



# INNOVATIONS IN HYBRID SAMPLING TECHNIQUES

## Improving Representation of Teens and Young Adults

---

**Mansour Fahimi, Ph.D.**  
**Marketing Systems Group**

# PRESENTATION LAYOUT

- **Big Picture**
- **Traditional Inferential Paradigm**
- **What Has Changed**
- **Survey Sampling in the Digital Age**
- **Mathematics of Sample Mixology**
- **Takeaways**

# BIG PICTURE

## Total Survey Error

### Errors of Non-observation

- Sampling Frame
- Sampling Method
- Nonresponse

### Errors of Observations

- Questionnaire
- Data Collection
- Respondent

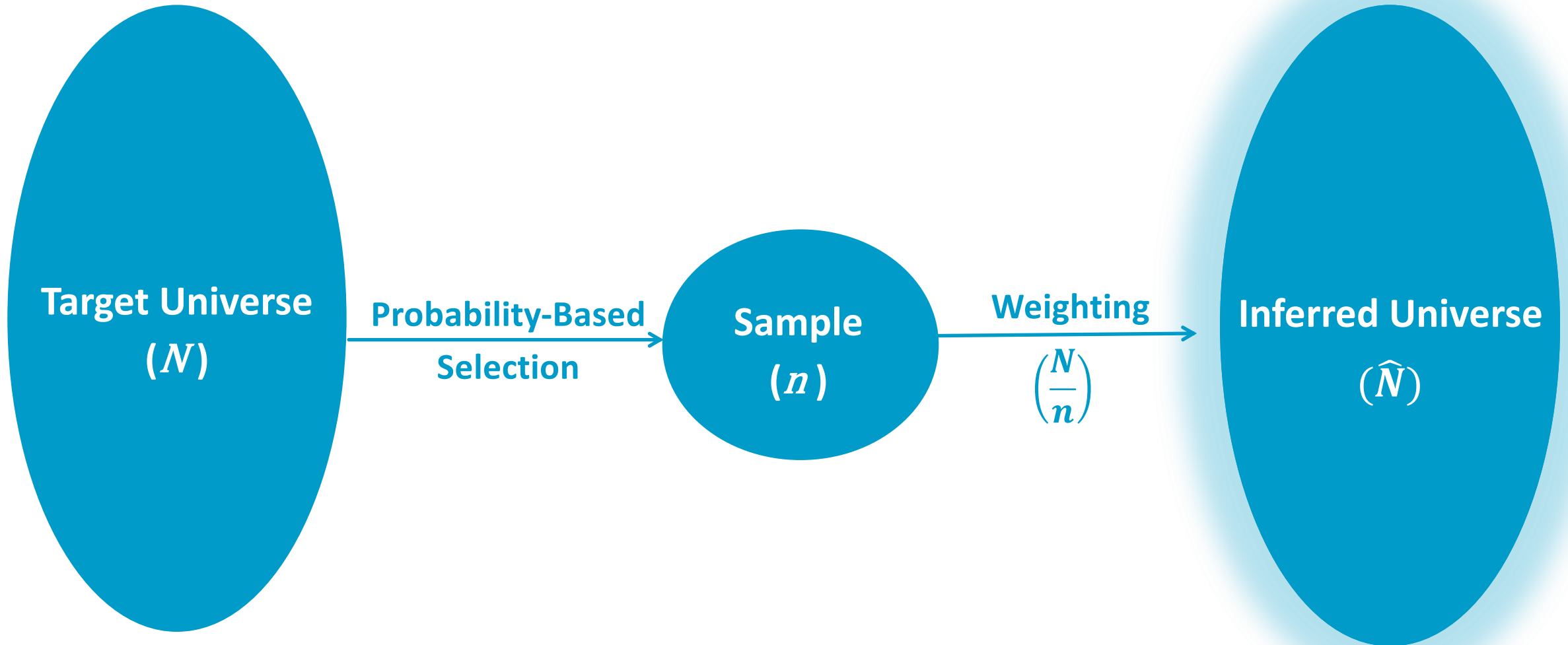
### Errors of Processing

- Data Editing
- Imputation
- Weighting

### Errors of Dissemination

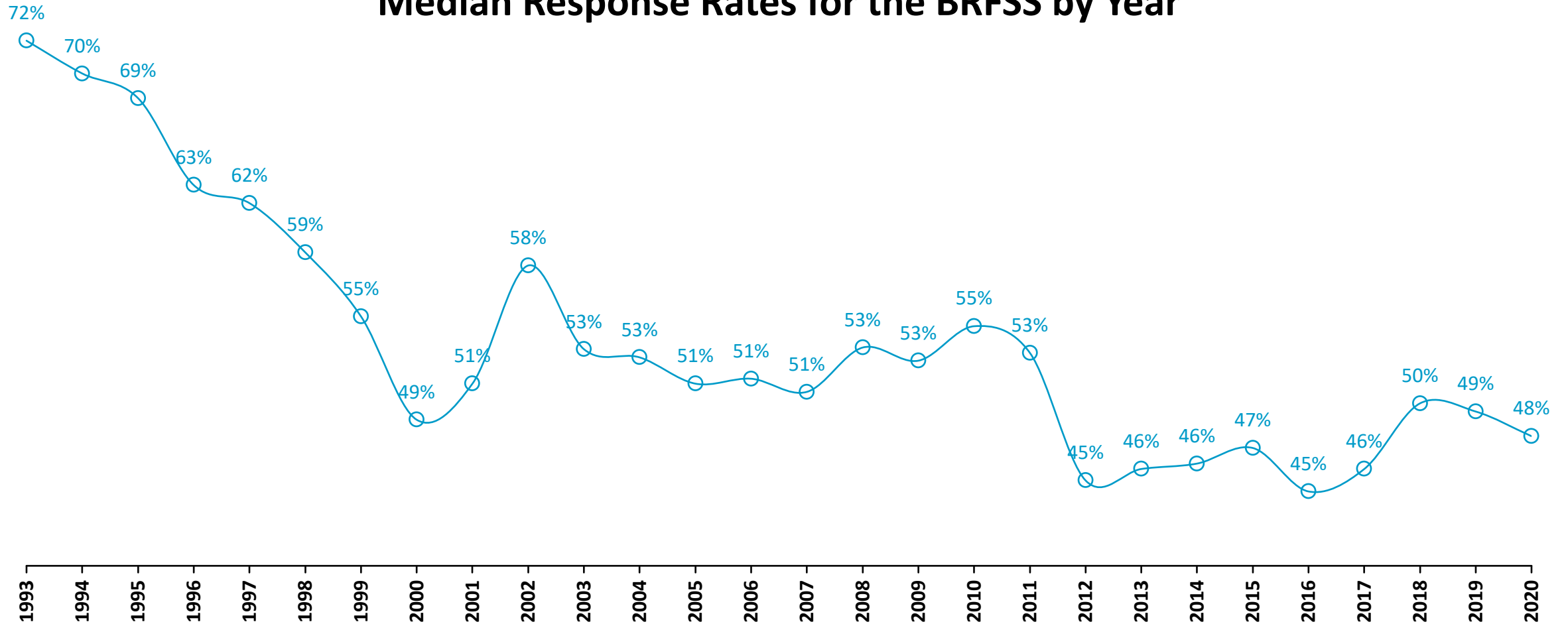
- Field to Researchers
- Researchers to Users
- Analysis & Reporting

