately estimate the load reduction for each event or each event hour.

For DR-1, it also difficult to develop estimates for day types representing \textit{ex ante} event conditions. DR-1 is called when prices exceed a certain threshold. Historically, the frequency of eligible hours has varied from close to 2,000 hours per year to 132 hours in 2009. For such a program, defining normal and extreme year conditions solely based on weather would be inaccurate due to strong correlations between weather, market prices, event frequency, and participant load response. In short, there is insufficient market and event history to fully define \textit{ex ante} load conditions for DR-1. Lastly, because load response for DR-1 is tied to event frequency, duration and avoided wholesale market costs, modifying the strike price would not only reduce the number of event hours but could fundamentally influence participation and overall load response. This is a particularly perplexing analysis problem for which there are no easy solutions.

A.1. Load Reduction Regression Model Development

Customer load patterns were carefully analyzed using regression methods. Regression techniques have several distinct advantages over the day-matching or baseline methods typically used for DR program settlement. First, they can help identify the key drivers and predictors of load patterns and load reductions. Second, regression results provide more robust estimates of load reductions and are not as sensitive to biases in the reference load. Third, they can be used to predict load reductions for operations or for long term planning by factoring in expected weather, system, and market conditions, if available. In other words, they can better predict load reductions on a day-of or day-ahead basis in order to assist operations as well as predict load reduction resources under system peaking conditions in order to assist long term planning.
The high concentration of enrolled load and bid behavior in DR-1 was a significant factor in selecting the particular regression methods used to evaluate program reductions. All of the four largest customers were in the same industry and had similar processes and load patterns. Their variation in load was step-like, with load levels clustered around specific values (e.g., low, medium, or high) that reflected different industrial processes in operation.

The load usage patterns and load reductions were estimated using 2008 and 2009 hourly demand data for both curtailment and non-curtailment periods. Individual customer regressions were developed for each participant. The regressions were developed with the primary goal of accurately estimating *ex post* load reductions and energy use patterns. The focus was primarily on prediction accuracy and on the robustness of the variables reflecting load reductions. Because of the variation in bids submitted by individual customers, the bid amounts were used as variables to predict the expected load reduction for a given event period. This method allowed the analysis to incorporate participant projections of load reduction and assess the extent to which participants exceeded or failed to meet those short term estimates.

The individual customer models consisted of four sets of variables:

- Hourly and day-of-week variables that identified the average load shape of customers absent curtailments;
- Variables designed to explain variation in load patterns, including hourly prices, weekly price patterns and seasonal effects;
- Variables representing prior participant load were used to predict current load (formally known as dynamic lags), including the same hour of the prior day, and the average same hour load for non-event days in the prior four weeks; and
- Variables designed to quantify the average load reductions and variation in load reductions, including curtailment bid amounts and interactions with onset time, hour of day, and time elapsed since event start.

Several potential explanatory variables were tested for their effect on performance relative to the bid amount, including overall market demand, market prices and weather. The final regression equations are shown below and predict hourly demand on event and normal days relatively well.
\[ MW_i = a + \sum_{j=1}^{24} b_{ij} \times \text{hour}_i \times \text{daytype}_j + \sum_{j=2}^{12} c_i \times \text{month}_j \]
\[ + \sum_{i=1}^{24} d_i \times \text{hour}_i \times \text{bidamount}_i + e \times \text{bidamount}_i \times \text{LnEventDuration}_i \]
\[ + f \times \text{L24MW}_i + \sum_{i=1}^{24} g_i \times \text{hour}_i \times \text{TwoWeekAvgMW}_i + h \times \text{DayLagThreeHour}_i \]
\[ + k \times \text{DayLagSystemLoad}_i + l \times \text{TwoWeekAvgThreeHour}_i + m \times \text{TwoWeekAvgSystemLoad}_i \]
\[ + \sum_{i=1}^{24} n_i \times \text{hour}_i \times \text{DailyPriceRatio}_i + o \times \text{DailyAvgThreeHour}_i \]
\[ + \sum_{i=1}^{24} p_i \times \text{hour}_i \times \text{CDD}_i + \sum_{i=1}^{24} q_i \times \text{hour}_i \times \text{HDD}_i + \sum_{i=1}^{24} r_i \times \text{hour}_i \times \text{CDDsqr}_i + \sum_{i=1}^{24} s_i \times \text{hour}_i \times \text{HDDsqr}_i \]
\[ + u_i \]

