CIMM STUDY RELEASE:

Best Practices in Combining Smart TV and Set-Top Box Data

Howard Shimmel, Janus Strategy & Insights
CIMM Set Top Box-Smart TV ACR
Best Practices for Commingling

02.04.21
CIMM Study Overview

**Goal** - Assess strengths and weaknesses of Smart TV ACR and Set Top Box data to inform best practices for combining them at the household level to create granular nationally representative data sets for linear TV programming and advertising use cases

• Phase 1 – **Review of Smart TV ACR and Set Top Box providers** - collection of vital landscape statistics including sample size, data captured and reported, data processing rules

• Phase 2 – **Review of existing methods used to integrate Smart TV ACR and Set Top Box providers** - collection of detailed account of methods for integrating Smart TV ACR and Set Top Box data, covering matching methods at the device and household level as well as the co-mingled processing of viewing data
Feedback gathered from 18 entities

**MVPDs**
- attice
- NYI NEW YORK INTERCONNECT
- Major
- MVPD
- dish
- AT&T
- xandr
- pre-meditated media

**OEMs/ Smart TV ACR Providers**
- VIZIO
- Roku

**Third-party integrators**
- FOURTHWALL MEDIA
- SAMBA TV
- iSpot.tv
- alphonso
- comscore
- 6IX
- 5IVE
- TIVO
- gracenote
- videoamp
- JANUS

Phase 1 - 18 firms
Phase 2 - 9 firms
Background - The Case for Commingling

- Set Top Box and Smart TV ACR data sets have quickly gained an influential marketplace position as metric sources for planning, scheduling, stewardship and post evaluation of TV transactions.

- Demand for analytics that allow advertisers and agencies to precisely plan digital video and CTV on top of linear is rapidly accelerating.

- There is a growing industry need to improve TV measurement systems that supports the exploration and discovery of best practices to commingle these two complementary data sources.

<table>
<thead>
<tr>
<th>Complementary Audience Measurement</th>
<th>Set Top Box</th>
<th>Smart TV ACR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic representation</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Contiguous U.S. representation</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Sample Size</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>People Measurement</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Cross Device Measurement</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Delayed/non-live viewing</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>DVR, VOD</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Household-level match rate</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Reporting speed</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>CTV data collection</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Multi-set data capture</td>
<td>★</td>
<td>★</td>
</tr>
</tbody>
</table>
The universe for commingling Set Top Box and Smart TV ACR

- Nearly 42% of US homes are equipped with both Set Top Boxes and internet-enabled Smart TVs

Source: Nielsen, September 2020
Data set skews

There are unique demographic skews to the contributing data sets, and to the segments that are not captured in contributing data sets

