

Approaches to Increase Survey Participation and Data Quality in an At-risk, Youth Population

**Presentation at the FedCASIC Conference
Washington, D.C.**

March 19, 2014

Lisbeth Goble • Jillian Stein • Lisa K. Schwartz

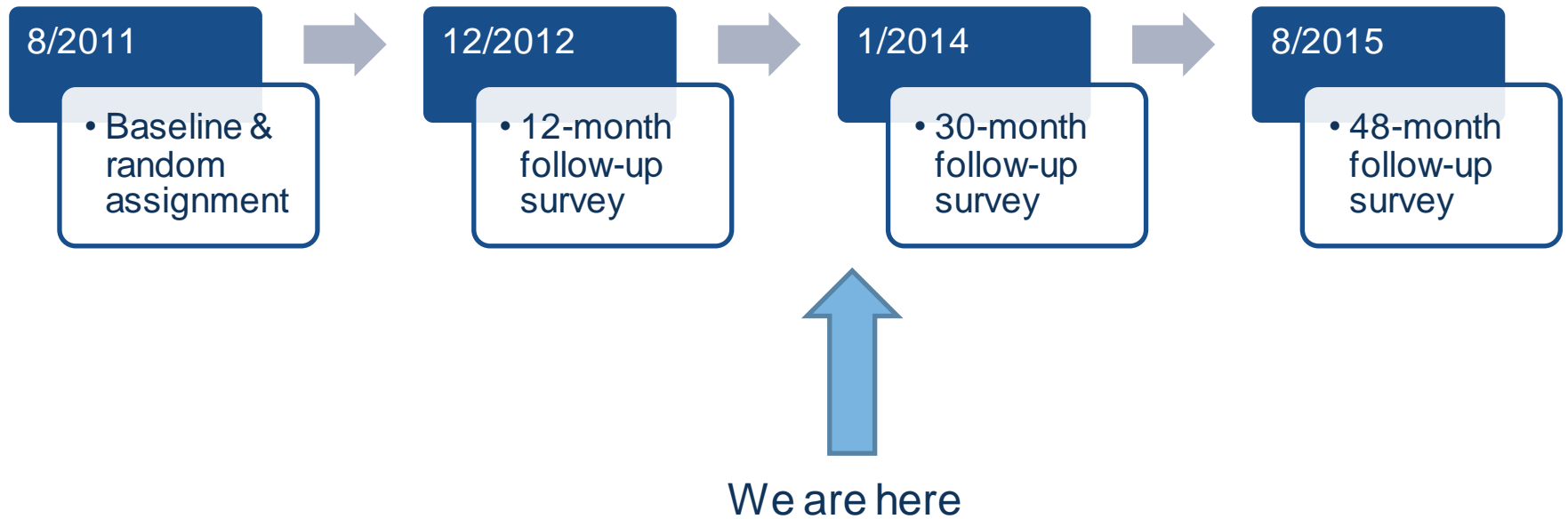
Overview of Presentation

- **Background**
- **Research Questions**
- **Data and Methods**
- **Findings**
- **Conclusions and Next Steps**

Evaluation of the YouthBuild Program

- **Seven-year random assignment evaluation**
 - Funded by DOL, ETA
 - MDRC is prime contractor with Mathematica and SPR
 - 77 sites across the country, ~ 3,400 youth enrolled over 18 months
- **“Second-chance” program serving primarily high school dropouts**
- **Primary outcomes of interest are educational attainment, employment, and involvement with criminal justice system**
- **Secondary outcomes include social and emotional development, community and civic engagement**

YouthBuild: Data Collection Timeline



YouthBuild: Data Collection Strategy

- **Multi-mode data collection for rolling sample of cohorts**
 - Web-only phase (~ 5 weeks)
 - Web + CATI phase (~ 3 months)
 - Web + CATI + Field Locating phase (~ 3 months)

YouthBuild: Sample Characteristics at Baseline

Characteristics	
Evaluation Treatment Group	64%
Male	64%
Age (avg)	20.2
White, non-Hispanic	15%
Black, non-Hispanic	62%
Other, non-Hispanic	8%
Hispanic	14%
Has child/children	31%
Housing: In stable housing	76%
Housing: In temporary housing	18%
Housing: Other/unknown	6%

YouthBuild: Data Collection Challenges

- Achieve an 80 percent response rate with hard-to-reach population

YouthBuild: Data Collection Challenges

Stage of Data Collection	Key Challenges
Locating	<ul style="list-style-type: none">• Highly mobile• Limited “electronic footprint”
Contacting	<ul style="list-style-type: none">• Cell-only “households”• Housing instability• Nontraditional modes of communication
Gaining cooperation	<ul style="list-style-type: none">• Appeals to civic duty, benefits of research unlikely to be compelling

Adaptive Design

- **Responsive design (Groves & Heeringa, 2006)**
 - Analysis and mid-course corrections during data collection based on real-time data about survey processes (paradata) and survey response
 - Implemented between phases of data collection
 - Sample design, protocols, modes remain fixed during a phase
- **Adaptive design (Wagner, 2008)**
 - Modify survey processes in real-time
 - Tailor protocols to subgroups to maximize response

Key Research Questions

- 1.** To what extent are baseline characteristics predictors of response at follow-up (age, race, sex, contact information)?
- 2.** Of those who complete, what are the characteristics of those who complete on the web? On the phone?
- 3.** Do targeted approaches such as differential incentives affect response rates?
- 4.** How might we use this information to inform future rounds of data collection?

