

# Using Known Benchmarks to Inform the Accuracy of Online Research

---

September 10, 2007



# The online research landscape

---

After 20+ years, panels have emerged as main vehicle for custom studies

Repeated use has created sense of comfort; panels now being used for even high-value business decisions

Clients have also gained enough experience to discover that results are not always comparable to other modes or within mode

Commonly mentioned causes include “professional respondents,” “cheaters,” other culprits

**In many areas, online panel research has jettisoned our historical assumptions about research methods, quality**

# How do we study the accuracy of online research and what to do about it?

---

Longstanding research practices offer one solution: Comparisons to known and accepted benchmarks.

# Landmark study of data quality and vendor accuracy

---

Sponsored by the Stanford Institute for the Quantitative Study of Society

All vendors fielded identical questionnaires in '04.

Objective benchmarks (Census, CDC, U.S. Dept. of State, MRI, Other sources) gathered for comparison.

- Suggested for ongoing evaluation of accuracy

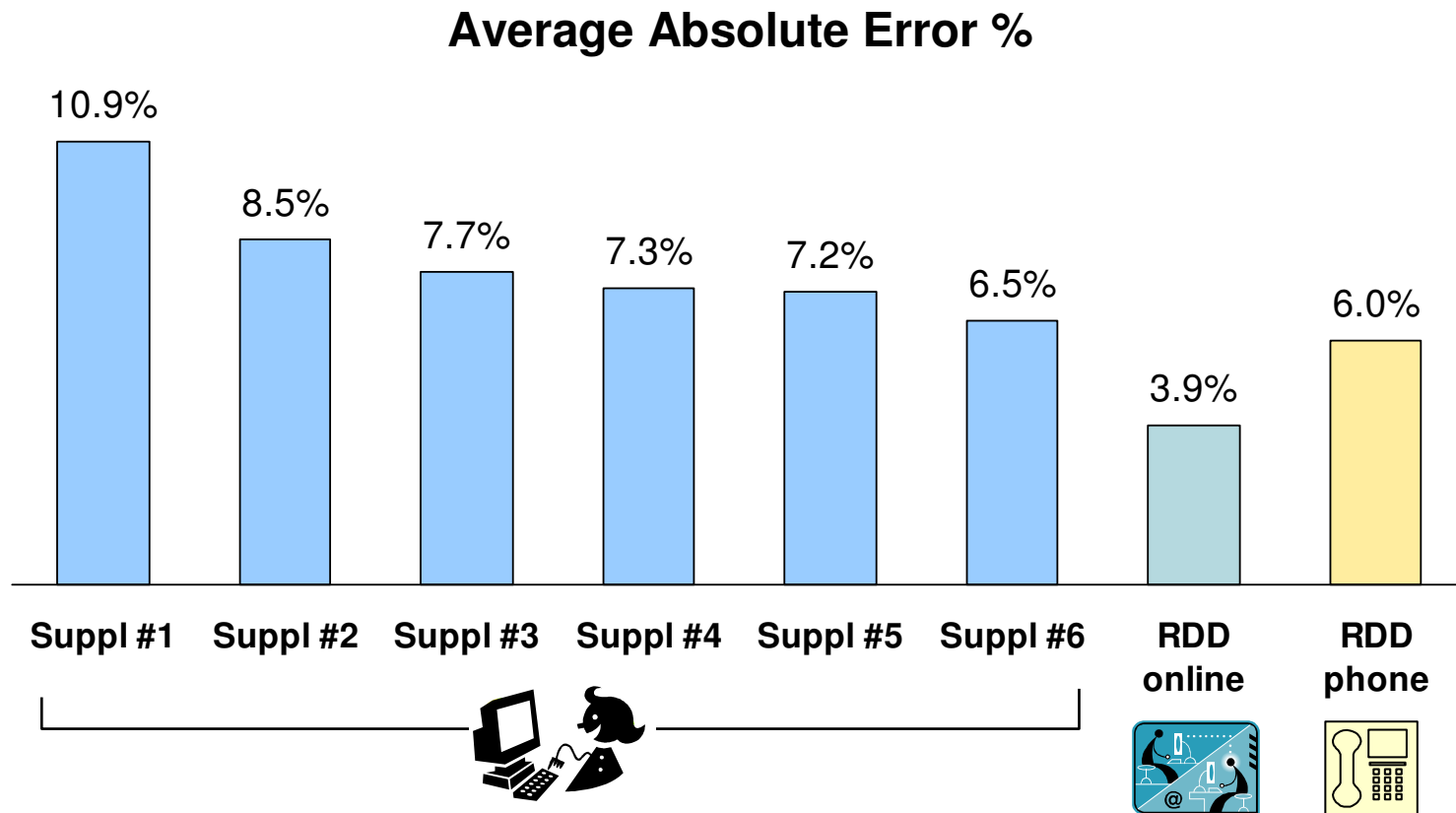
Apples-to-apples comparison of data from seven online research companies, one high-priced telephone vendor.