Table A-1 defines the regression variables and describes the effects they seek to identify.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MW</td>
<td>Participants’ estimated hourly energy usage</td>
</tr>
<tr>
<td>A</td>
<td>An estimated constant</td>
</tr>
<tr>
<td>b-s</td>
<td>Estimated parameters</td>
</tr>
<tr>
<td>monthi</td>
<td>Dummy variables for each month, designed to pick up seasonal effects</td>
</tr>
<tr>
<td>daytypej</td>
<td>Dummy variables designed to pick up day-of-week effects. Days of the week are classified into five day types: Sunday, Monday, Tuesday through Thursday, Friday and Saturday</td>
</tr>
<tr>
<td>houri</td>
<td>Dummy variables designed to estimate the effect of variation across the hours of the day (i.e., operating schedule) on bid behavior</td>
</tr>
<tr>
<td>bidamount</td>
<td>MW bid amount submitted by a participant</td>
</tr>
<tr>
<td>LnDuration</td>
<td>The natural log of the length of an event in hours.</td>
</tr>
<tr>
<td>L24MW</td>
<td>A Lag variable representing a participant’s hourly energy use 24 hours previous</td>
</tr>
<tr>
<td>TwoWeekAvgMW</td>
<td>The two-week same-hour moving average of a participant’s hourly energy use</td>
</tr>
<tr>
<td>Variable</td>
<td>Description</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>DayLagThreeHour</td>
<td>The 24 hour lag of the market price of electricity, designed to capture the effect of short term changes in price patterns and levels on electricity consumption</td>
</tr>
<tr>
<td>DayLagSystemLoad</td>
<td>The 24 hour lag of system load, designed to capture the effect of short term changes in system load levels on electricity consumption</td>
</tr>
<tr>
<td>TwoWeekAvgThreeHour</td>
<td>The trailing, two-week, same-hour average market price of electricity, designed to capture the effect of sustained changes in price patterns and levels on electricity consumption</td>
</tr>
<tr>
<td>TwoWeekAvgSystemLoad</td>
<td>The trailing, two-week, same-hour average system load, designed to capture the effect of sustained changes in price patterns and levels on electricity consumption</td>
</tr>
<tr>
<td>DailyPriceRatio</td>
<td>The ratio of the three-hour dispatch price to the daily average of the three-hour dispatch price, designed to capture the substitution effect of hourly prices within a day</td>
</tr>
<tr>
<td>AvgDailyThreeHour</td>
<td>The daily average of the three-hour dispatch price, designed to capture the effect of daily levels of market prices on electricity consumption</td>
</tr>
<tr>
<td>CDD</td>
<td>Cooling degree days (defined as the maximum of 0 or the average daily temperature minus 65 F°), designed to reflect the impact of cooling load</td>
</tr>
<tr>
<td>CDDsqr</td>
<td>The square of CDD, designed to reflect nonlinearities in the relationship between temperature and bid behavior</td>
</tr>
<tr>
<td>HDD</td>
<td>Heating degree days (defined as the maximum of 0 or 65 F° minus the average daily temperature), designed to reflect the impact of heating load</td>
</tr>
<tr>
<td>HDDsqr</td>
<td>The square of HDD, designed to reflect nonlinearities in the relationship between temperature and bid behavior</td>
</tr>
<tr>
<td>U</td>
<td>The error term</td>
</tr>
</tbody>
</table>

### A.2. Assessment of Load Reduction Model Accuracy and Precision

In order to ensure that the regression-based load reductions were not an artifact of the final regression model, multiple regression specifications and techniques were tested and their effect on the reduction variables was carefully analyzed. For most participants, the coefficients of the load reduction variables did not vary much in magnitude or statistical significance with the inclusion or exclusion of explanatory variables. In other words, the load reduction estimates were robust. In addition, the regression based predictions were compared to actual load under both event and non-event conditions in order to assess the ability of the models to accurately predict load patterns and reductions.
Figure A-1 compares actual and predicted load for event hours. Because event periods vary across individual days, the comparison averages load across various days. As seen in Figure A-1, the regression predicted load mirrors the actual load during event days quite well. The graph also shows the hourly variation in the expected load reductions. The expected load reductions are larger during off peak hours than during system peak hours. This is related to the naturally occurring load shapes of industrial participants. In general, the aggregate participant load is higher during night hours, partly because the largest participants are responsive both to hourly price changes and to changes in wholesale market price patterns, adjusting their usage as peak price periods change between the winter and summer months and as overall market prices increase or decrease.

![Figure A-1: Comparison of Regression Predicted Load and Actual Load for Event Hours](image)

Figure A-2 shows the program level actual and regression predicted load for non-event hours. The two lines are nearly indistinguishable, indicating the model used to estimate reductions predicts load very well. Results vary slightly by day and at the individual customer level, but they generally mirror actual load.
As Figure A-3 shows, the regressions predict relatively accurately for the range of market prices and system load conditions.

Lastly, Figure A-4 shows the ability of the regression models to accurately predict load across temperature conditions.
Overall, the regressions explain 78% of the variation in aggregate participant load. While the explanatory power of the regressions varies from customer to customer, from a policy perspective, the performance of individual customer regressions is of less importance than how the regressions perform in aggregate and for specific customer segments. In general, it is more difficult to explain fully how a specific participant behaves on an hourly basis than it is to explain how the aggregate program load behaves on an hourly basis.
A.3. **Bid Behavior Regression Model Development**

Various estimation techniques were tested for the bid amount model. The non-normal distribution presents challenges and lends itself to more complex regression methodologies. For the largest customers, individual multinomial logit models were tested as their bids tend to be clustered highly around several values. This is because participants bid amounts that relate to specific processes they plan to shut down. A two part model was also tested, and proved more accurate than the multinomial logit approach. When running regressions on all participants’ bids instead of individual participants, the distribution of bid amounts begins to resemble the normal distribution, however the significant amount of zero values where participants chose not to bid prevents a single-stage model from being feasible. An accurate alternative approach is to first estimate a propensity to bid logit model to determine whether or not a participant would bid into an event. Nextly a GLM model designed to account for non-normal distributions was run to predict participants’ bid amounts given they did choose to enter a bid. Both regressions are structured to cluster estimated values by participant. By pooling the data like this, it was also possible to identify how customer characteristics related to participation levels. A participant’s estimated bid in any eligible hour is equal to his estimated propensity to bid multiplied by his estimated bid amount were he to bid.