<table>
<thead>
<tr>
<th></th>
<th>STB</th>
<th>ACR</th>
<th>Calibration</th>
<th>Calibration</th>
<th>ACR</th>
<th>Not Captured</th>
<th>ACR</th>
<th>Not Captured</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pay TV</td>
<td>Enabled Smart TV</td>
<td>Pay TV and Enabled Smart TV</td>
<td>Pay TV w/o Enabled Smart TV</td>
<td>Over-The-Air</td>
<td>Over-The-Air and Enabled Smart TV</td>
<td>Over-The-Air w/o Enabled Smart TV</td>
<td>Broadband Only</td>
</tr>
<tr>
<td>HOH Age Range &lt; 35</td>
<td>55</td>
<td>102</td>
<td>70</td>
<td>37</td>
<td>95</td>
<td>122</td>
<td>71</td>
<td>216</td>
</tr>
<tr>
<td>HOH Age Range = 35-54</td>
<td>96</td>
<td>121</td>
<td>117</td>
<td>69</td>
<td>110</td>
<td>130</td>
<td>91</td>
<td>126</td>
</tr>
<tr>
<td>HOH Age Range = 55+</td>
<td>120</td>
<td>84</td>
<td>98</td>
<td>146</td>
<td>94</td>
<td>70</td>
<td>117</td>
<td>39</td>
</tr>
<tr>
<td>HOH Origin = Hispanic</td>
<td>87</td>
<td>113</td>
<td>102</td>
<td>68</td>
<td>162</td>
<td>187</td>
<td>96</td>
<td>105</td>
</tr>
<tr>
<td>HOH Race = Black</td>
<td>95</td>
<td>96</td>
<td>95</td>
<td>95</td>
<td>130</td>
<td>113</td>
<td>146</td>
<td>82</td>
</tr>
<tr>
<td>HOH Race = White</td>
<td>106</td>
<td>102</td>
<td>104</td>
<td>108</td>
<td>93</td>
<td>94</td>
<td>91</td>
<td>101</td>
</tr>
<tr>
<td>Asian Household</td>
<td>63</td>
<td>91</td>
<td>78</td>
<td>45</td>
<td>73</td>
<td>87</td>
<td>60</td>
<td>136</td>
</tr>
<tr>
<td>Language Class = Spanish Dominant</td>
<td>72</td>
<td>91</td>
<td>76</td>
<td>68</td>
<td>210</td>
<td>210</td>
<td>210</td>
<td>56</td>
</tr>
<tr>
<td>HHLD Income = &lt;$75,000</td>
<td>92</td>
<td>79</td>
<td>72</td>
<td>117</td>
<td>128</td>
<td>105</td>
<td>149</td>
<td>94</td>
</tr>
<tr>
<td>HHLD Income = $75,000+</td>
<td>109</td>
<td>125</td>
<td>133</td>
<td>80</td>
<td>67</td>
<td>94</td>
<td>42</td>
<td>107</td>
</tr>
</tbody>
</table>

Source: Nielsen 9/28/20
Data Integration Pathway – Mapping Data Loss Example

As each integration layer is added, the number of viewing records enriched with ad and content, demographic and consumer data diminishes, with the original population potentially misrepresented.

**REPORTING ELIGIBILITY**

**MATCHING - DEMO**

**MATCHING - PRODUCT**

---

**ELIGIBLE**

Cable operators - 1%-3% homes don’t return data
Satellite - 50% of MVPD A homes without RPD
Consumer data opt-outs extremely low
Residential segment filtered by customer records

Customer record-to-household match rates 90%+ (-10%)

Net Drop-Off – 47%

For both, Set Top Box and Smart TV ACR, match rates 20% - 40% depend on consumer target coverage and match partner – Match rates for third-party data are likely to be higher than for first-party brand consumers.

---

Consumer data opt-outs - 10% - 20%
Unconnected to internet - NA
Residential segment filtered by IP match and algorithms

IP-to-household match rates 65% (-35%)

Net Drop-Off – 85%
STB/ACR Integration - Recommendation

Stage 1. Select data sets - situation analysis, key considerations, specific use cases

Stage 2. Establish commingling match cell design, example:
   - 1. matched STB/ACR, 2. unmatched STB, 3. unmatched ACR

Stage 3. Execute matches
   - Household (demographics)
   - Device match (STB to ACR)
   - Match validation

Stage 4. Calibration and Weighting
   - Calibrations - STB to ACR data; ACR to STB data
   - Weighting - Demographics, TV access universe, tuning metrics to all match cells
   - Scaling - Scaling ACR to network reach levels from STB data
   - Consolidation - combine all match cells to form one reporting data set

Stage 5. Validation
   - Universe estimates
   - Core tuning metrics to industry benchmark
Establish Match and Commingling Design

- Most commingled approaches use tuning data from two or more sources where device-to-device integration provides insights for calibrating the combined data set.

Two-source example: Three data sets underlie the combined data set:

1. Set Top Box only
2. Set Top Box/Smart TV ACR
3. Smart TV ACR only
Full Presentation and White Paper available on CIMM Website

www.cimm-us.org
CIMM STUDY DISCUSSION PANEL:
Lessons in Creating Scaled and Representative Granular TV Datasets

David Algranati, Comscore
Josh Chasin, VideoAmp
Caroline Horner, 605
Tom Weiss, MarketCast/Deductive
Gerard Broussard, Pre-Meditated Media