Sampling, WTP and demand assessment: findings from a household survey in 3 urban slums of Mumbai

Luisa M. Mimmi, GPOBA

Project Objective: OBA subsidy for affordable access to Improved Electricity connection in 4 Mumbai Slums

- **Project Signed**: April 27, 2009
- **GPOBA contribution**: US$1.65 million grant
- **Other Stakeholders**: USAID Reliance Power, ICPCI
- **Original Output**: ~ 25,000 households with affordable LEGAL connections (+ safe wiring)

What is SAMPLING?

Asking a small group of individuals (*households / schools…*) about a SPECIFIC HYPOTHESIS (*relevant characteristic or outcome*) and TRUST that their answer is ALSO TRUE for the target population of reference

SAMPLING design: KEY INGREDIENTS

INGREDIENTS ->	RESULT

1) "EFFECT" ≈ Hypothesis that I am hoping to test
   (a) Effect size and
   (b) Effect dispersion (variability)
 	
2) "PRECISION of the test" -> usually **= 5%**
   Significance level α = probability that we will conclude that the intervention has an effect, when in reality it has no effect
 	
3) "POWER of the test" -> usually **= 80% or 90%**
   ($1-\beta$) = probability that we will conclude that the intervention has an effect, when in fact it does have an effect
 	
4) SAMPLE SIZE (n) -> it will depend from all of the above
	POWER CALCULATION =
	Compute 1) + 2) + 3) ingredients to get the minimum necessary (n)
	
	* Any statistical software (even some free ones) will do this for you….

SAMPLING INGREDIENT (1)
"Hypothesis": Question to ask the sample…

▸ *HYPOTHESIS* / "Desired Effect":

Affordability is the ISSUE:

[COST for LEGAL connection > COST ILLEGAL Connection]

Es: *[Rs 2,000 > Rs. 1,000]*

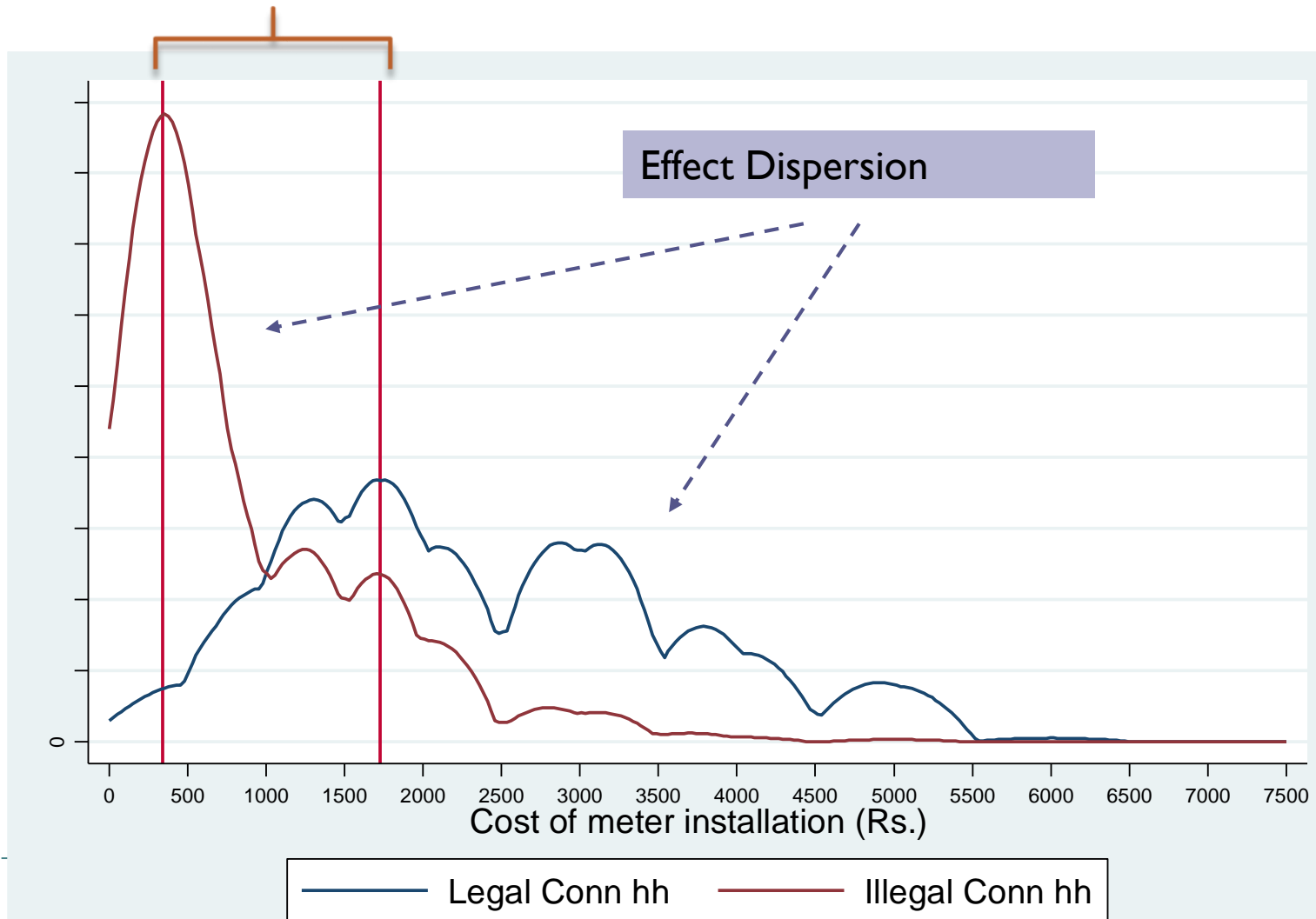▸ *If this is TRUE*: A **subsidy of Rs. 1,000** could make the LEGAL connection affordable for the poor

