Presentation to W3C on Audience Measurement

Josh Chasin; Chief Research Officer| May 6, 2013

comSCORE.

Agenda

Placing audience measurement in a historical context

How audience measurement integrates census data (tags and cookies) into reported data

Historical Context

Media vehicle audiences are not, a priori, known

- How many people watched the TV broadcast?
- How many people listened to the radio station?
- How many people actually read this month's issue?
- How many people passed by the billboard?

The business of media, generally dependent on advertising, requires some consensus third party measure of audience size and composition, in order for that advertising to be bought and sold

Audience measurement emerged to fill this need

- First, based on surveys
- Over time, incorporating technology (i.e., TV meters)
- Ultimately incorporating what the industry calls "naturally occurring data" (e.g. TV STB data)
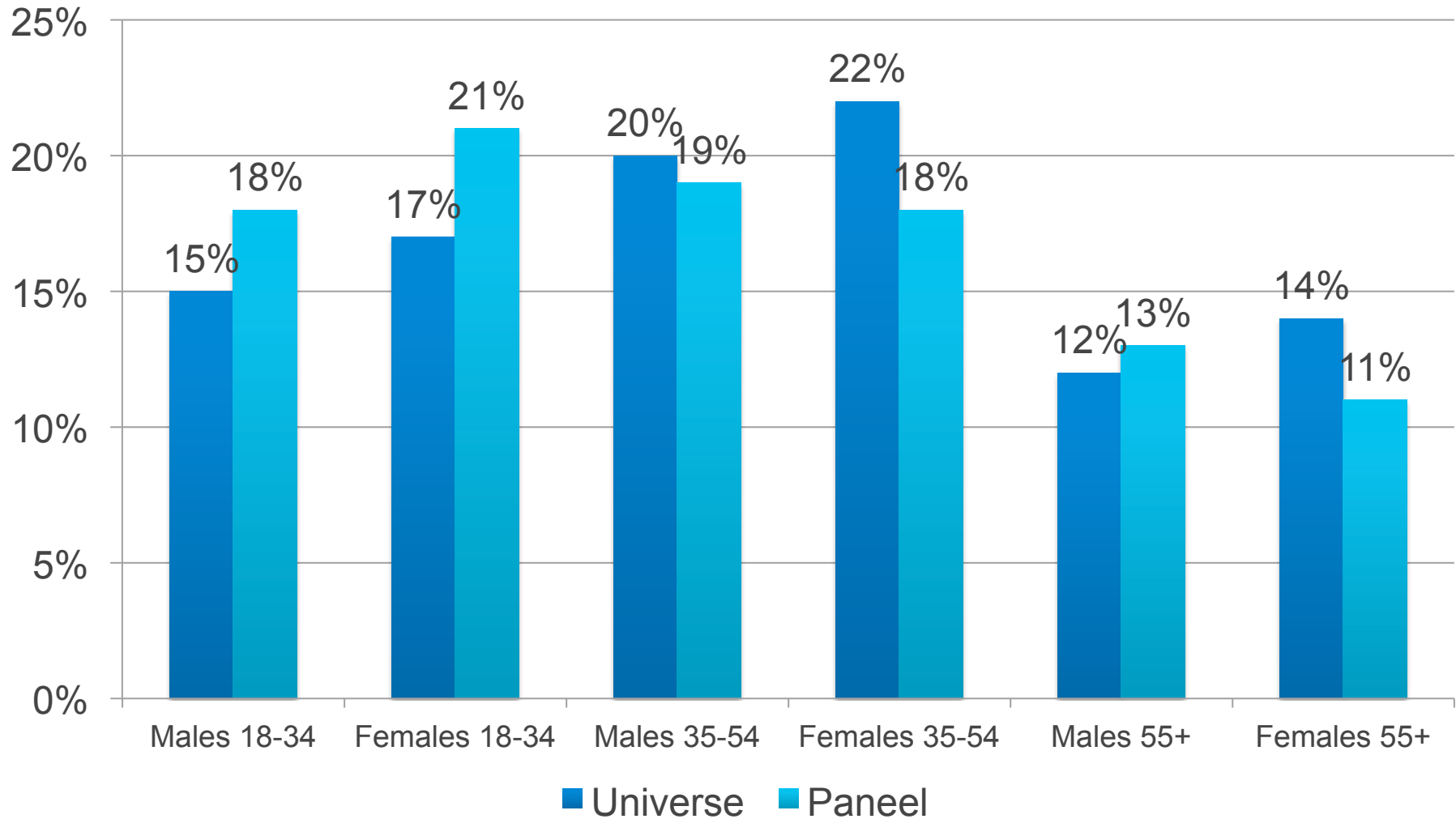
Audience measurement neither supports nor leads to cookie targeting