Data Sources

Data Source	Information
Baseline Information Form (BIF)	<ul style="list-style-type: none">• Respondent demographics• Contact information• Contact preferences
Paradata	<ul style="list-style-type: none">• Data on data collection efforts including:<ul style="list-style-type: none">• Locating• Notifications• Random Assignment status• Incentive Experiment status
12-month follow-up survey	<ul style="list-style-type: none">• Complete status• Status by mode

Sample

- **Complete Sample – N=3,436**
- **Analytical Sample – n=2,562**
 - Data collection on-going; using an interim data file
 - Restricted to cases released at least four-weeks prior to when data was pulled
 - No differences in demographics of complete and analytical sample

Developing an Adaptive Design Plan

- **Planning Baseline Data Collection to Inform Adaptive Design**
 - Multiple modes of contact information
 - Preferred mode of contact

18. How would you prefer to be contacted in the future?

1 Regular Mail

5 Facebook

2 Email

6 Twitter

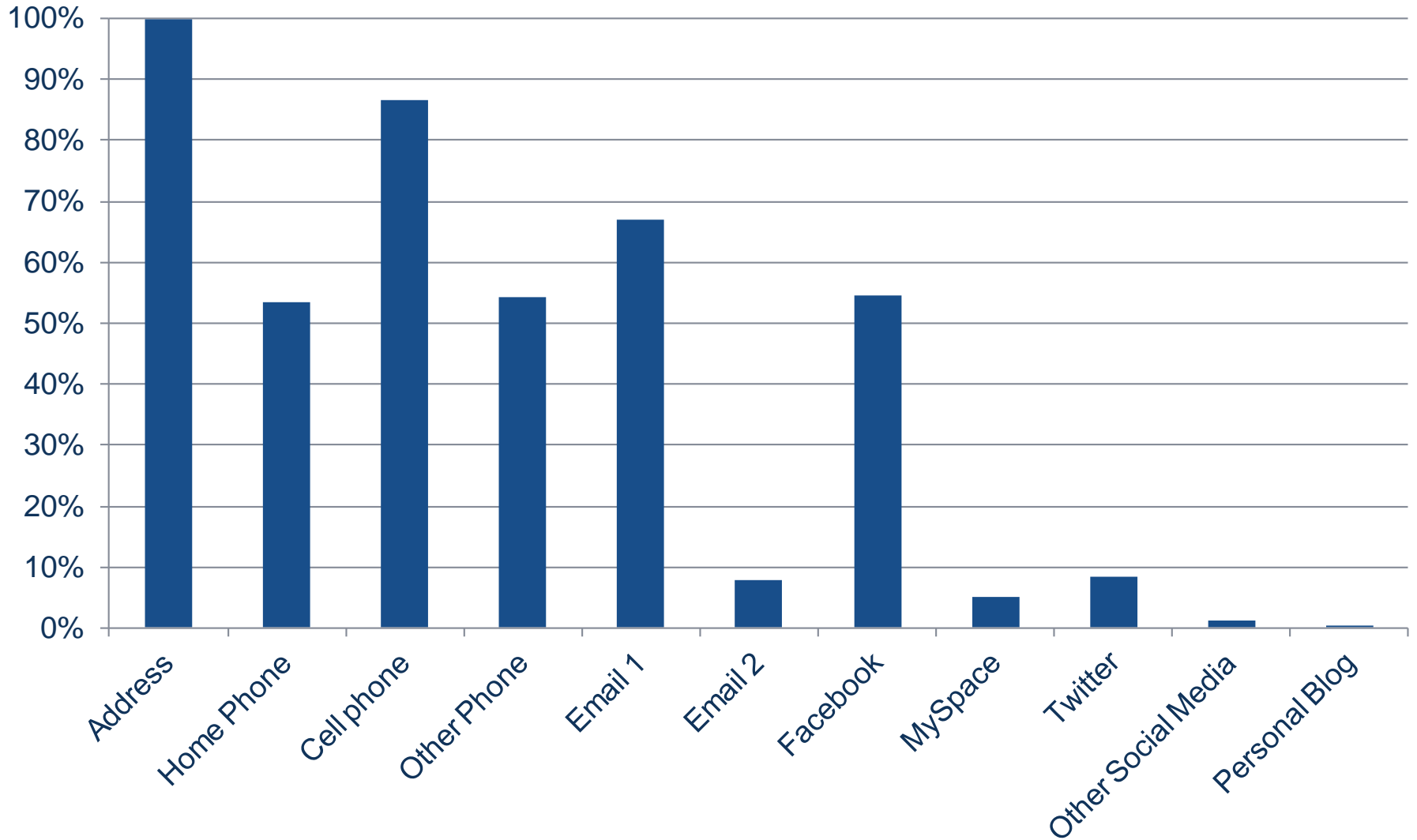
3 Cell Phone

7 MySpace

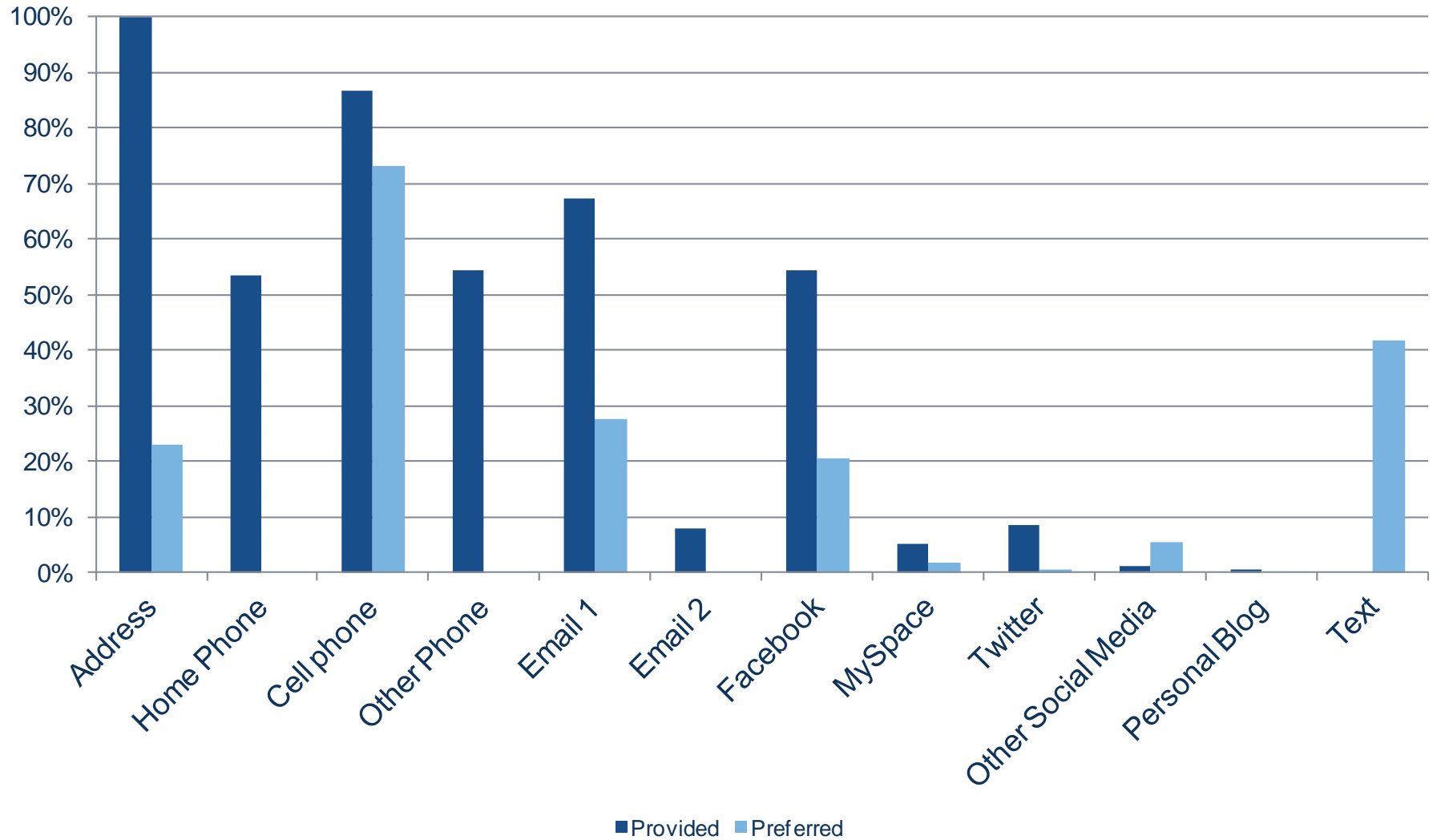
4 Text

8 Other _____

Contact Information from BIF



Provided contact vs preferred contact



Findings

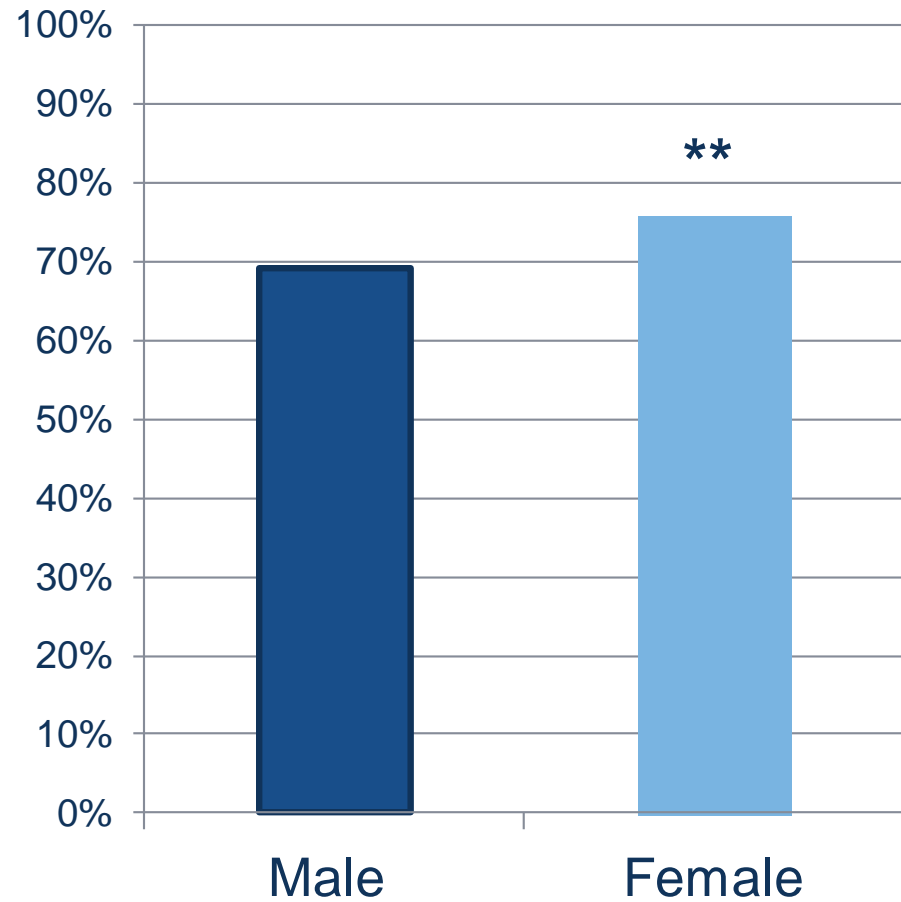
- **To what extent are baseline characteristics predictors of response at follow-up?**
 - **Sex**
 - **Age**
 - **Race/Ethnicity**
 - **Housing**
 - **RA Status**
 - **Contact information**