I am comparing 2 groups, so I will need 2 sub-samples

(n1, n2)

SAMPLING INGREDIENT (1)
Effect (Difference in COSTs) has 2 qualities:
(a) Effect size and (b) Effect dispersion

SAMPLING: COMPUTE POWER CALCULATION

INGREDIENTS ->	RESULT

1) "EFFECT" ≈ **Hypothesis**
  (a)Effect size
  [COST for LEGAL Conn > COST ILLEGAL Conn]
 	
  (b) Effect dispersion (variability or STD DEVIATION)*
 	
2) "PRECISION of the test" -> Significance level α
   usually **= 5%**
 3) "POWER of the test" -> (1– β) =
   usually **= 80% or 90%**
 	
4) SAMPLE SIZE (n) -> ???
	[COST for LEGAL conn > COST ILLEGAL Conn]
	Mean (legal) = Rs 1,7505
	Mena (illegal) = Rs. 340
	
	Std. Dev. (LEGAL) = 1225
	Std. Dev. (ILLEGAL) = 790
	
	   Significance level (α) = 5%
	      Power level (1– β) = 80%
	
	POWER CALCULATION =
	The minimum necessary sample size is:
	n1= 13
	n2 = 13
	
	(Extremely small n size is sufficient bc effect size is HUGE!)