The final regression model for bid decisions is shown below:

\[
\text{EventBid}_{ij} = a + \sum_{i=1}^{22} b_i \times \text{participant}_{ij} + \sum_{i=2}^{24} c_i \times \text{hour}_{ij} + \sum_{i=2}^{12} d_i \times \text{dayofweek}_{ij} + \sum_{i=2}^{12} e_i \times \text{month}_{ij} \\
+ f \times \text{CDD}_{ij} + g \times \text{CDDsqr}_{ij} + h \times \text{HDD}_{ij} + k \times \text{HDDSqr}_{ij} \\
+ l \times \text{LnMaxMW}_{ij} + m \times \text{Lnimplicitprice}_{ij} + n \times \text{LnDuration}_{ij} \\
+ o \times \text{LnActiveDays}_{ij} + p \times \text{LnDaysSince}_{ij} + q \times \text{TwoWeekAvg3Hour}_{ij} \\
+ r \times \text{TwoWeekAvg3Hoursqr}_{ij} + s \times \text{L24EligibleDay}_{ij} + v \times \text{RecentCurtailment}_{ij} \\
w \times \text{WeeklyEligibleHours}_{ij} + y \times \text{WeeklyEligibleHoursSqr}_{ij} + \text{participant}_{ij} + u_{ij}
\]

The final regression model for bid amount is as follows:

\[
\text{BidAmount}_{ij} = a + \sum_{i=1}^{22} b_i \times \text{participant}_{ij} + \sum_{i=2}^{24} c_i \times \text{hour}_{ij} + \sum_{i=2}^{12} d_i \times \text{dayofweek}_{ij} + \sum_{i=2}^{12} e_i \times \text{month}_{ij} \\
+ f \times \text{CDD}_{ij} + g \times \text{CDDsqr}_{ij} + h \times \text{HDD}_{ij} + k \times \text{HDDSqr}_{ij} \\
+ l \times \text{LnMaxMW}_{ij} + m \times \text{Lnimplicitprice}_{ij} + n \times \text{LnDuration}_{ij} \\
+ o \times \text{LnActiveDays}_{ij} + p \times \text{LnDaysSince}_{ij} + q \times \text{TwoWeekAvg3Hour}_{ij} \\
+ r \times \text{TwoWeekAvg3Hoursqr}_{ij} + s \times \text{L24EligibleDay}_{ij} + v \times \text{RecentCurtailment}_{ij} \\
w \times \text{WeeklyEligibleHours}_{ij} + y \times \text{WeeklyEligibleHoursSqr}_{ij} + \text{participant}_{ij} + u_{ij}
\]
Table A-2 defines the variables and describes the effects they seek to identify.

### Table A-2: Description of Bid Behavior Regression Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>a is an estimated constant</td>
</tr>
<tr>
<td>b-y</td>
<td>b-y are estimated parameters</td>
</tr>
<tr>
<td>monthi</td>
<td>Dummy variables for month of the year, designed to pick up seasonal effects</td>
</tr>
<tr>
<td>dayofweeki</td>
<td>Dummy variables designed to pick up day-of-week effects</td>
</tr>
<tr>
<td>houri</td>
<td>Dummy variables representing the hours of the day, designed to estimate the effect of operating schedule on bid behavior</td>
</tr>
<tr>
<td>participantx</td>
<td>Dummy variables that capture individual participants unique predisposition to bid</td>
</tr>
<tr>
<td>CDD</td>
<td>Cooling degree days (defined as the maximum of 0 or the average daily temperature – 65 F°), designed to reflect the impact of cooling load</td>
</tr>
<tr>
<td>CDDsqr</td>
<td>The square of CDD, designed to reflect nonlinearities in the relationship between temperature and bid behavior</td>
</tr>
<tr>
<td>HDD</td>
<td>Heating degree days (defined as the maximum of 0 or 65 F° - the average daily temperature), designed to reflect the impact of heating load</td>
</tr>
<tr>
<td>HDDsqr</td>
<td>The square of HDD, designed to reflect nonlinearities in the relationship between temperature and bid behavior</td>
</tr>
<tr>
<td>LnMaxMW</td>
<td>The natural log of a participant's maximum MW for the year, designed to estimate the effect of customer size on bidding behavior;</td>
</tr>
<tr>
<td>Lnimplicitprice</td>
<td>The natural log of the hourly implicit price, defined as the 3-hour ahead price plus the strike price during eligible hours, designed to capture the effect of DR-1 incentives and avoided costs on bidding behavior</td>
</tr>
<tr>
<td>LnDuration</td>
<td>The natural log of the length of the event in hours</td>
</tr>
<tr>
<td>LnActiveDays</td>
<td>The natural log of the number of days since a participant signed up for the DR-1 program, designed to reflect the effect of program experience on bid behavior</td>
</tr>
<tr>
<td>LnDaysSince</td>
<td>The natural log of the count of days since a participant’s last event took place, designed to capture the effect of event frequency on bidding behavior</td>
</tr>
<tr>
<td>TwoWeekAvg3Hour</td>
<td>The trailing, two-week, same-hour average market price of electricity, designed to estimate the effect of sustained changes in price patterns</td>
</tr>
<tr>
<td>Variable</td>
<td>Description</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>TwoWeekAvg3Hoursq</td>
<td>The square of the two-week, same-hour market price, designed to identify nonlinearities in the effect of sustained changes in price patterns and levels on bid behavior</td>
</tr>
<tr>
<td>L24EligibleDay</td>
<td>A binary variable designed to assess the effect of back-to-back event days</td>
</tr>
<tr>
<td>WeekEligibleHours</td>
<td>The sum of eligible hours over the trailing week, designed to capture whether frequent events lead to higher participation or participant fatigue</td>
</tr>
<tr>
<td>WeekEligibleHoursSqr</td>
<td>The square of the sum of eligible hours over the trailing week, designed to identify nonlinearities</td>
</tr>
<tr>
<td>RecentCurtailment</td>
<td>RecentCurtailment is a weighted variable from 0 - 1 that represents when the most recent curtailment by the customer occurred, up to one week previous.</td>
</tr>
<tr>
<td>U</td>
<td>The error term</td>
</tr>
</tbody>
</table>

### A.4. Assessment of Bid Behavior Model Accuracy and Precision

To assess the accuracy of the bid behavior regression models, the predicted bid amounts were compared with actual bid amounts. Despite variation in event participation and bid amounts, the regressions predicted the total load curtailment bids relatively well. The regression explains 67% of the variation in aggregate bid levels. In other words, the regression models perform significantly better than a simple average of aggregate bids.

