

Can We Replace Employment Questions with Auxiliary Data from LinkedIn or Résumés?

Jennifer Sinibaldi, PhD

Center for Behavioral Science Methods

Overview of Experiment

- Interested in reducing respondent burden
- Job information asked across multiple surveys
- 4 options to provide current job information
 - Answer questions
 - Upload résumé
 - Copy/paste résumé
 - Provide LinkedIn account
- Do people prefer to provide auxiliary data rather than answer survey questions?
- Unique: Involves *respondent-provided* auxiliary data, rather than linkage or scraping from an external source

Research Questions

1. What did respondents choose to do?
2. Were there demographic differences between what respondents chose to provide?
3. How many respondents provided:
 - any data,
 - valid data,
 - valid data for all 3 details about their job?
4. Why did respondents choose what they did?

HH Panel Survey Questions

Data Collection

- Questions on Census Household Panel (now HTOPS)
- Collected November 2024 (19 Nov – 3 Dec)
- Only people who replied that they were working were eligible

Auxiliary information or survey questions – V1

JOB1_FILTER. You can provide a LinkedIn account or résumé, or answer questions about your main job. Which of these would you like to do?

- Provide a link to your LinkedIn account
- Upload your résumé
- Copy and paste information from your résumé
- Answer questions



Provide LinkedIn account

JOB_LNK. Provide a link to your LinkedIn page. The url should look something like: `linkedin.com/in/first-last-####/`



Upload résumé

JOB_CV. Use the link below to upload your current résumé.

Drop files or click here to upload



Copy/Paste résumé

Copy and paste the information about your main job as a/an Survey Methodologist from your current résumé. Make sure to include the dates of employment, job title, and duties.



What is your job title?

Survey Questions

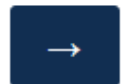
When did you start working as a/an Survey Statistician for your current employer?

Month

Year

Describe your typical activities or duties at this job.

Examples: instruct and evaluate students and create lesson plans, assemble and install pipe section and review building plans for work details.



Methodology & Data

Brief overview

- A lot of analysis was done manually
 - LinkedIn analysis: Search, view, and screenshot each profile page
Manually code data into Excel logs
Details of methods in DC-AAPOR Review Preview Conference, Sept '25
 - Résumé upload and copy/paste: Visually viewed each file and recoded in Stata to determine missing and valid data
 - Survey questions: Reviewed titles and duties to determine valid data
- Different decisions as to whether something “counted” or didn’t count would change the analysis slightly

Data for analysis

- HH Panel response: 9,500
 - Pathed to Job experiment: 7,000 (74%)
 - Working: 4,000 (57%)
 - Received the question asking them to provide auxiliary info: **4,000** (100%)

Results

1. What did respondents choose to do?

n=4,000

Aux Data	Chose	n
LinkedIn	11%	450
Resume upload	5%	200
Resume copy	1%	40
Survey Q	82%	3,300
Missing	1%	30

2. Were there demographic differences?

n=4,000		LinkedIn n=450	Resume Upload n=200	Resume Copy n=40	Survey Qs n=3,300
Sex	n=3700				
Male	44%	13%	4%	1%	82%
Female		9%	5%	1%	84%
Age	n=4000				
18-49	55%	14%	6%	1%	79%
50 or older		8%	3%	1%	87%
Language	n=4000				
English	98%	11%	5%	1%	83%
Spanish		10%	10%	10%	80%
Education	n=2800				
< College	37%	4%	5%	1%	89%
College degree	34%	15%	4%	2%	79%
Grad degree	30%	15%	5%	1%	80%

All X² tests significant

3. How many provided any data?

	n	Provided Data
LinkedIn	450	94%
Resume Upload	200	78%
Resume Copy	40	60%
Survey Qs	3,300	
Job title		99.3%
Start date		98.5%
Job duties		96.9%
Data for all 3		96.5%
Data for none		0.5%

3. How many provided valid data?

	n	Provided Data	Provided Valid Data
LinkedIn	450	94%	83%
Resume Upload	200	78%	78%
Resume Copy	40	60%	50%
Survey Qs	3,300		
Job title		99.3%	99.2%
Start date		98.5%	98.4%
Job duties		96.9%	96.0%
Data for all 3		96.5%	--
Data for none		0.5%	