In fact, cookie targeting circumvents audience measurement entirely

Integrating Panel Data and Census Data

Strategy 1: Calibration

How Does Weighting Work? Illustration: Sample Composition versus Universe

How Does Weighting Work?

Each person in the panel represents a number of people in the universe. If the Universe is 10 million, and the panel is 10,000 people, then (without weighting) each person represents (10 million / 10,000), or 1,000 people

- Note: the size of the universe is typically ascertained through surveys

But the market requires that the panel is "weighted" to reflect the universe for age and gender (so as not to penalize or benefit unfairly media vehicles with specific age or gender skews)

We saw that 15% of the panel was male 18-34, versus 18% of the Universe

Basically, weighting for demographics means we need to "weight up" the sample Males 18-34 by (18/15, or 1.2)

So we apply a weight of 1.2 to each male 18-34 panelist, and now instead of having them each represent 1,000 people, they each are "weighted" to represent 1200 people

- After weighting, male 18-34 panelists will collectively carry 18% of the weight, matching universe

Calibration of a panel to cookie and beacon data works in a similar fashion.

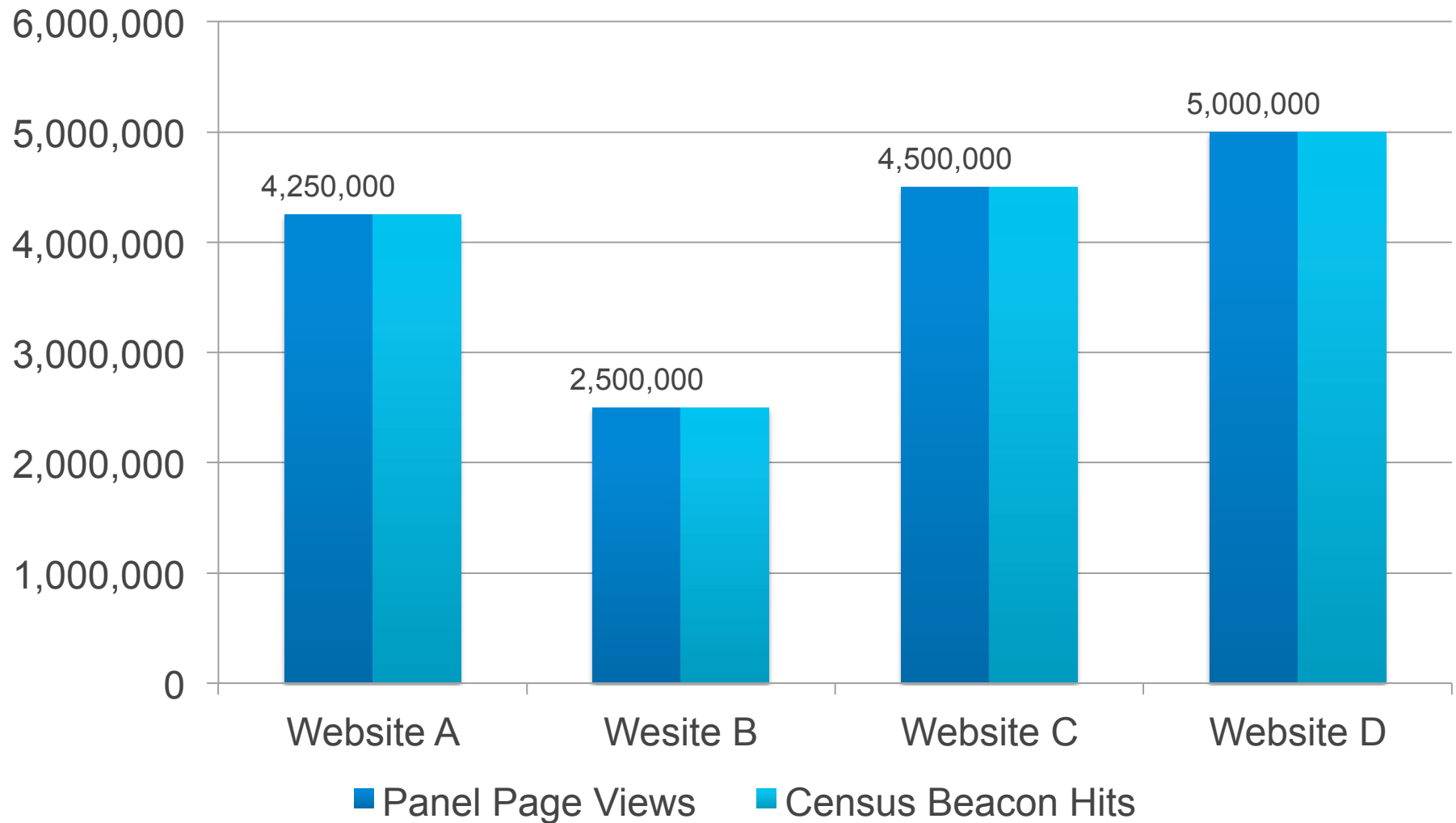comScore.