SAMPLING DESIGN: (1) **STRATIFICATION** of target population
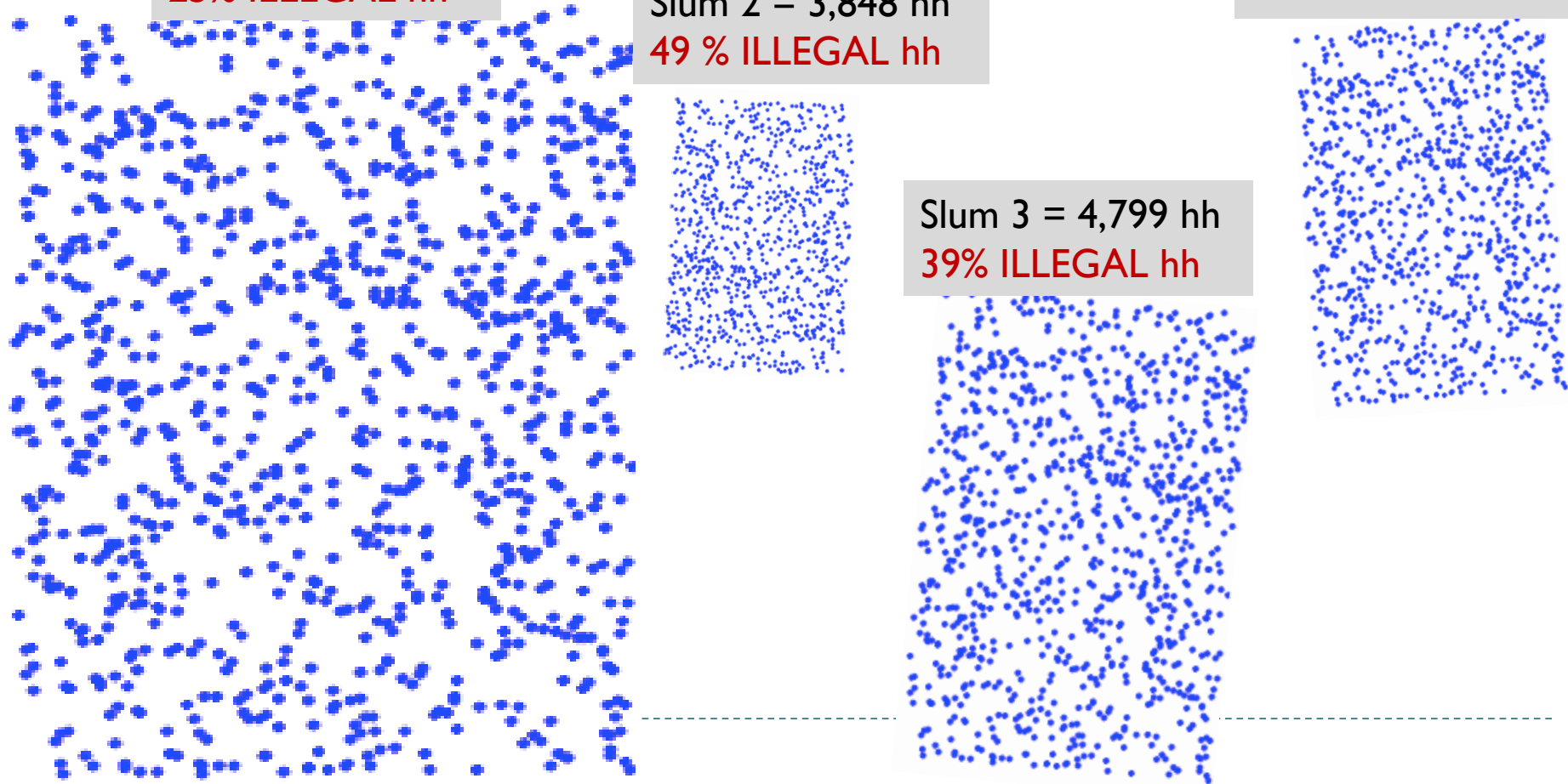
▸ N = Population of 4 slums = **41,984 hh**