Methods

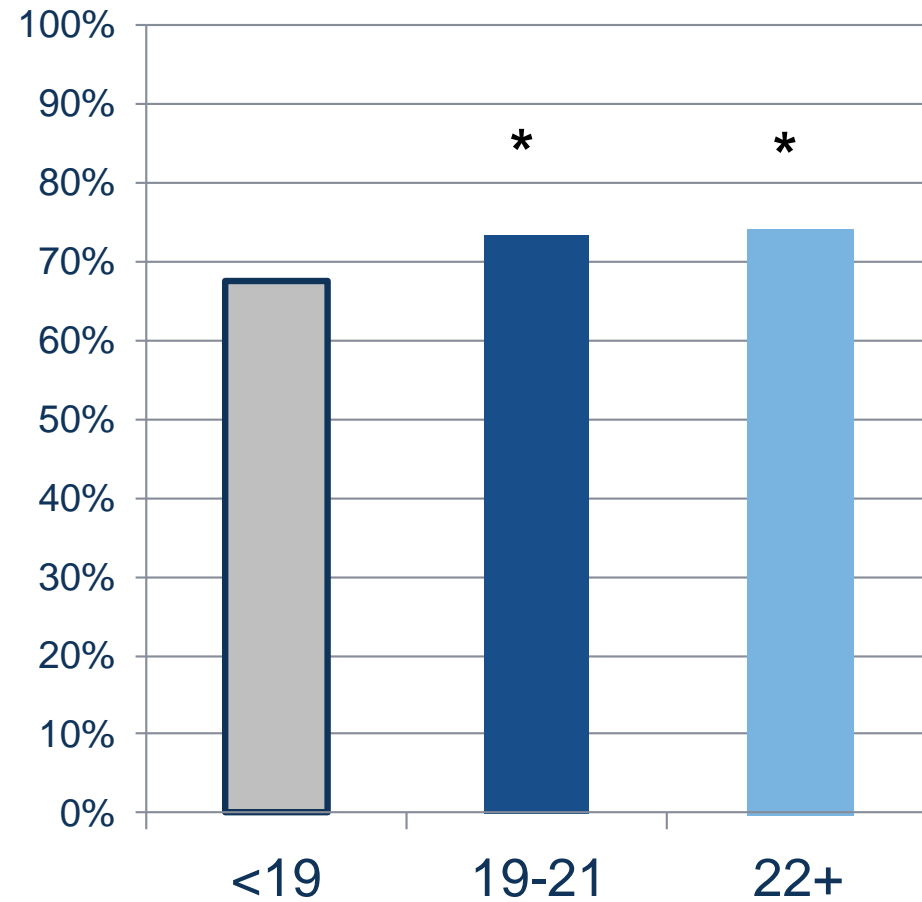
- **Logistic Regression models with completion as the dependent variable**
 - Ever complete
 - By phone
 - By Web
- **Controlling for demographics and contact information**
- **Reporting predicted probabilities of completion holding controls at their mean**