# THE MAGIC OF STATISTICAL INFERENCE



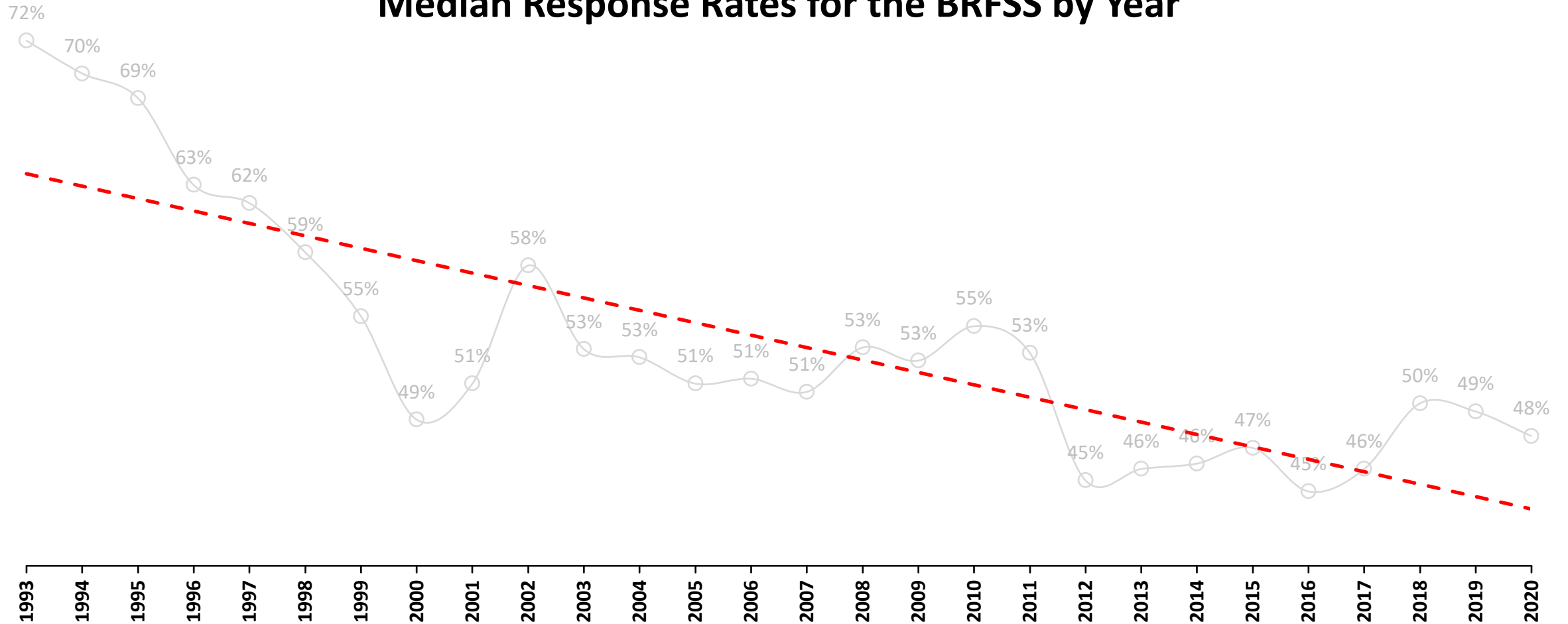
# WHAT HAS CHANGED

## Median Response Rates for the BRFSS by Year



# WHAT HAS CHANGED

## Median Response Rates for the BRFSS by Year



# SURVEY SAMPLING IN THE DIGITAL AGE

## ➤ **Growing interest in alternative survey sampling methods:**

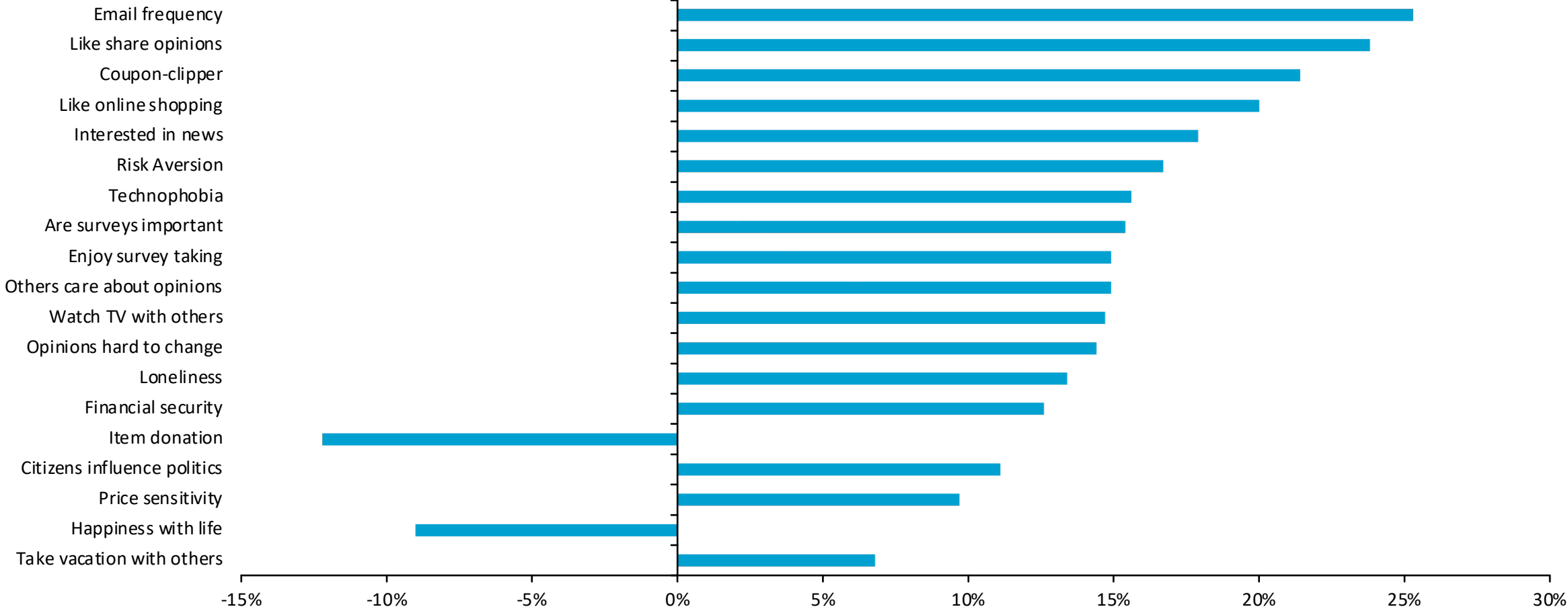
- Reduce cost
- Speed up administration
- Enhance survey experience
- Improve coverage
- Address eroding rates of response

## ➤ **Emerging alternatives sampling methods:**

- Online panels with “*organic*” representation
- Online panels with “*synthetic*” representation
- Hybrid sampling supplementing ABS with samples from online panels

# SAMPLES WITH COMPROMISED REPRESENTATION

## Inadequacy of Traditional Geodemographic Weighting





# CALIBRATION 1.0 (UNIVARIATE)

## ➤ Correction for Higher Propensity of Early Adoption (DiSogra & Cobb 2011)

- **E1:** I usually try new products before others
- **E2:** I often try new brands because I like variety
- **E3:** I look for what is new when I shop
- **E4:** I like to be the first to try something new
- **E5:** I like to tell others about new brands

## ➤ Correlation Coefficients Among Early Adoption Measurers

Measure	E1	E2	E3	E4	E5
E1	1.0	0.7	0.7	0.9	0.7
E2		1.0	0.7	0.7	0.8
E3			1.0	0.7	0.7
E4				1.0	0.7
E5					1.0

# NEED FOR MULTIVARIATE CALIBRATIONS

## ➤ **Social Engagement:**

- Take vacation with others
- Exercise/play sports with others
- Have a meal in someone's home

## ➤ **Self-Importance:**

- Importance of sharing opinions
- My opinions are hard to change
- Feel confident in social settings