Figure A-5 shows the accuracy of the regression models under different market price conditions. Figure A-6 shows how well the bid behavior regressions predict by temperature. In each case, the bid behavior regressions predicted actual bids relatively well. Despite the significant historical variation in bid amounts across eligible hours, the models enable OPA to estimate the expected participant bids with a relatively high degree of accuracy.
As anticipated, higher market prices are associated with larger bid amounts. This is partly because participants receive greater compensation for load reductions under DR-1 when prices are high. However, because almost all DR-1 participants are direct wholesale market participants, they also avoid the cost associated with consuming higher priced power. As such, they are motivated to reduce load not just because of the DR-1 program features, but also because of market conditions. This finding was further supported by the individual participant load reduction regressions, which also found that participants, particularly very large
participants, were responsive to both short-term market prices and sustained changes in market price patterns.

Another interesting finding is that participants bid higher load reductions when an event had been called on the previous day and when sustained high priced periods were experienced. A potential explanation is that, under those circumstances, participants were primed to anticipate events.

Finally, it was found that the level of event participation increased over time for some participants and decreased for others. This finding illustrates the value of DR-1 as a training ground for demand response. With experience, participants appear to have learned whether they were able to shift or reduce load and adjusted participation levels accordingly. The more frequent participants have since moved to DR programs that require a contractual commitment. On the opposite spectrum, several enrolled customers participated so infrequently that they were effectively non-participants. The changes in bid behavior also highlight that, without experience, customers’ ability to estimate their load response capabilities may be limited.
APPENDIX B  DR-2 DETAILED ANALYSIS METHODOLOGY AND VALIDATION

DR-2 is unique and present several challenges from an evaluation perspective. In order to calculate load reductions for DR-2, the participants load in the absence of load shifting – the counterfactual – must be estimated. For most DR programs, this is accomplished by using pre-enrollment data, observing behavior during non-event days (a within customer control), or through a control group. It is not possible to observe the naturally occurring facility behavior while customers are participating in DR-2 since the program is designed to produce load shifting on a daily basis. Given the limited number of facilities of similar size and industry, it is also not possible to develop an adequate control group.

In this instance, the pre-enrollment load contains a large number of days in which natural load patterns were disrupted due to participation in DR-1. However, the time span when customers participated in DR-1 includes many non-event periods and enables the estimation of naturally occurring load patterns through regression analysis. In addition, the naturally occurring facility behavior can be observed for the time periods prior to participation in any DR program.

Technically, this analysis approach is referred to as an interrupted time series – a quasi-experimental method. The key to an interrupted time series analysis is knowing the exact point at which an intervention occurred. An intervention can be an event day or a payment stimulating a change in behavior – in this case load shifting - after a specific point. Part of the challenge for DR-2 is the multiple interventions to participant load that occurred since 2007. In 2007 and 2008, they participated in DR-1 which altered their behavior during event days. From March to October 2009, they were engaging in load shift during the transition period. For November and December of 2009, there was a different intervention in place – namely, the DR-2 full program rules were in effect and the settlement baseline calculations were updated to incorporate the DR-1 evaluation data. A strong point of this evaluation is that the timing of each intervention (i.e., shift in DR participation) is well documented. The downside is that disentangling the effect of each of the program effects requires us to rely on historical participant that is distant from the evaluation period.

The key weakness of an interrupted time series design is potentially confounding program effects with other factors that happen to occur during the participation period. In this case, the threat is real. The economic downturn occurred at the same time as the DR-2 program launched, first in transition mode and, subsequently, in its final form. Anecdotal evidence from participant indicates that the economic downturn affected the DR-2 participant profoundly. In addition, all of the participants face wholesale market prices, which were substantially lower in 2009 than in prior years.
As is discussed in section 4.3.1, the careful analysis of the load data show that the key difference between participant loads prior to and during DR-2 participation was the amount of days facilities were shut down for the day. While enrolled in DR-2, participants provided the IESO notice of non-performance days. These were typically, though not always, days in which the facility was shut down. For the pre-enrollment time frame, days when the facility was shut down can be identified by their load levels. In order to provide comparability across the changing economic conditions, day in which the facilities were shut down were excluded from DR-2 transition and enrollment period as well as from the pre-enrollment data. Facility shut downs are factored into the long term planning reduction estimates by de-rating based on the historical frequency of non-performance days.

The regressions provide the hourly reductions of the program but the load shifting to off-peak periods does not always match the reduction during the peak period. This reflects the fact that some customers cannot fully shift their load reductions to off-peak periods due to the physical capacity of the plant. To obtain the amount of load shifting, we had to normalize the pre-enrollment, transition, and enrollment results. This process is simple and, in essence, subtracts out differences in overall consumption and allows us to focus solely on changes in the load shape. By construction, both the reference load shape and the transition period load shapes are directly comparable in terms of consumption.