Females and 'older' respondents more likely to complete

Sex

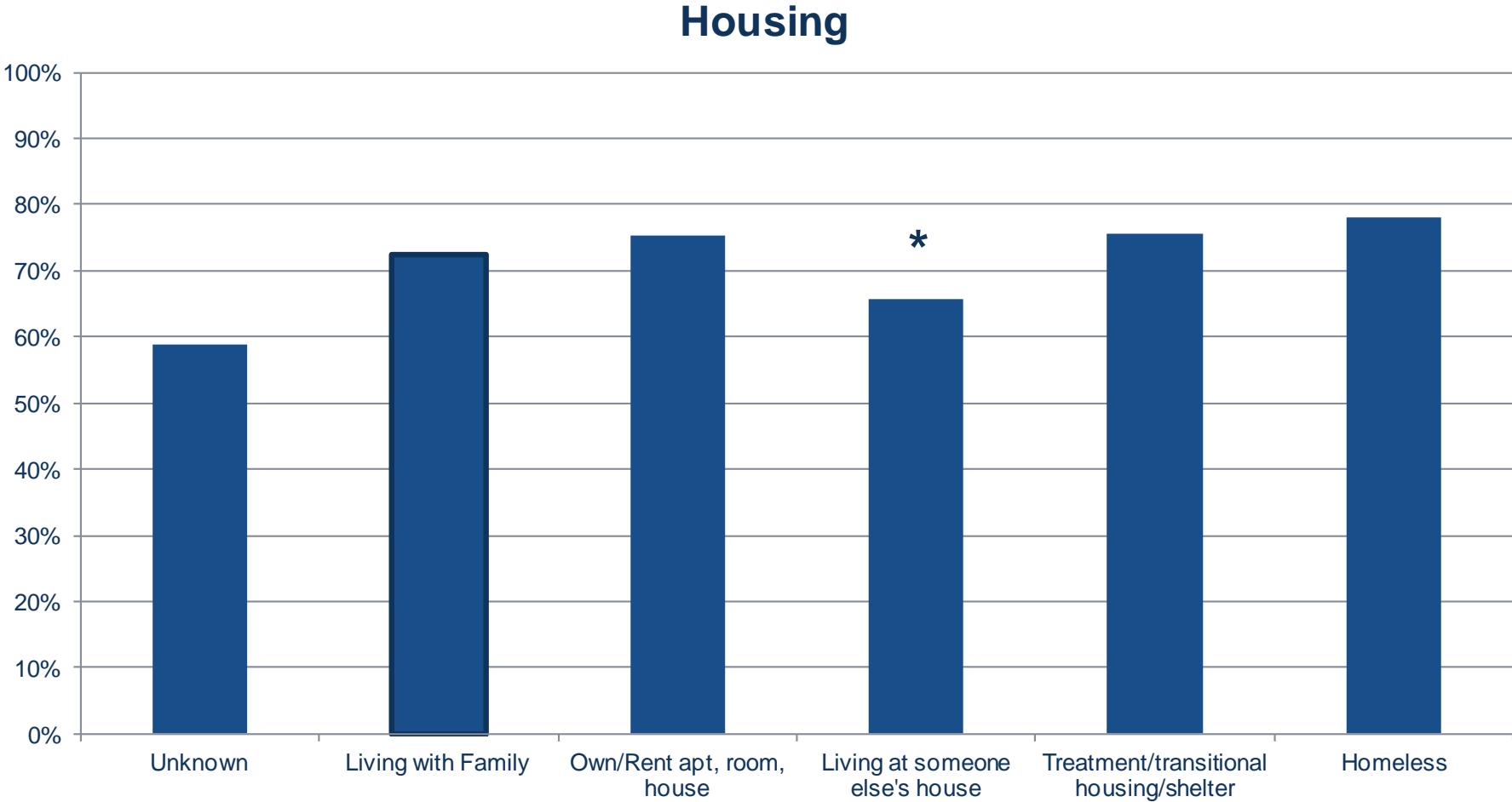


Age



* p<.05; **p<.01

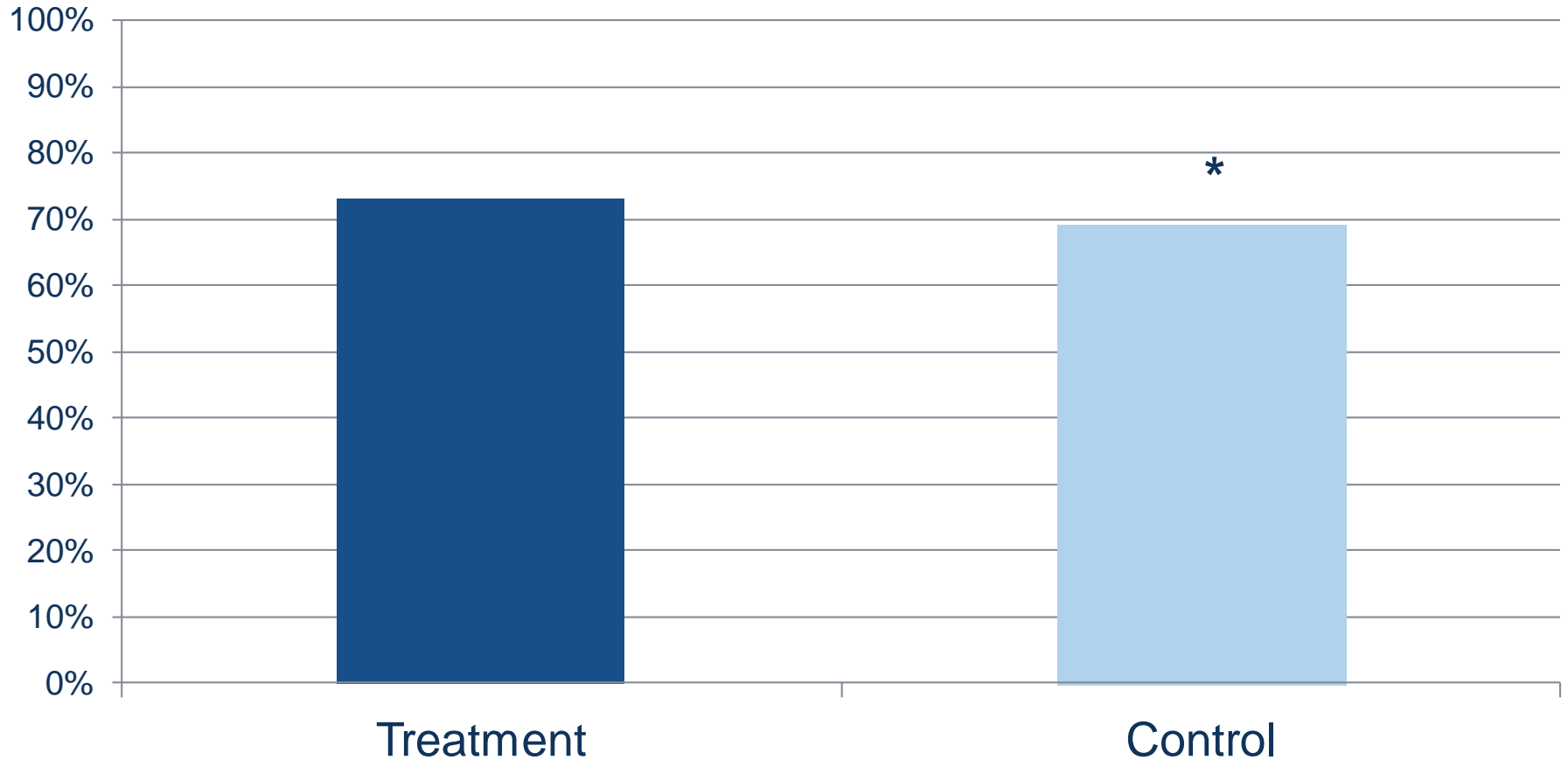
Those living with someone else less likely to complete