## ➤ **Shopping Habits:**

- Use of coupons when shopping
- Enjoying shopping online
- Importance of brand compared to price

## ➤ **Happiness and Security:**

- Happiness with life
- Feeling insecure and lonely
- Concerned about cyber theft

## ➤ **Politics:**

- Having influence on politics
- Views on efficacy of government
- Closely following the news

## ➤ **Community:**

- Feeling part of a community
- Number of moves in past 5 years
- Religiosity

## ➤ **Politics:**

- Having influence on politics
- Views on efficacy of government
- Closely following the news

## ➤ **Altruism:**

- Donating blood
- Donating items
- Volunteering without pay

## ➤ **Survey Participations:**

- Experience with online surveys
- Important of taking opinion surveys
- Frequency of online surveys in a month

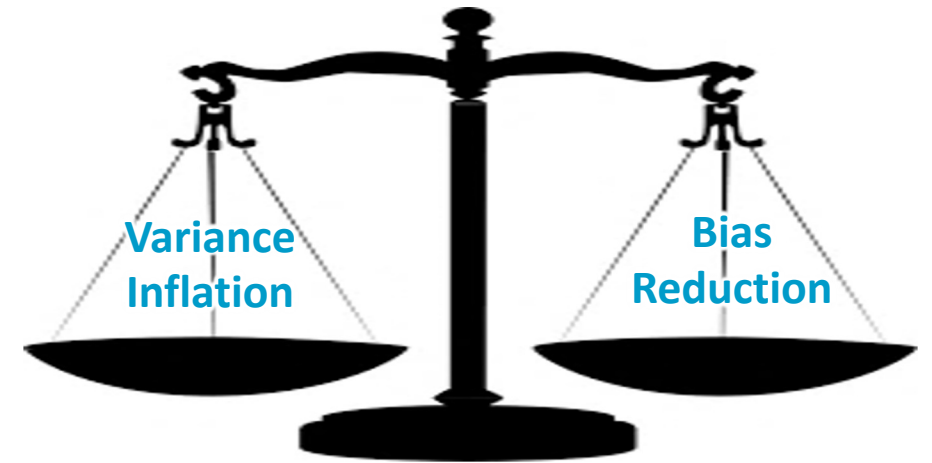
## ➤ **Internet and Social Media Usage:**

- Frequency of accessing personal email
- Frequency of accessing Internet
- Time spent watching TV per day

# CALIBRATION 2.0 (MULTIVARIATE)

## ➤ Pragmatic Parsimony:

- Stepwise Logistic Regression
- Classification and Regression Trees
- Random Forests



## ➤ Beyond Early Adoption Differences (Fahimi, Barlas, and Thomas 2015)

- Amount of time spent online
- Interest in expressing opinions online
- Coupon usage
- Amount of TV watching
- Frequency of relocations

## CALIBRATION 2.1 (YOUNGER RESPONDENTS)

### ➤ **Effective Adjustments for Teens & Young Adults (Ipsos & Truth 2021)**

- ~~■ Amount of time spent online~~
- ~~■ Interest in expressing opinions online~~
- ~~■ Coupon usage~~
- ~~■ Amount of TV watching~~
- Housing type
- Financial comfort
- Living with parents
- Interest in influencing social media
- Frequency of relocations