- Online vendors: SPSS, GoZing, Greenfield, Harris, Knowledge Networks, SSI, SurveyDirect
- Telephone vendor: SRBI

Also reveals key differences on participation rates and breakoffs that affect accuracy of results.

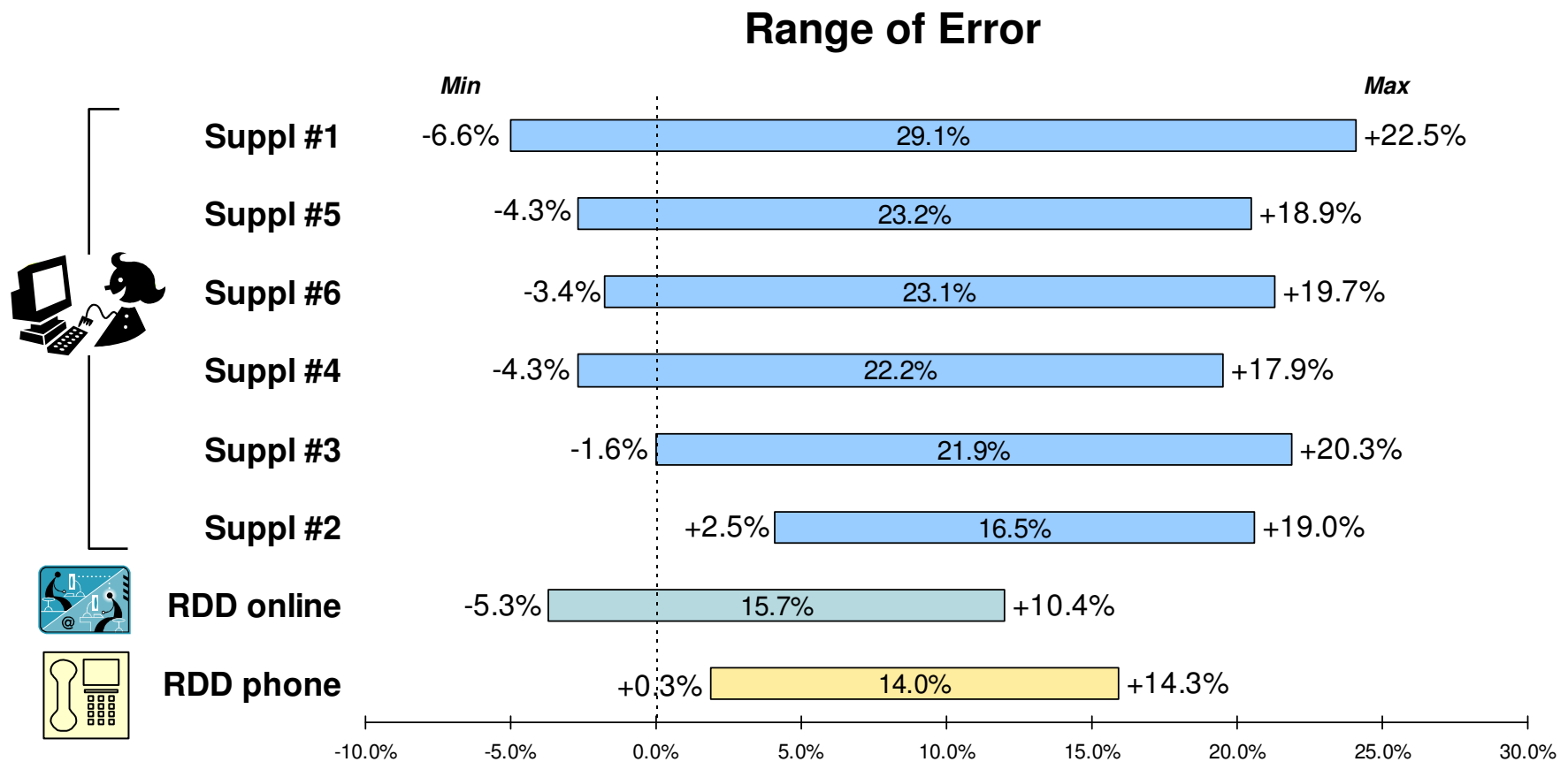
# Finding: Error rates in volunteer online panels do not vary widely

---



Includes: Primary and Secondary Demographics as well as product usage and attitude items.  
Average Absolute Error = Absolute Value (Benchmark – Survey Estimate)

# Finding: Volunteer samples tend to over-estimate benchmarks



# Summaries: Accuracy Against Benchmarks

---




*We focus here on modal response categories; average errors are much smaller than those that follow*