* p<.05; **p<.01

Treatment group more likely to complete

YouthBuild Treatment Status

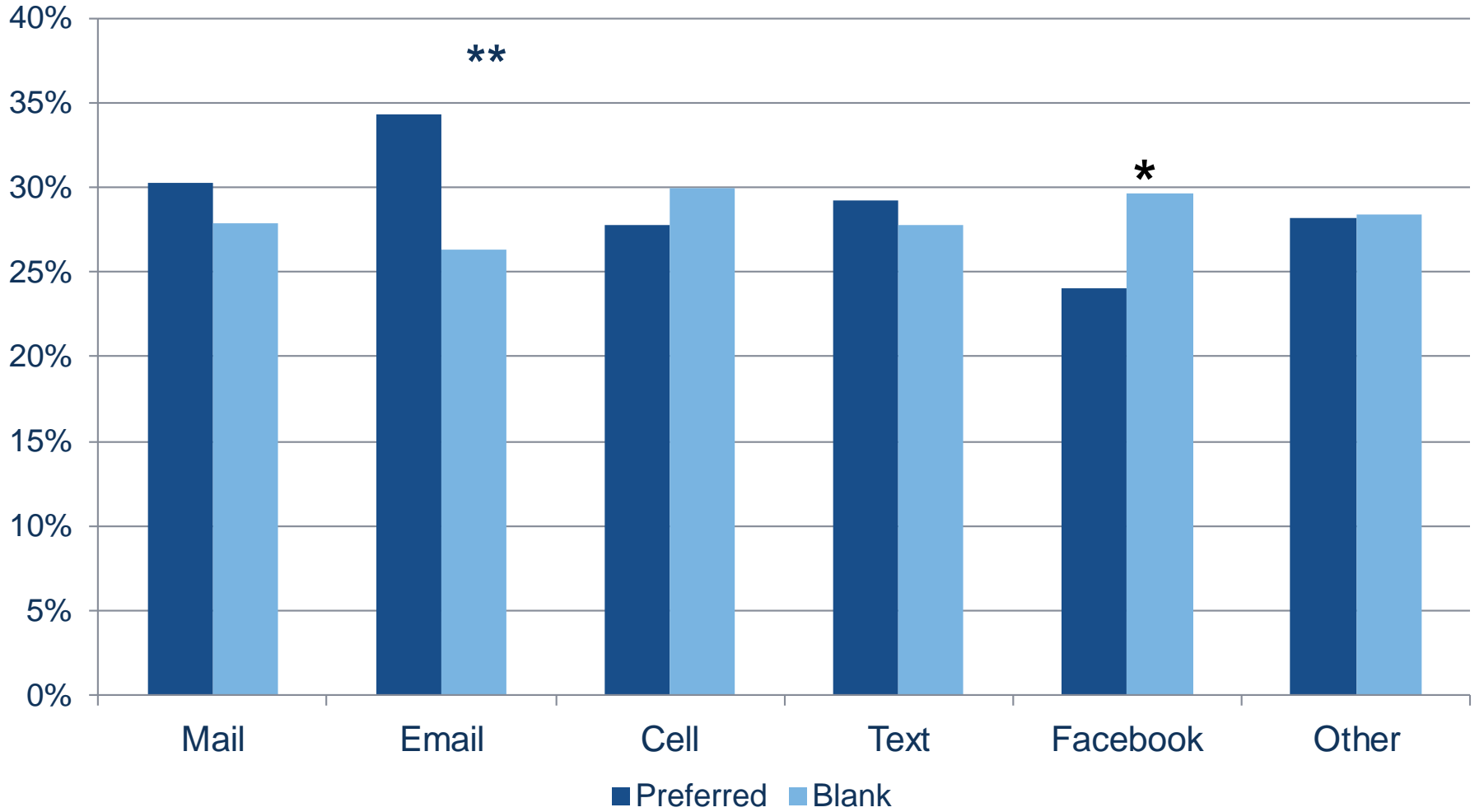


* p<.05; **p<.01

Characteristics of web respondents

- **Black, non-Hispanic** *less likely*
- **Housing (compared to living with family)**
 - Living with someone else *less likely*
 - Living in transitional/treatment housing *more likely*