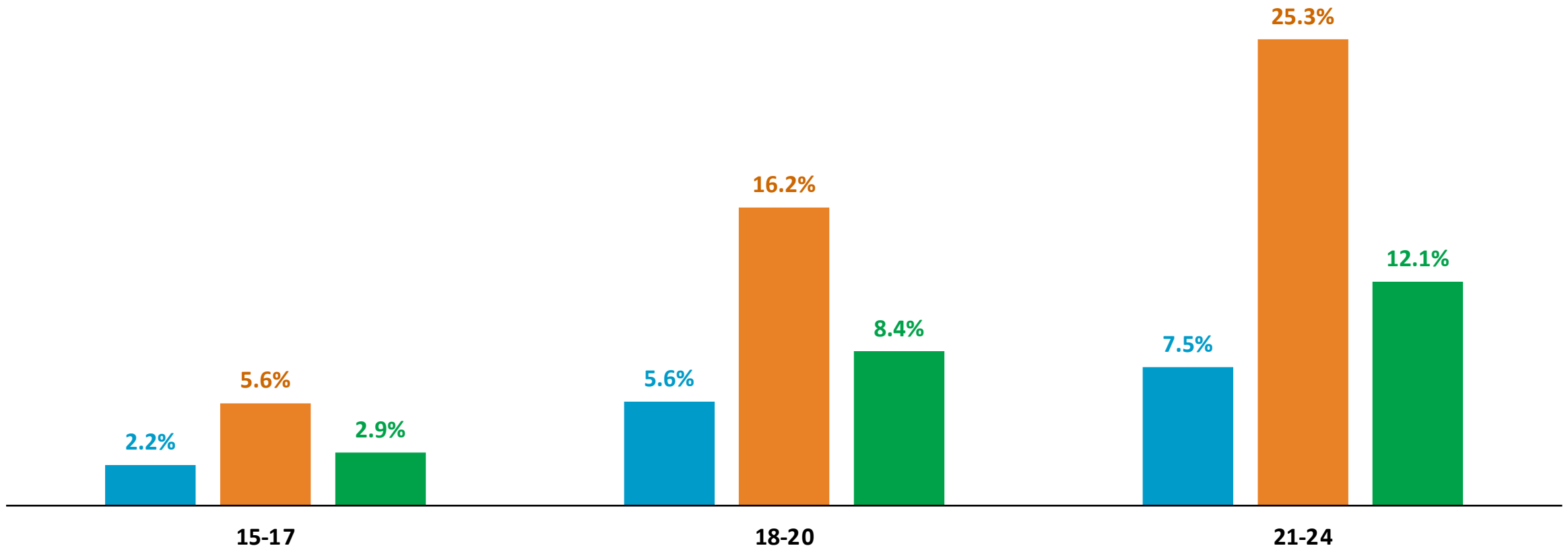
# MATHEMATICS OF SAMPLE MIXOLOGY

- Optimal blending of two independent sample means  $\bar{y}_1$  and  $\bar{y}_2$
- Composite Estimator:  $\bar{y} = \alpha\bar{y}_1 + (1 - \alpha)\bar{y}_2$
- For example, if  $\alpha = 0.5$  then:  $\bar{y} = \frac{\bar{y}_1 + \bar{y}_2}{2}$
- Variance of  $\bar{y}$ :  $V(\bar{y}) = \alpha^2V(\bar{y}_1) + (1 - \alpha)^2V(\bar{y}_2) + 2\alpha(1 - \alpha)Cov(\bar{y}_1, \bar{y}_2)$
- Since  $\bar{y}_1$  and  $\bar{y}_2$  are independent:  $Cov(\bar{y}_1, \bar{y}_2) = 0$
- Therefore:  $V(\bar{y}) = \alpha^2V(\bar{y}_1) + (1 - \alpha)^2V(\bar{y}_2) = \alpha^2V(\bar{y}_1) + V(\bar{y}_2) + \alpha^2V(\bar{y}_2) - 2\alpha V(\bar{y}_2)$
- Minimizing  $V(\bar{y})$  with respect to  $\alpha$ :
  - $\frac{\partial[V(\bar{y})]}{\partial\alpha} = 0 \Rightarrow 2\alpha V(\bar{y}_1) + 2\alpha V(\bar{y}_2) - 2V(\bar{y}_2) = 0 \Rightarrow \alpha_{optimal} = \frac{V(\bar{y}_2)}{V(\bar{y}_1) + V(\bar{y}_2)}$

