

Sampling homeless youth

Committee on Law and Justice

May 4, 2023

Daniela Golinelli

dgolinelli@mathematica-mpr.com



Objective

- / Obtain large probability sample of homeless youth**
- / Target population: emancipated homeless youth ages 13-23 in Los Angeles County**
- / Estimate rates of risky behaviors**



Challenges

- / A sampling frame of homeless youth is not available**
- / Most studies rely on convenience samples or sample youth from a limited set of locations**
- / Not all homeless youth use services**
- / Surveying homeless youth is costly**



Solution: Location sampling

/ Constructing a frame of locations with good coverage is complex and expensive

- Multiple location frames: service and natural sites

/ Participants can enter the sample at multiple time points and places (multiplicity issue)

- weight adjustment to obtain “visitors” rather “visits”
- collect additional information to derive the weight adjustment



Focus on multiple location frames

/ **Build list of sites by working with**

- Service providers
- Outreach agencies
- Police

/ **Assess bias induced from using a limited locations frame**

- Restricting the type of locations or sites
- Restricting the geographic regions



The Connect! Study

- / Probability sample of 419 homeless youth from Los Angeles County**
- / 41 sites: 15 shelters, 7 drop-in centers, and 19 street hangouts**
- / Four regions: Westside, Hollywood, Long Beach, and Pasadena**



Distribution of locations by type and region

Region	Types of sites			
	Shelters	Drop-in Centers	Hangouts	Total
Westside	0	3	7	10
Hollywood	8	2	5	15
Long Beach	4	2	4	10
Pasadena	3	0	3	6
Total	15	7	19	41





Participants distribution by site type

	Sites visited in last 30 days				
	Shelter?	Drop-in?	Hangouts?	# respondents	Weighted percentage
Shelter only →	Yes	No	No	36	8%
Drop-ins only →					
Hangouts only →					
# respondents					
Weighted percentage	44%				





Participants distribution by site type

	Sites visited in last 30 days				
	Shelter?	Drop-in?	Hangouts?	# respondents	Weighted percentage
Shelter only →	Yes	No	No	36	8%
	Yes	Yes	No	9	3%
	Yes	No	Yes	97	16%
	Yes	Yes	Yes	90	17%
Drop-ins only →	No	Yes	No	4	1%
	No	Yes	Yes	137	37%
Hangouts only →	No	No	Yes	46	17%
# respondents	232	240	370	419	
Weighted percentage	44%	59%	87%		



Restricting the location frame by type

Respondent characteristic		Overall	Shelters	Drop-ins	Hangouts
Age:		20.1	19.4*	20.9*	20.3*
Male (%):		61.2	57.7	70.3*	61.8
Race/ethnicity(%):	African American				
	White				
	Hispanic				
	Other/Mixed				
Ever spent night in jail (%)					
Traveler (%)					
Ever in foster care (%)					
Total years homeless					
Child abuse (%):	Physical				
Heavy drinking, past month (%)					
Marijuana use, past month (%)					
Hard drug use, past month (%)					
Any unprotected sex					
# of sex partners, past 3 months					
# of alters homeless					
# of alters drink or use drugs					
# of alters isolates					



Restricting the location frame by type

Respondent characteristic		Overall	Shelters	Drop-ins	Hangouts
Age:		20.1	19.4*	20.9*	20.3*
Male (%):		61.2	57.7	70.3*	61.8
Race/ethnicity(%):	African American	23.9	35.6*	27.6*	21.9*
	White	34.0	16.4	31.8	37.5
	Hispanic	20.0	27.9	15.5	18.4
	Other/Mixed	22.1	20.1	25.1	22.2
Ever spent night in jail (%)		68.0	50.7*	75.1*	70.9*
Traveler (%)		34.2	14.1*	38.6	39.1*
Ever in foster care (%)		34.6	38.3	39.7*	33.7
Total years homeless		3.0	2.5*	3.4*	3.2*
Child abuse (%):	Physical	47.7	47.7	53.6*	47.9
Heavy drinking, past month (%)		40.2	22.0*	40.9	42.9*
Marijuana use, past month (%)		66.4	44.3*	69.8	72.2*
Hard drug use, past month (%)		36.2	19.7*	38.5	39.3*
Any unprotected sex		52.7	38.4*	56.0	54.0
# of sex partners, past 3 months		1.7	1.3*	2.0*	1.9*
# of alters homeless		6.3	3.4*	7.2*	6.9*
# of alters drink or use drugs		12.1	9.6*	12.3	12.7*
# of alters isolates		6.4	5.3*	7.1*	6.8*