Slum1 = 30,067 hh
25% ILLEGAL hh

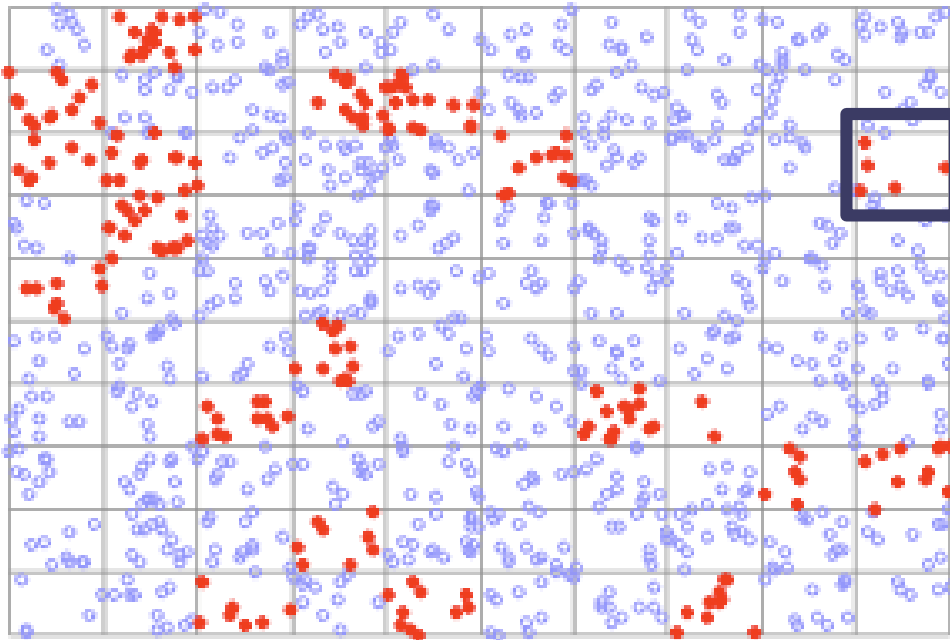Slum 2 = 3,848 hh
49 % ILLEGAL hh

Slum 3 = 4,799 hh
39% ILLEGAL hh

Slum 4 = 3,270 hh
49% ILLEGAL

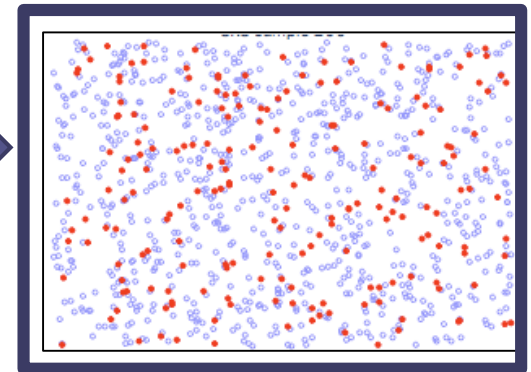# SAMPLING DESIGN (cont.): 2 STAGES with **CLUSTERS**

▸ **STAGE 1**:

Divide hh POPULATION in **BLOCKS**/Cluster **of 150 hh each**)

▸ 183 BLOCKS (out of 298) selected

▸ **STAGE 2**:

From each Block **pick RANDOMLY ~ 35 hh**

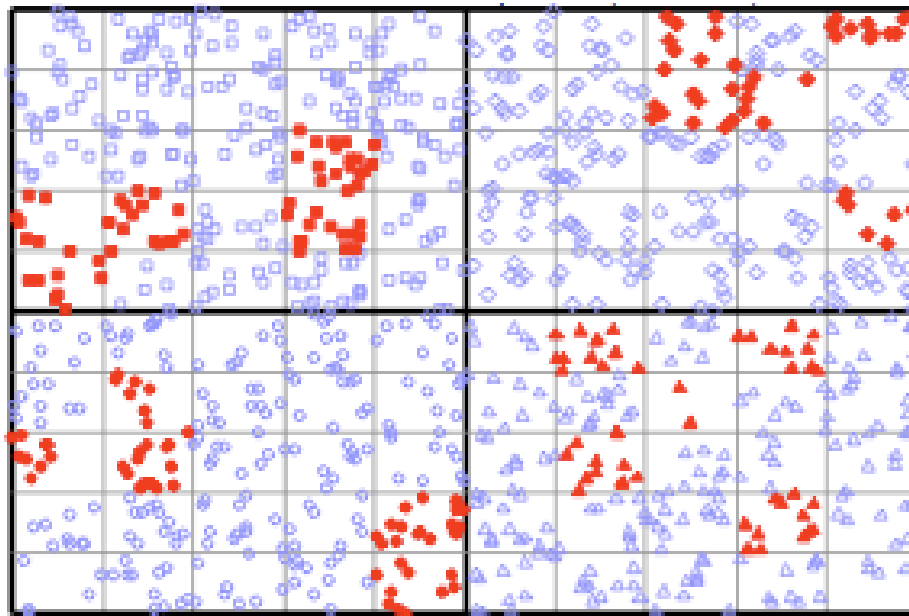▸ In each block ~ 35 (out of 150) households selected