---




Variability exists on issues  
key to consumer business,

Research demonstrates the stakes  
behind accuracy




# Accuracy of primary demographics

								 RDD online	 RDD phone
	Bench- mark %	Suppl #1	Suppl #2	Suppl #3	Suppl #4	Suppl #5	Suppl #6		
		Percentage Error							
Aged 35-44	20.7%	0.5%	0.8%	-0.9%	-1.6%	-3.7%	-3.5%	-0.4%	0.1%
White	82.2%	0.1%	7.5%	-0.2%	-35.7%	5.0%	6.6%	-0.2%	-3.6%
Female	51.8%	-2.2%	-3.1%	0.3%	3.6%	0.4%	-1.1%	0.5%	1.3%
High school graduates	28.6%	-8.2%	-12.4%	11.6%	-12.0%	-9.6%	-12.1%	3.7%	5.2%
Income \$50-75,000	19.5%	-3.5%	3.4%	-2.9%	-2.1%	-0.3%	1.4%	0.4%	1.0%
<i>Mean absolute error</i>		2.9%	5.4%	3.2%	11.0%	3.8%	5.0%	1.0%	2.3%




# Accuracy of secondary demographics

								 RDD online	 RDD phone
	Bench- mark %	Suppl #1	Suppl #2	Suppl #3	Suppl #4	Suppl #5	Suppl #6		
		Percentage Error							
Married	56.6%	-3.0%	4.9%	2.9%	-11.0%	1.6%	0.2%	0.6%	3.1%
Worked for pay last week	63.4%	-3.1%	3.6%	-1.4%	-0.1%	-7.6%	-2.2%	-0.5%	-4.1%
Owns home	67.9%	-2.9%	3.8%	3.9%	-15.4%	1.4%	3.3%	3.0%	9.3%
3 Bedrooms in home	43.2%	-2.1%	1.9%	0.6%	-6.3%	-1.5%	2.0%	2.9%	3.0%
2 Vehicles in household	42.6%	-1.8%	4.3%	1.6%	-1.1%	-0.6%	-0.8%	5.4%	-0.8%
<i>Mean absolute error</i>		2.6%	3.7%	2.1%	6.8%	2.5%	1.7%	2.5%	4.1%