Web Completion Rates by Contact Preferences

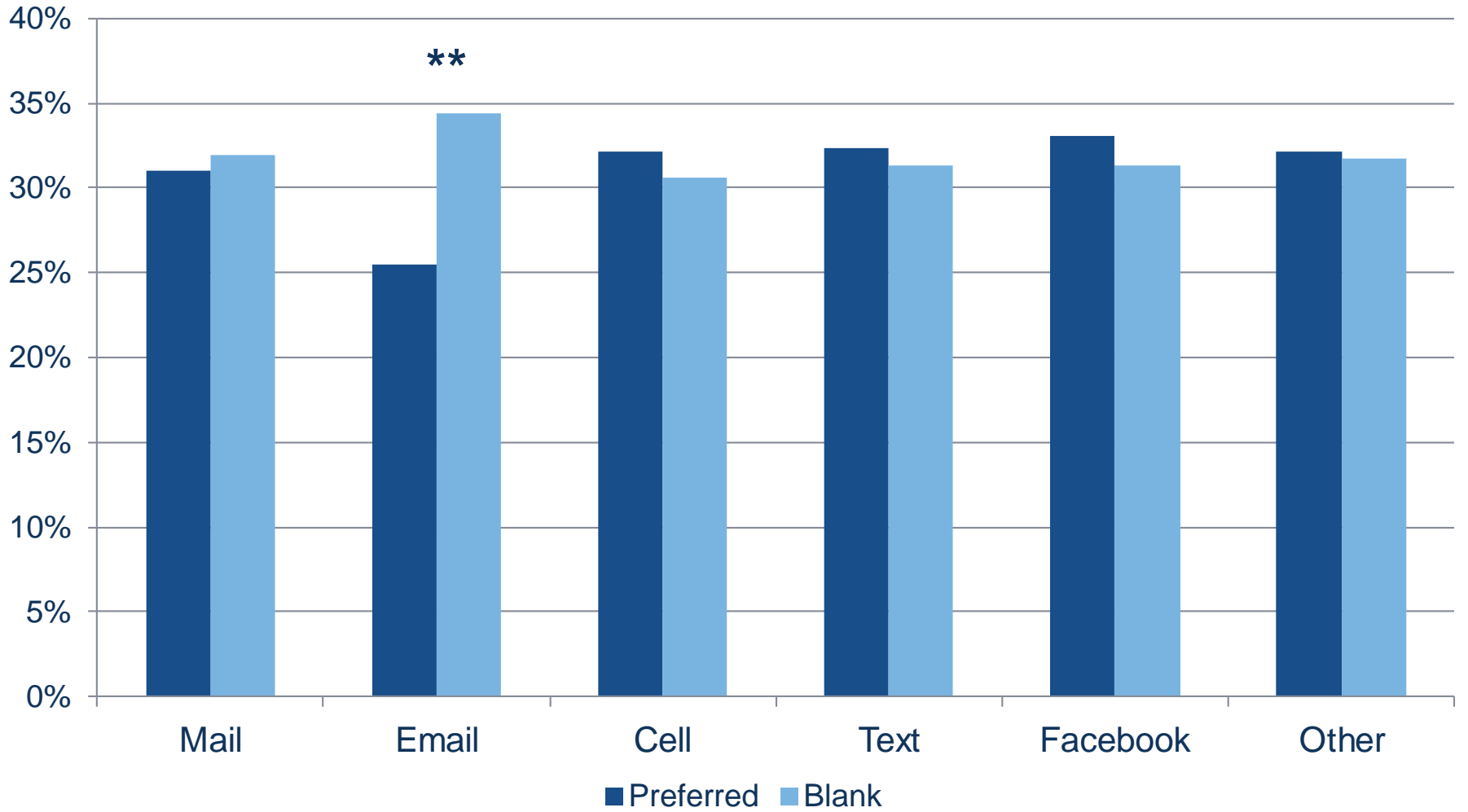


* $p < .05$; ** $p < .01$

Characteristics of phone respondents

- **Housing (compared to living with family)**
 - Own/rent housing more likely
 - Homeless more likely