Participants distribution by region

	Regions visited in last 30 days					
	Westside?	Hollywood?	Long Beach?	Other areas?	# of respondents	Weighted %
Westside only →	Yes	No	No	No	42	15.6%
	Yes	Yes	No	No	61	17.8%
	Yes	No	Yes	No	3	0.7%
	Yes	No	No	Yes	6	1.1%
	Yes	Yes	Yes	No	10	2.6%
	Yes	Yes	No	Yes	24	4.6%
	Yes	No	Yes	Yes	2	0.3%
	Yes	Yes	Yes	Yes	18	4.1%
Hollywood only →	No	Yes	No	No	135	33.2%
	No	Yes	Yes	No	14	2.2%
	No	Yes	No	Yes	33	6.4%
	No	Yes	Yes	Yes	4	0.4%
Long Beach only →	No	No	Yes	No	46	7.1%
	No	No	Yes	Yes	3	0.4%
Other areas only →	No	No	No	Yes	18	3.4%
# of respondents	166	299	100	108		
Weighted %	46.8%	71.4%	17.8%	20.7%		



Restricting the regions

Respondent characteristic		Overall	Westside	Hollywood	Westside or Hollywood
Age		20.1	20.4	20.1	20.2
Male (%):		61.2	61.1	58.9	61.1
Race/ethnicity(%):	African American	23.9	13.5*	28.9*	24.5
	White	34.0	47.6	25.5	33.8
	Hispanic	20.0	13.6	20.7	18.9
	Other/Mixed	22.1	25.3	24.9	22.7
Ever spent night in jail (%)		68	74.3	65.5	68.8
Traveler (%)		34.2	50.2*	31.3	37.2*
Ever in foster care (%)		34.6	31.1	35.5	35.0
Total years homeless		3.0	3.5*	2.9	3.1*
Heavy drinking, past month (%)		40.2	51.7*	37.8	40.8
Marijuana use, past month (%)		66.4	82.7*	62.6	67.8
Hard drug use, past month (%)		36.2	49.2*	37.1	37.0
Any unprotected sex		52.7	59.8*	50.8	53.6
# of sex partners		1.7	1.9	1.8	1.8
# of alters who are homeless		6.3	8.6*	5.8	6.5*
# of alters who drink or use drugs		12.1	14.4*	11.8	12.4*
# of alters who are isolates		6.4	6.8	6.3	6.6



Focus on multiplicity issue

- / **Homeless youth can enter the sample in multiple ways and times during the field period**
- / **Additional questions used to measure visit frequency:**
 - most recent episode of homelessness, its length, and how much was spent in the LA area;
 - Number of days in the past 30 days were spent at the various sites and regions
- / **Use this information to estimate a measure of visit frequency to correct the sampling weights**



Sampling design

/ **A 3-stage design:**

- Stage 1: Selection of sites
- Stage 2: Selection of site-days within sites
- Stage 3: Selection of youth within site-days



Constructing the weights

$$SW_{ij} = \frac{D_i}{d_i} \times$$

$$VF_{ijk} \propto \textit{months homeless} \times \# \textit{ visits past month}$$



Constructing the weights

$$SW_{ij} = \frac{D_i}{d_i} \times \frac{N_{ij}}{n_{ij}} \times$$

$$VF_{ijk} \propto \textit{months homeless} \times \# \textit{ visits past month}$$



Constructing the weights

$$SW_{ij} = \frac{D_i}{d_i} \times \frac{N_{ij}}{n_{ij}} \times \frac{1}{VF_{ijk}}$$

$VF_{ijk} \propto \text{months homeless} \times \# \text{ visits past month}$



Conclusions

- Location sampling is a successful strategy for sampling hard-to-reach populations
- Location sampling requires a good knowledge of the target population and a lot of planning for developing location frames with good coverage
- Often multiple location frames are necessary
- Limiting the location frame to services only or only one region can bias estimates
- Obtaining a measure of visit frequency is necessary for correcting the sampling weights and obtain a sample of youth rather than youth-visits
- Omitting the correction can bias the sample
- Location sampling can be costly



Sampling homeless youth

Committee on Law and Justice

May 4, 2023

Daniela Golinelli

dgolinelli@mathematica-mpr.com