3. How many provided 3 valid pieces of data?

	n	Provided Data	Provided Valid Data	Provided All 3 Data Points
LinkedIn	450	94%	83%	35%
Resume Upload	200	78%	78%	59%
Resume Copy	40	60%	50%	40%
Survey Qs	3,300			
Job title		99%	99%	--
Start date		99%	98%	--
Job duties		97%	96%	--
Data for all 3		96%	--	95%

4. What was fastest to provide?

Fastest	Overall	LinkedIn	Upload	Copy	Questions
	n=3,900	n=450	n=200	n=30	n=3,200
LinkedIn	14%	94%	5%	10%	4%
Resume Upload	7%	2%	86%	20%	3%
Resume Copy	1%	0%	2%	40%	0%
Survey Qs	78%	4%	7%	30%	92%

4. Why did respondents choose what they did?

Reason	n	Overall
Easiest	3,900	62%
Fastest	3,900	44%
Info less detailed	3,900	10%
Info more detailed	3,900	6%
Info less up-to-date	3,900	1%
Info more up-to-date	3,900	12%
Control what info provided	3,900	23%
Other options harder on device	3,900	13%
Reduced # of questions to answer	650	25%
No LinkedIn account	3,500	27%
LinkedIn is private	3,500	6%
No resume	3,700	10%
Resume not up-to-date	3,700	25%

4. Why did respondents choose what they did?

Reason	LinkedIn n=450	Upload n=200	Copy n=40	Questions n=3,300
Easiest	84%	81%	50%	58%
Fastest	67%	68%	30%	40%
Info less detailed	4%	3%	10%	11%
Info more detailed	14%	28%	10%	4%
Info less up-to-date	0%	1%	0%	1%
Info more up-to-date	22%	26%	10%	10%
Control what info provided	6%	8%	30%	26%
Other options harder on device	12%	2%	10%	14%
Reduced # of questions to answer	28%	22%	10%	
No LinkedIn account		14%	20%	27%
LinkedIn is private		3%	10%	6%
No resume	3%			11%
Resume not up-to-date	11%			27%



4. Why didn't respondent provide data?

~100 cases overall, including ~30 who didn't answer the filter question

- **Thought I did:** ~ 10% or less of nonrespondents for each data type
- **Technology problem:** 40% LinkedIn nonrespondents selected it
- **Didn't understand:** ~ 10% or less of nonrespondents for each data type
- **Skipped by accident:** ~ 10% or less of nonrespondents for each data type
- **Didn't want to:** Highest for each data type except LinkedIn
 - 53% Overall; 70% for upload; 60% copy/paste; 60% survey questions; 30% LinkedIn; 40% nonrespondent to filter question
- **Took too much time:** Highest for those who skipped survey questions (20%)

Conclusions

Conclusions

- Panelist prefer to answer survey questions (and the data are valid)
 - Why?
 - Conditioning– they are used to it, and can do it quickly
 - they know they can skip what they don't want to answer
 - Self-selection – these are already people who like to answer questions
- LinkedIn may be an option but will involve a lot of manual look up and the duties will be missing

Thank you

Jennifer.Sinibaldi@census.gov

Main job

JOB. What is your job? If you have more than one job, choose your main job.

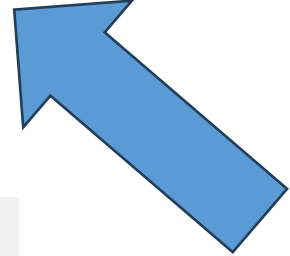
Examples: fourth-grade teacher, entry-level plumber, web developer, etc.

Job:



Auxiliary information or survey questions – V2

JOB2_FILTER. You can provide a LinkedIn account or résumé instead of answering questions about your main job. Which of these would you like to do?



- Provide a link to your LinkedIn account
- Upload your résumé
- Copy and paste information from your résumé
- Answer questions



1. Was there a difference between the 2 versions?

	V1 “or” n=2,000	V2 “instead of” n=2,000
LinkedIn	12%	10%
Resume upload	5%	5%
Resume copy	1%	1%
Survey Q	81%	83%
Missing	1%	1%

No significant difference