Phone Completion Rates by Contact Preference

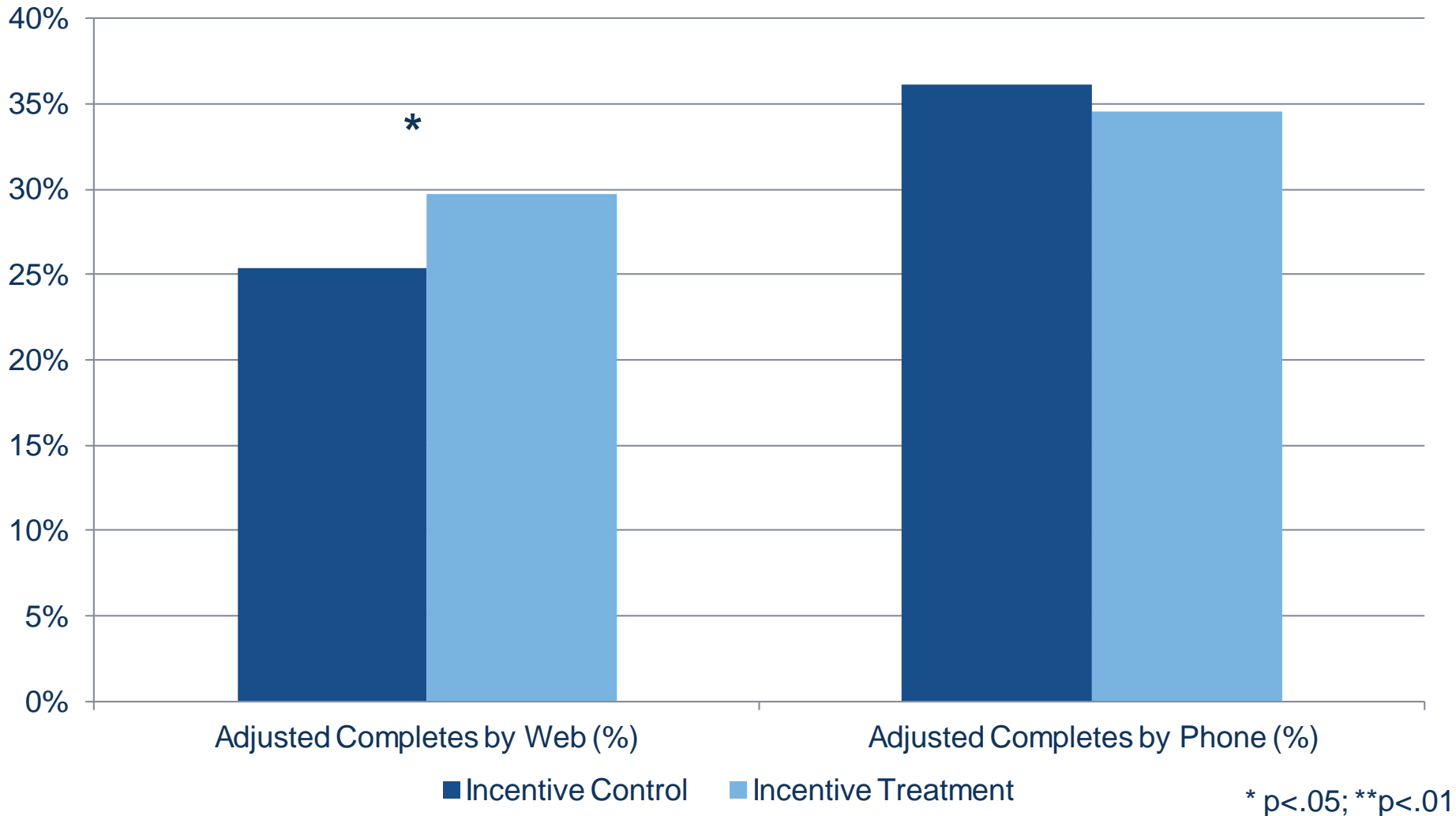


* p<.05; **p<.01

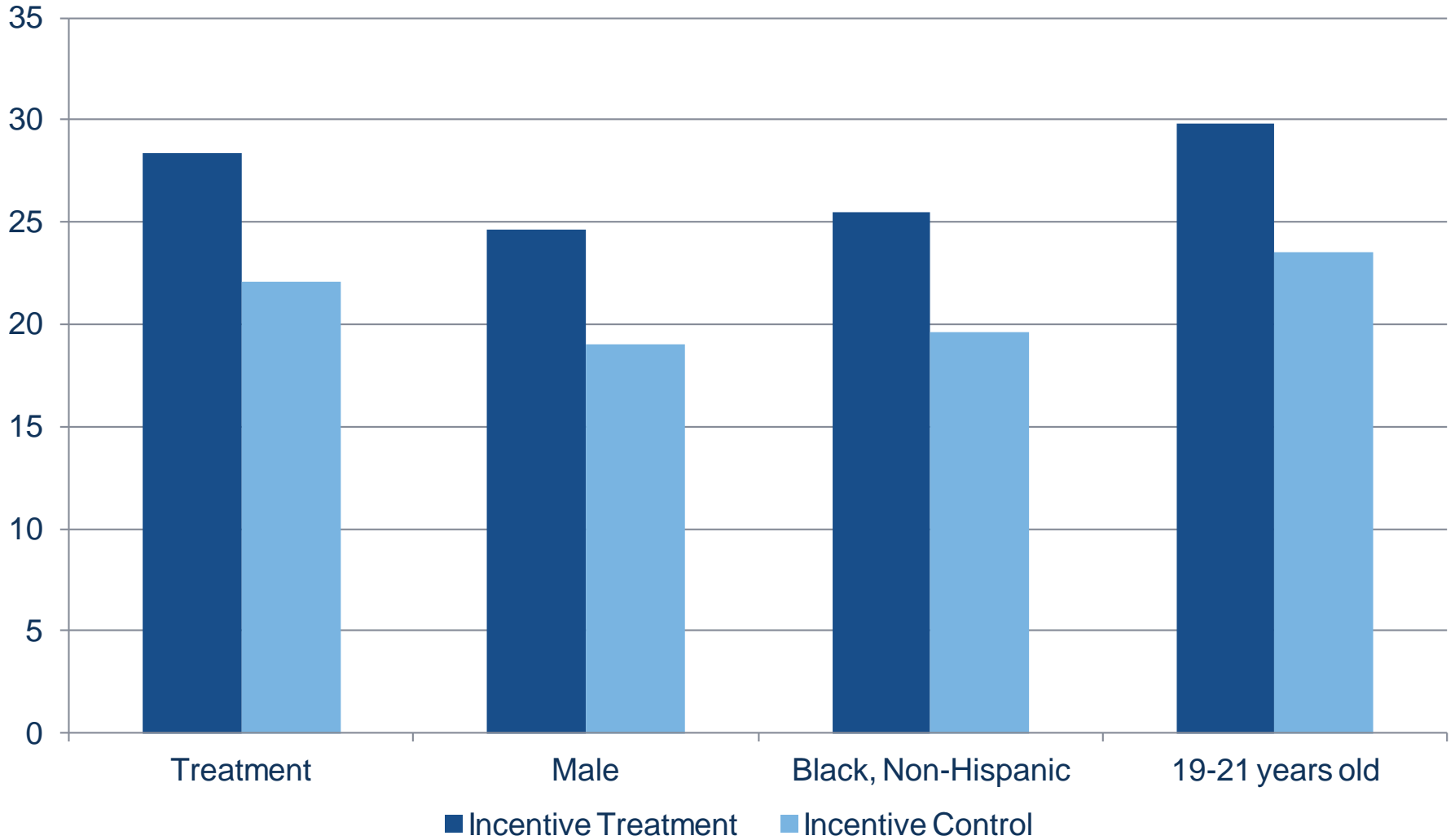
Incentive Experiment

- **Do targeted approaches such as differential incentives affect response?**
- **Early bird special**
 - **Treatment: \$40 gift card for completing in first 4 weeks of data collection**
 - **Control: \$25 regardless of time of completion**

Early Completion by Mode



Early Completion by Demographics



12-Month Follow-up Survey: Where Things Stand

- Achieving 80 percent response rate as cohorts close down

Response Rates Based on Eight Closed Cohorts (n = 2, 175)

RA Outcome	Web	Phone	Field	Overall
Treatment	27.6	38.8	16.6	83.0
Control	26.2	38.0	16.1	80.2
Total	27.1	38.5	16.5	82.1

Conclusions and next steps

- **What can be tailored for whom?**
 - **Tailor approaches based on contact preferences**
 - Delay phone follow-up for those who prefer email contact
 - Limit the web effort for those who do not prefer email
 - **Incentivizing early completion works for groups that are typically hard to get**
- **Coming soon...**
 - **Who completes by field?**
 - **How do our various outreach attempts (mail, email, phone, text, Facebook) relate to survey completion?**
 - **What are the impacts on cost of these strategies?**

Special Thanks To....

- MDRC
- U.S. Department of Labor
- Social Policy Research Associates (SPR)

For More Information

- Please contact
 - Lisbeth Goble
 - LGoble@mathematica-mpr.com
 - Jillian Stein
 - JStein@mathematica-mpr.com
 - Lisa Schwartz
 - LSchwartz@mathematica-mpr.com

Questions