# Lifestyle and other variables

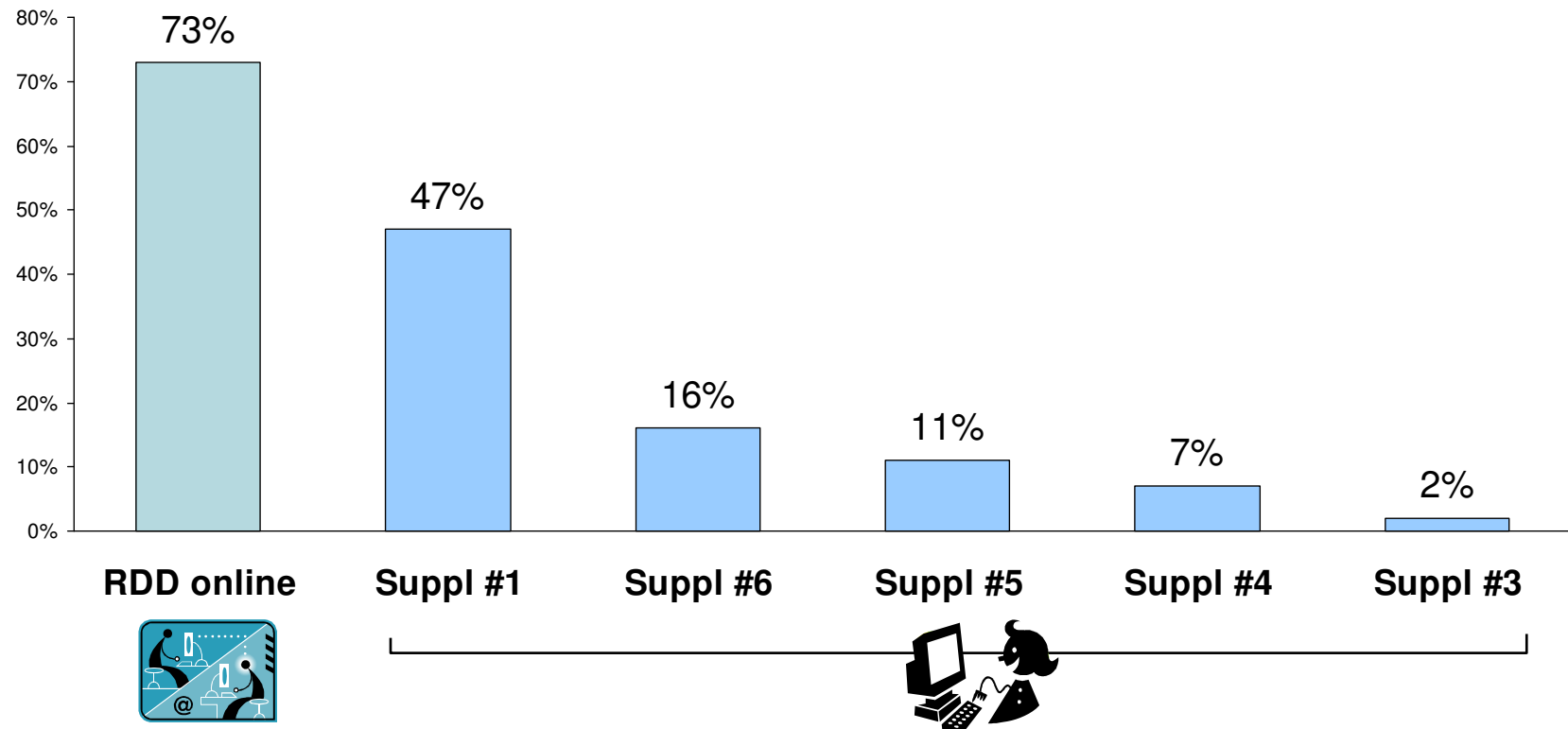
				
	<b>Benchmark</b>	<b>RDD phone</b>	<b>RDD online</b>	<b>All Volunteer</b>
	<u>%</u>	<u>%</u>	<u>%</u>	<u>%</u>
Last trip on American Airlines (Bureau of Trans. Services)	17.3	31.6	<b>16.7</b>	20.1
Smoke every day/occasionally (CDC)	21.6	25.7	<b>24.0</b>	29.7
Own valid U.S. passport (U.S. Dept. of State)	20.3	31.5	<b>27.2</b>	34.1
Have current driver's license (U.S. Census, Stat. Abstract)	89.1	92.7	<b>88.9</b>	94.2
Traveled in last year (MRI)	55.1	65.4	<b>65.5</b>	74.8
Member frequent flyer program (MRI)	17.8	21.5	<b>22.6</b>	32.7
Rented a car (MRI)	18.4	18.7	<b>19.3</b>	27.1
Exercised in last month (CDC)	76.0	80.3	<b>70.7</b>	76.2
Magazine Subscriptions (Audit Bureau of Circulation)	4.9	9.0	<b>6.2</b>	11.4
Seen movie last month, age 60+	8.0	15.2	<b>14.5</b>	15.6

# Lifestyle and other variables

									
	Bench- mark	RDD phone	RDD online	Suppl #1	Suppl #2	Suppl #3	Suppl #4	Suppl #5	Suppl #6
	%	%	%	%	%	%	%	%	%
Last trip on American Airlines (Bureau of Trans. Services)	17.3	31.6	<b>16.7</b>	21.5	17.8	18.6	24.5	20.1	17.8
Smoke every day/occasionally (CDC)	21.6	25.7	<b>24.0</b>	26.8	26.3	29.6	29.8	31.3	34.2
Own valid U.S. passport (U.S. Dept. of State)	20.3	31.5	<b>27.2</b>	30.7	34.6	34.0	37.4	31.1	36.7
Have current driver's license (U.S. Census, Stat. Abstract)	89.1	92.7	<b>88.9</b>	93.7	95.0	96.1	94.8	93.0	92.7
Traveled in last year (MRI)	55.1	65.4	<b>65.5</b>	74.8	74.0	75.4	74.1	73.0	77.6
Member frequent flyer program (MRI)	17.8	21.5	<b>22.6</b>	31.3	32.2	35.7	35.3	29.2	32.3
Rented a car (MRI)	18.4	18.7	<b>19.3</b>	24.8	22.9	26.3	31.5	23.9	33.1
Exercised in last month (CDC)	76.0	80.3	<b>70.7</b>	72.6	71.7	74.4	80.8	71.7	86.2
Magazine Subscriptions (Audit Bureau of Circulation)	4.9	9.0	<b>6.2</b>	9.5	10.6	10.0	10.4	11.8	16.1
Seen movie last month, age 60+	8.0	15.2	<b>14.5</b>	15.1	14.5	10.1	22.6	20.9	1.4

# Effective panel size varies by cooperation rate

Source: *Comparing the Results of Probability and Non-Probability Sample Surveys*, presented at 2005 AAPOR Conference



Thought experiment: How big is your panel?

The effective size of a 2-million-person panel is  $2,000,000 \times 2\% \rightarrow 40,000$

# Conclusions

---

Across many measures there was remarkable comparability of results, yet for known benchmarks the differences were not trivial

We cannot jettison longstanding research assumptions, methods that are crucial to our understanding of accuracy

Comparisons to benchmarks affirm client concern about variation of results and projectability

There is an ongoing need to keep this research current and use it as an ongoing metric

Transparency is essential to establishing true measures of quality, building client confidence