SAMPLING DESIGN (cont.): *bottom line...*

"STRATIFIED TWO-STAGES random sampling"

(with 50% legal connection / 50% illegal connection in each slum/stratum)

n1 = 1357 hh

n4 = 600 hh

n2 = 443 hh

n= 600 hh

HOW will the 'SURVEY DESIGN' AFFECT THE ANALYSIS?

(and therefore the study conclusions)?

How do I correct for the "SAMPLING ERROR"?

1. Sampling WEIGHTS (correct for the over/underrepresentation of certain sub-groups)
   1. But I need to **know EXACTLY the probability of each unit to be picked** (i.e. I need a **SAMPLING FRAME**)

2. Any good Statistical Software can help "correcting" for the other **"Survey Design"** (strata + clusters) **errors**

Now let's Ask some questions....

Willingness To Pay = WTP

Measuring how much "utility" one derives from a good or service ….

In *normal conditions* it is captured by how much *I would spend* for it….

Luisa's WTP for a train ticket DC-NY?

AMTRAK Trip cost $150

SURVEY " I would pay …. $99.99 max"

CONCLUSION "If I get a $50 discount … next time I go to NY by train"
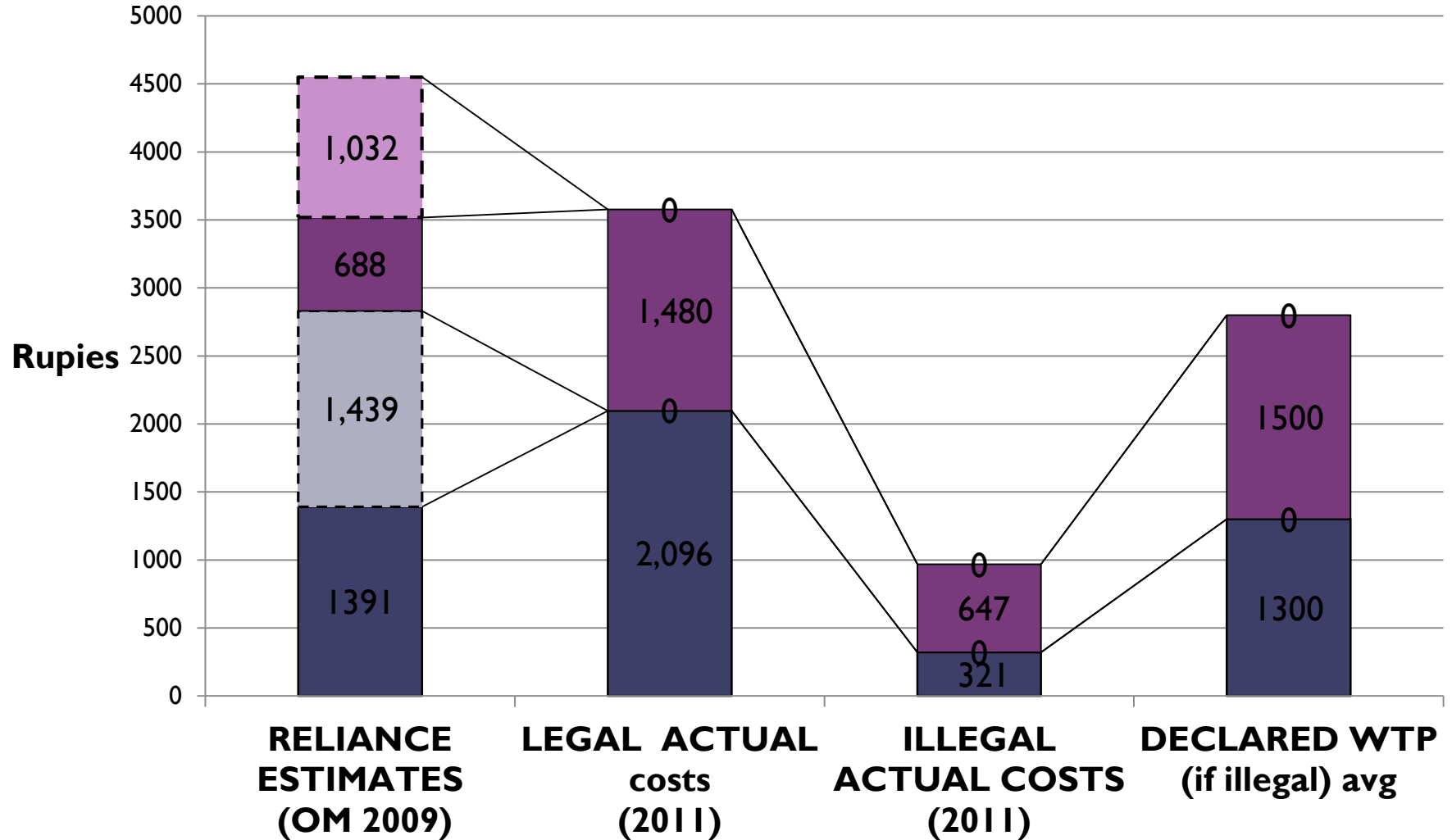
*Except: I **actually** take the DC2NY bus $25…*
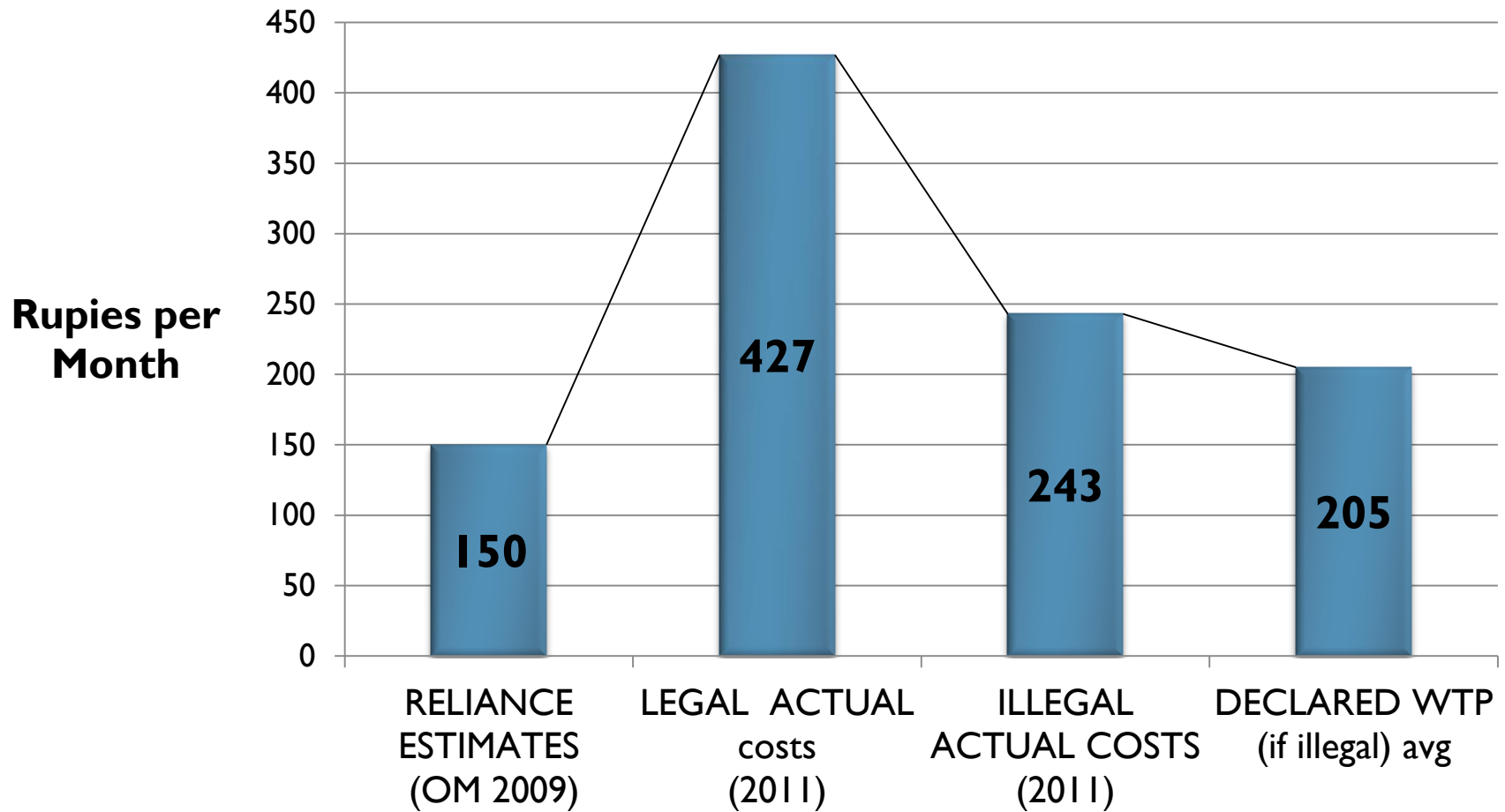*& GET ONE FREE every 5 rides !!!*