B.1. Regression Model Development
The load reductions of the program were analyzed using regression analysis. The load usage patterns and load reductions were estimated using 2006 through 2009 hourly demand data for pre-enrollment, transition, and enrollment period. As noted earlier, plan shut downs were excluded from the analysis and are factored through de-rating for the frequency of non-performance days. Individual customer regressions were developed for each participant. The regressions were developed with the primary goal of accurately estimating ex post load reductions and energy use patterns. The focus was primarily on prediction accuracy and on the robustness of the variables reflecting load reductions.

Variation in the demand for the three participants was step-like, with load levels clustered around specific values that reflect different industrial processes in operation. This creates a unique data pattern as the variation in demand is not continuous. The expected load level is best reflected by the probability that one, two, or multiple processes are operating jointly. As a result, a specific regression technique, ordinal probit regression, was applied for these customers. The approach essentially predicts the likelihood that the facility is at specific demand levels. It also allows us to identify the most likely facility load level and produce the expected load by factoring the likelihood of each demand level.
At a high level, the regression models can be thought of in terms of three main components:

- **Variables that reflect the average load shape of customers (load shape variables).** These are typically accounted by various shape indicator variables by hour of day and day of week;
- **Variables that explain deviation in hourly usage from the average load shape:** These include factors such as economic conditions, wholesale market prices, temperature, and seasonality; and
- **Variables that estimate the hourly load shifting while participants were on DR-2.** These variables are designed to estimate the hourly effects of the DR-2 program throughout the transitioning and enrollment period. Separate reductions were estimated for each month and hour of weekdays.

Mathematically, the model can be expressed by the following equation:

\[
MW_i = a + \sum_{j=2}^{24} b_{ij} \cdot \text{hour}_i \cdot \text{daytype}_j + \sum_{i=2}^{24} c_i \cdot \text{month}_i
\]

\[
+ \sum_{i=1}^{24} d_i \cdot \text{Daily PriceRatio}_i + e \cdot \text{Daily Avg Price}_i + f \cdot \text{Two Week Avg Price}_i
\]

\[
+ \sum_{i=1}^{24} g_i \cdot \text{hour}_i \cdot \text{CDD}_i + \sum_{i=1}^{24} h_i \cdot \text{hour}_i \cdot \text{HDD}_i + \sum_{i=1}^{24} k_i \cdot \text{hour}_i \cdot \text{CDDsqr}_i + \sum_{i=1}^{24} l_i \cdot \text{hour}_i \cdot \text{HDDsqr}_i
\]

\[
+ \sum_{i=1}^{24} m_i \cdot \text{hour}_i \cdot \text{TSX}_i + \sum_{i=1}^{24} n_i \cdot \text{hour}_i \cdot \text{TSXsqr}_i
\]

\[
+ \sum_{i=3}^{24} \sum_{j=2}^{24} o_{ij} \cdot \text{hour}_i \cdot \text{month}_j \cdot \text{transition period} + \sum_{i=11}^{24} \sum_{j=2}^{24} p_{ij} \cdot \text{hour}_i \cdot \text{month}_j \cdot \text{enrolled period}
\]

\[+ u_i\]

Table B-1 defines the regression variables and describes the effects they seek to identify.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MW</td>
<td>Participants’ estimated hourly energy usage</td>
</tr>
<tr>
<td>a</td>
<td>An estimated constant</td>
</tr>
<tr>
<td>b-p</td>
<td>Estimated parameters</td>
</tr>
<tr>
<td>month_i</td>
<td>Dummy variables for each month, designed to pick up seasonal effects</td>
</tr>
<tr>
<td>Variable</td>
<td>Description</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>daytype&lt;sub&gt;i&lt;/sub&gt;</td>
<td>Dummy variables designed to pick up day-of-week effects. Days of the week are classified into five day types: Monday, Tuesday through Thursday, Friday and weekends and holidays</td>
</tr>
<tr>
<td>hour&lt;sub&gt;i&lt;/sub&gt;</td>
<td>Dummy variables designed to estimate the effect of variation across the hours of the day (i.e., operating schedule)</td>
</tr>
<tr>
<td>DailyPriceRatio</td>
<td>The ratio of the three-hour dispatch price to the daily average of the three-hour dispatch price, designed to capture the substitution effect of hourly prices within a day</td>
</tr>
<tr>
<td>TwoWeekAvgPricer</td>
<td>The trailing, two-week, same-hour average market price of electricity, designed to capture the effect of sustained changes in price patterns and levels on electricity consumption</td>
</tr>
<tr>
<td>DailyAvgPrice</td>
<td>The daily average of the three-hour dispatch price, designed to capture the effect of daily levels of market prices on electricity consumption</td>
</tr>
<tr>
<td>CDD</td>
<td>Cooling degree days (defined as the maximum of 0 or the average daily temperature minus 65 F°), designed to reflect the impact of cooling load</td>
</tr>
<tr>
<td>CDDsqr</td>
<td>The square of CDD, designed to reflect nonlinearities in the relationship between temperature and electricity consumption</td>
</tr>
<tr>
<td>HDD</td>
<td>Heating degree days (defined as the maximum of 0 or 65 F° minus the average daily temperature), designed to reflect the impact of heating load</td>
</tr>
<tr>
<td>HDDsqr</td>
<td>The square of HDD, designed to reflect nonlinearities in the relationship between temperature and electricity consumption</td>
</tr>
<tr>
<td>TSX</td>
<td>The Toronto Stock Exchange index, designed to capture the effects of the overall economy on electricity consumption</td>
</tr>
<tr>
<td>TSXsqr</td>
<td>The Toronto Stock Exchange index squared, designed to reflect nonlinearities in the relationship between the TSX and electricity consumption</td>
</tr>
<tr>
<td>TransitionPeriod</td>
<td>The treatment effect of participation in DR-2's transitional period</td>
</tr>
<tr>
<td>EnrolledPeriod</td>
<td>The treatment effect of participation in DR-2's enrolled period</td>
</tr>
<tr>
<td>U</td>
<td>The error term</td>
</tr>
</tbody>
</table>