# MATHEMATICS OF SAMPLE MIXOLOGY

## Cigarette Smoking Prevalence by Age and Sample Type

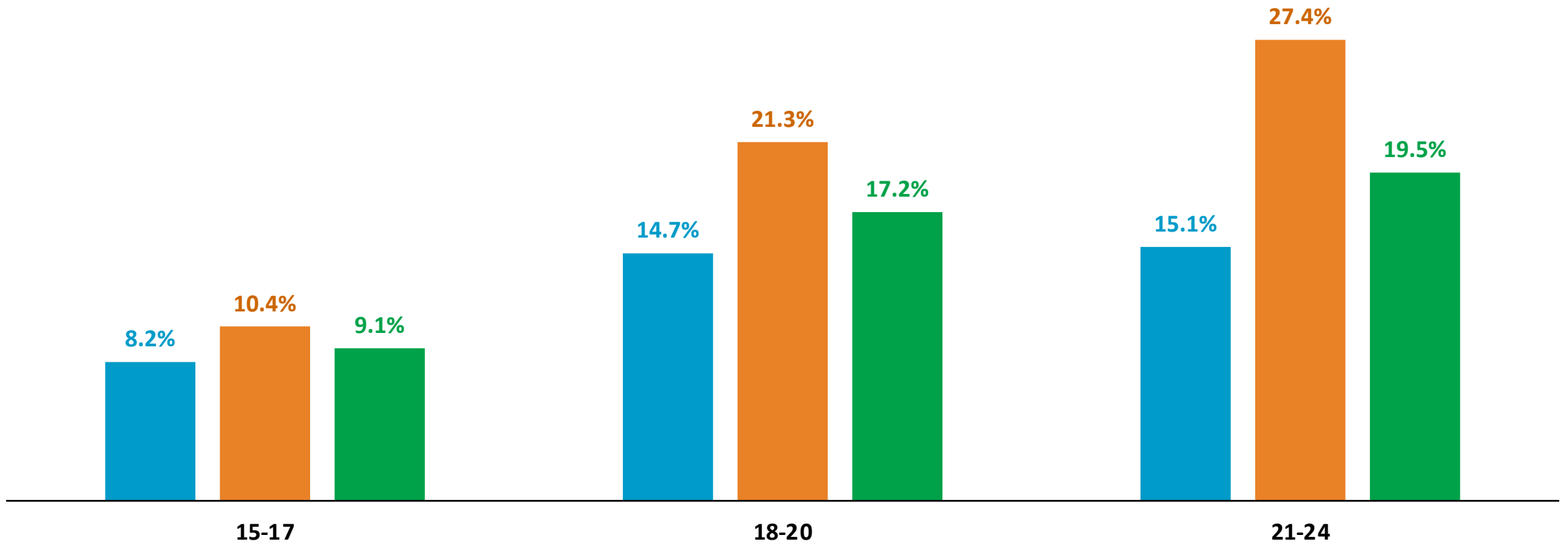
■ ABS ■ Panel ■ ABS + Panel



# MATHEMATICS OF SAMPLE MIXOLOGY

## e-Cigarette Smoking Prevalence by Age and Sample Type

■ ABS ■ Panel ■ ABS + Panel



# TAKEAWAYS

- **New Realities of the Accelerating *Digital Age*:**
  - Traditional sampling methods have issues: coverage, time, and cost
  - Perfect coverage and high response rates are fading luxuries
  - The days of lengthy surveys are over
  - Alternative (hybrid) sampling methods are becoming more effective
- **Pragmatic Rigor:**
  - Respecting the strengths of traditional methods
  - Appreciating the potential gains of alternative methods
  - Dealing transparently with the shortfalls of alternative methods
  - Embracing parsimonious instrumentation
  - Recognizing the need for pragmatic sampling & effective weighting adjustments
  - Relying on analytics that are proper & responsive to the above





THANK YOU



**Mansour Fahimi, Ph.D.**  
**EVP & Chief Data Scientist**  
[mfahimi@m-s-g.com](mailto:mfahimi@m-s-g.com)