Panel-Only Page Views Versus Census Beacon Hits

We know empirically how many census beacon hits (analogous to Page Views) were generated at sites A, B, C, and D

We can thus weight the panelists such that after weighting, the panel projections for Page Views match the known counts for census beacon hits

Integrating Panel Data and Census Data

Strategy 2: Hybrid Integration

comSCORE.

The Industry Has Demanded That Audience Measurement Companies Integrate Site-Centric Data Into the Audience Measurement Process

"Imagine my surprise when I came to the IAB and discovered that the main audience measurement companies are still relying on panels, a media-measurement technique invented for the radio industry exactly seven decades ago, to quantify the Internet."

--Randall Rothenberg; CEO, US IAB; Open letter to comScore and Nielsen, April 20, 2007

But site-centric audience measurement is cookie-based, and audiences are comprised of persons

Integration of Cookie Data and Panel Data to Size Website Audiences

Given a publisher tags content for the measurement company (let's call that company, "comScore")

Step 1: Count comScore census cookies observed at that site in a month

- Allocate by country (that is, in creating the US audience, limit analysis to US cookies only)
- Eliminate bots, spiders etc. (known IAB lists, plus active filtration)
- Net result: in-country filtered cookies (let's call it, 3 million cookies)

Step 2: Observe cookies per person from the panel

- Among panelists, research company can see cookie behavior
- Develop a panel-based cookies-per-person metric (let's call this 1.5)

Step 3: Reported audience is (census filtered cookies) / (Cookies per person)

- In this case, 3 million / 1.5 = 2 million persons

So Where Do Demographics Come From?

- **Demographics come entirely from the panel**

- **Suppose the panel (post-weighting and calibration, if applicable) tells us that 20% of the site's audience is female 25-54**

- **Apply the % composition from the panel (20%) to the site's total projected audience (2 million, previous slide), or 400,000 women 25-54**

comSCORE.

A Few Points

1. We don't use, or even attempt to know, anything about the persons behind these cookies

2. All we are doing is counting the cookies

3. We do not believe that counting is tracking

Questions & Conversation