### B.2. Assessment of Model Accuracy and Precision

The accuracy of the regression models was assessed by testing the robustness of the load reduction coefficients and comparing regression predicted load with metered load.
In order to ensure that the regression-based load reductions were not an artifact of the final regression model, multiple regression specifications were tested and their effect on the reduction variables was carefully analyzed. The load reduction variables did not vary much in magnitude or statistical significance with the inclusion or exclusion of explanatory variables. In other words, the load reduction estimates were robust.

Figure B-1 compares actual load, regression-predicted load with DR, and regression predicted load without DR for the DR-2 transition and official contractual period in 2009. The predicted load with DR closely matches the metered load for both periods. The graphs exclude scheduled non-performance days. The program load for the contractual period (Nov-Dec 2009) is substantially lower than in the transition period. This simply reflects the fact that one of the direct participants shut down their facility for the last four months of 2009.

![Figure B-1: R-2 Comparison of Regression Predicted and Actual Load for 2009 Transition and Contractual Period](image)

Figure B-2 compares actual load and regression predicted load for pre-enrollment weekdays. The accuracy of the regression model with the pre-enrollment data is critical since it reflects the ability of the regressions to estimate load outside of DR-2 conditions and produce the counterfactual.
Figure B-2:  
DR-2 Comparison of Regression Predicted and Actual Load for Pre-Enrollment Weekdays

![Graph showing comparison of regression predicted and actual load for pre-enrollment weekdays.]

Figure B-3 shows the actual and regression-predicted load across the temperature spectrum for all days, including pre-enrollment, transition, and contract period. The regression predicts participant load accurately for both hot temperatures and for cold temperatures. The ability to accurately predict participant load during summer peaking conditions provides added confidence that the regression will accurately quantify summer load reductions. Figure C-4 shows how well the regression predicts load by market electricity prices and Ontario system demand.

Figure B-3:  
DR-2 Comparison of Regression Predicted and Actual Load By Temperature

![Graph showing comparison of regression predicted and actual load by temperature.]

Figure C-4 shows how well the regression predicts load by market electricity prices and Ontario system demand.
Figure B-4: 
DR-2 Comparison of Regression Predicted and Actual Load By Temperature

By Market Price

Ontario System Load

Program Load (MW)

3 Hour Ahead Electricity Price

Program Load (MW)

Ontario System Load

Actual MW

Predicted MW

Actual MW

Predicted MW
APPENDIX C  DR-3 DETAILED ANALYSIS METHODOLOGY AND VALIDATION

The DR-3 load reductions were estimated through regression methods using data from event and non-event days and the available pre-enrollment data. Individual participant regressions were developed for each of the 121 participants in 2008 and 2009 and aggregated to the settlement account level and to the program as a whole.

Regression methods were used because they have several distinct advantages over day matching or baseline methods. First, regression results provide more robust estimates of load reductions and are not as sensitive to biases in the reference load. Second, they can help identify the key drivers and predictors of participant load and load reduction behavior, including factors such as market prices and weather. Third, they can be used to predict load reductions for operations or for long term planning by factoring in expected weather, system, and market conditions, if available. In other words, they can better predict load reductions on a day-of or day-ahead basis in order to assist operations, as well as predict load reduction resources under system peaking conditions in order to assist long term planning. Finally, they are not subject to gaming behavior and can, in fact, detect gaming behavior.

The regression models were developed with the primary goal of accurately estimating ex post load reductions and energy use patterns. The focus was primarily on the accuracy of predictions and on the robustness of the variables that drive the load reductions.

C.1. Regression Model Development

Individual participant regressions were used rather than aggregate or panel models for several reasons. The high concentration of load and participation in DR-3 was a significant factor in the selection of individual participant regressions. Changes in the number and mix of participants, as well as differences in the available data, made it difficult to analyze aggregated data. Moreover, individual participant regressions can better address variation in the size, production schedules, and processes of participants.

A primary focus of the analysis was the actual load reductions relative to the contractual load reductions – the realization rate. The contractual load reductions did not vary for most participants. Comparing the event coefficients to contractual load reductions allowed the regressions to assess the extent to which participants exceeded or failed to meet contractual load reductions.

The final individual participant models included the following explanatory variables:

- Hour and day-of-week variables that identified the average load shape of participants in the absence of curtailment events;
- Month variables to capture seasonal variation in load;
- Hour and year variables that modeled how load has changed over the years in which data is available;
- Weather variables that model variation in load due to facility heating and cooling;
- The 3-hour ahead price and price ratio if the participant participated in the wholesale market;
- Pre-event and post-event variables that capture load shifting to non-event hours on event days;
- Event variables for the four event hours; and
- Variables representing prior participant load to predict current load including the average same-hour load for non-event weekdays in the prior four weeks.

The final model used to estimate load reductions is shown below:

$$ kW_i = a + \sum_{j=1}^{24} b_{ij} \times \text{hour}_i \times \text{daytype}_j + \sum_{i=1}^{24} c_i \times \text{month}_i + \sum_{j=1}^{2009} \sum_{i=1}^{24} d_{ij} \times \text{hour}_i \times \text{year}_j $$

$$ + \sum_{i=1}^{24} e_i \times \text{hour}_i \times \text{CDD}_i + \sum_{i=1}^{24} f_i \times \text{hour}_i \times \text{CDDsqr}_i $$

$$ + \sum_{i=1}^{24} g_i \times \text{hour}_i \times \text{HDD}_i + \sum_{i=1}^{24} h_i \times \text{hour}_i \times \text{HDDsqr}_i $$

$$ + \sum_{i=1}^{24} i_i \times \text{hour}_i \times \text{PRICE}_i + \sum_{i=1}^{24} j_i \times \text{hour}_i \times \text{PRICEsqr}_i $$

$$ + \sum_{i=1}^{24} k_i \times \text{hour}_i \times \text{PRICERATIO}_i + \sum_{i=1}^{24} l_i \times \text{hour}_i \times \text{PRICERATIONsqr}_i $$

$$ + \sum_{i=1}^{24} m_i \times \text{hour}_i \times \text{FourWeekkW}_i + \sum_{i=1}^{24} \sum_{j=1}^{21} n_{ij} \times \text{hour}_i \times \text{EventDay}_j + u_i $$

Table C-1 defines the regression variables and describes the effects they seek to identify.
### Table C-1: Description of Load Reduction Regression Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>kW</td>
<td>Participant’s estimated hourly energy usage</td>
</tr>
<tr>
<td>A</td>
<td>An estimated constant</td>
</tr>
<tr>
<td>b through n</td>
<td>Estimated parameters</td>
</tr>
<tr>
<td>hour_i</td>
<td>Dummy variables designed to estimate the effect of each hour of the day (i.e., operating schedule) on usage</td>
</tr>
<tr>
<td>daytype_j</td>
<td>Dummy variables designed to pick up day-of-week effects (Tues-Thurs are grouped together because these days are not different from each other)</td>
</tr>
<tr>
<td>month_i</td>
<td>Dummy variables for month of the year, designed to pick up seasonal effects</td>
</tr>
<tr>
<td>year_j</td>
<td>Dummy variables for each year in the estimating sample (2006 - 2009) to reflect trends in the general economy</td>
</tr>
<tr>
<td>CDD</td>
<td>Cooling degree days (defined as the maximum of 0 or average daily temperature minus 65 F°) designed to reflect the impact of cooling load</td>
</tr>
<tr>
<td>CDDsqr</td>
<td>The square of CDD designed to reflect nonlinearities in the relationship between temperature and usage</td>
</tr>
<tr>
<td>HDD</td>
<td>Heating degree days (defined as the maximum of 0 or 65 F° minus average daily temperature) designed to reflect the impact of cooling load</td>
</tr>
<tr>
<td>HDDsqr</td>
<td>The square of HDD designed to reflect nonlinearities in the relationship between temperature and usage</td>
</tr>
<tr>
<td>PRICE</td>
<td>3-hour ahead wholesale market price (only included if the participant is a wholesale market customer)</td>
</tr>
<tr>
<td>PRICEsqr</td>
<td>The square of the 3-hour ahead wholesale market price (only included if the participant is a wholesale market customer)</td>
</tr>
<tr>
<td>PRICERATIO</td>
<td>Ratio of price and the maximum price for the day to capture that customers tend to shift load to hours that are relatively less expensive for the day (only included if the participant is a wholesale market customer)</td>
</tr>
<tr>
<td>PRICERATIOsqr</td>
<td>The square of the price ratio variable (only included if the participant is a wholesale market customer)</td>
</tr>
<tr>
<td>FourWeekkW</td>
<td>Four-week moving average of a participant’s hourly energy usage</td>
</tr>
<tr>
<td>EventDay_j</td>
<td>Dummy variables that capture the impact of each of the 21 event days in the estimating sample</td>
</tr>
<tr>
<td>U</td>
<td>The error term</td>
</tr>
</tbody>
</table>
C.2. Assessment of Model Accuracy and Precision

The accuracy of the regression models was assessed by testing the robustness of the load reduction coefficients and comparing regression predicted load with metered load. In order to ensure that the regression-based load reductions were not an artifact of the final regression model, multiple regression specifications and techniques were tested and their effect on the reduction variables was carefully analyzed. For most participants, the load reduction variables did not vary much in magnitude or statistical significance with the inclusion or exclusion of explanatory variables. In other words, the load reduction estimates were robust. The largest participant exhibited variations of approximately 10% in the coefficients based on the variables included or excluded in the analysis. In particular, the relative performance increased with the inclusion of variables reflecting short-term economic conditions and the same hour average load over the prior two weeks. The variables reflecting short-term economic conditions captured changes in participant load patterns that were correlated with economic conditions. The four-week, same-hour average load helped identify seasonal changes in participant operating schedules. The final model included both of these variables because they improved model accuracy.

Figure C-1 compares actual load, regression-predicted load with DR, and regression predicted load without DR for the average event day in 2009. Due to the change in participant mix over the evaluation period, average participant load is shown instead of the aggregate program load. Although event timing varied across event days, event duration was the same (e.g., four hours) for all of the 2008 and 2009 events. As a result, it was possible to normalize the event data based on pre-event, post event and event hours.

While the regression predicted load shows a small bias, it mirrors actual load on event days relatively well during event and non-event hours. Importantly, the small bias does not affect the load reduction estimates, which are determined by the regression load reduction variable coefficients. It is also clear that participants typically shift load to pre-event hours. While DR-3 can be called on short notice, as a courtesy, participants were notified a day in advance of potential events.
Figure C-1: Comparison of Regression Predicted and Actual Load for Event Days in 2009

Figure C-2 compares the regression-predicted and actual participant load for non-event days. On average, the regressions predict participant load accurately, although the degree of accuracy may vary across individual days.

Figure C-3 shows the average participant actual and regression-predicted load across the temperature spectrum for all days, including pre-enrollment, event, and non-event days. The regression predicts participant load accurately for both hot temperatures and for cold temperatures. While events were not experienced during the summer period, the ability to accurately predict participant load during summer peaking conditions provides added confidence that the regression will accurately quantify summer load reductions with additional event history.

Figure C-4 compares the actual and regression-predicted average participant load under various wholesale market price conditions. The regressions predict actual load for the range of market price conditions, indicating they perform accurately for both normal and extreme market conditions.
Figure C-2: Comparison of Regression Predicted and Actual Load for Non-Event Days

Figure C-3: Comparison of Regression Predicted and Actual Load by Temperature
While the explanatory power of the regressions varies across participants, from a policy perspective, the performance of individual customer regressions is of less importance than how the regressions perform for the overall program and for specific participant segments. A commonly used metric for assessing precision of regressions is the R-squared statistic, which measures the degree to which the regression predictions improve on predictions based on a simple average. The R-squared metric measures the amount of variation around the mean explained by the regression.

Overall, the regressions explain 88% of the variation in aggregate participant load. For aggregator participants as a group, the regressions explain 94% of the variation. For direct participants, the regressions explain 48% of the variation. The combined R-squared value is larger than for individual customer regressions because, with aggregation, the variation in participant behavior is smoothed out. Put differently, it is easier to explain aggregate load patterns than individual participant load patterns. The R-squared value for direct participants is smaller than for aggregator participants because they are large wholesale market customers that have relatively constant load and exhibit little variation, making it more difficult to identify the few instances when their load patterns deviate from their usual pattern. Despite the lower R-squared values, the regression predictions for direct participants tend to be fairly accurate. Figure C-5 presents the distribution of the individual regression R-squared values.
Figure C-5: Histogram of R-Squared Values
## APPENDIX D  GLOSSARY

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contracted day-ahead resources:</strong></td>
<td>Contracted load reduction amounts with planned non-performance factored in</td>
</tr>
<tr>
<td><strong>Contracted load reduction capability (contracted MW):</strong></td>
<td>Load reduction amounts customers contract to provide. In the absence of non-performance, this is the amount by which they will aim to reduce load</td>
</tr>
<tr>
<td><strong>DR</strong></td>
<td>Demand Response programs provide load reductions or load shifting for specified hours in exchange for incentives. They include both dispatchable and non-dispatchable programs.</td>
</tr>
<tr>
<td><strong>DR-1</strong></td>
<td>OPA's voluntary demand bidding program</td>
</tr>
<tr>
<td><strong>DR-2</strong></td>
<td>OPA's permanent load shifting program</td>
</tr>
<tr>
<td><strong>DR-3</strong></td>
<td>OPA's event-based load reduction program</td>
</tr>
<tr>
<td><strong>Event window</strong></td>
<td>Consecutive hours when a program provides load reduction capability. In DR-1 event days and hours are determined by comparing strike prices to the pre-dispatch price. In DR-3 it is determined ahead of time by the IESO based on the day-ahead supply cushion. For DR-2, the event window is based on load shifting contracts with participants.</td>
</tr>
<tr>
<td><strong>Ex ante load reductions</strong></td>
<td>Load reduction capability under a standard set of conditions. For contractual resources, the ex-ante impacts factor in historical scheduled non-performance and historical deviations between day-ahead committed resources and actual load reductions delivered by participants.</td>
</tr>
<tr>
<td><strong>Ex post load reductions</strong></td>
<td>Actual historical load reductions. They reflect the event conditions, dispatch strategy and customer mix at the time. They do not necessarily reflect existing load reduction capability.</td>
</tr>
<tr>
<td><strong>Load reductions</strong></td>
<td>Changes in load attributable to participation in a given demand</td>
</tr>
</tbody>
</table>
response program

**Load reductions for operations**
Load reduction estimates used for program operations. The estimates factor in historical differences between day ahead committed resources and the load reductions they actually deliver.

**Load reductions for planning**
This term is synonymous with ex-ante load reductions. They provide load reduction capability under a standard set of conditions and are recommended for planning and cost-effectiveness analysis. The estimates factor in historical scheduled non-performance and historical deviations between day-ahead committed resources and actual load reductions delivered by participants.

**Load shifting**
Substitution of electric load from one period to another.

**Scheduled non-performance**
Planned periods where customers will not participate in demand response programs. The system operator is informed in advance of scheduled non-performance. Scheduled facility shutdowns for maintenance are an example.

**Settlement Baseline**
The method for determining payment for participants. Settlement baselines produce load reductions estimates that may differ from load reductions estimated in impact evaluations. They are less complex and are designed to quick settle payments. Importantly, if the settlement baseline is biased, participants may still be in compliance for settlement but actually deliver less load reduction than contracted.

**Shift window**
Consecutive hours during which customers participating in load shifting programs will reduce load, shifting it to hours outside the shift window.

**Unscheduled non-performance**
Deviations between day-ahead contracted resources and actual load reductions delivered. The gap between the committed resource and actual load reductions can be due to multiple factors including biased settlement baselines and underperformance by participants.