Comparative Analysis WTP for Connection (Installation fee + Internal Wiring)

(for households consuming 0 - 100 KWh/month (2,222 or 74% of sample)

Comparative Analysis WTP for Monthly Consumption Charge

(for households consuming 0 - 100 KWh/month (2,222 or 74% of sample)

Rupies per Month

450
400
350
300
250
200
150
100
50
0

RELIANCE ESTIMATES (OM 2009)	LEGAL ACTUAL costs (2011)	ILLEGAL ACTUAL COSTS (2011)	DECLARED WTP (if illegal) avg
150	427	243	205

■ Avg (inputed) Consumption charge (Rs. per month)

Intangible determinants of SWITCH: ILLEGAL -> LEGAL (responses from 1,088 _ILLEGAL_ hhs (74% sample) consuming 0- 100 KWh/month)

DRIVERS for regularization	CHALLENGES of regularization

• **82%** declare they **WANT Legal connection** (and 95% of them WITH safe INT WIRING)
 	
• _REASONS_:
   • 52% Better service
   • 46% Continuous supply
 	
• 33% had tried to get regular connection
   • -> _REL refused (40% requests)_
   • -> _lacked documents (25% requests)_
	• **40% believe** the **MONTHLY ELECTRICITY BILL would be too high** (if Legal)
	   • _only 30% are concerned with CONNECTION COST_
	
	• **78%** of these hh got the (illegal) electricity **connection from a LOCAL PERSON / SLUM LORD**…
	   • _survey team encountered some hostility when interviewing…_
	
	• Pus, **Mumbai's 60yrs record of failed slum rehabilitation promises**…very little TRUST left in the communities!!!

Final "REALITY CHECK"… when you set out to estimate/ verify/evaluate an outcome/impact:

1. <u>Spend enough time</u> (A LOT!) getting a "sense" of the desired effect in your target population
   1. Big enough to be detected in a convincing way?
   2. How about confounding factors / variability/ heterogeneous response?

2. Which <u>reasonably small sample </u>of units could effectively represent your target population?
   1. Cluster, stratify & weight according to your question…
   2. Think VEEERY carefully about HOW to ask the right questions to elicit the desired answers (*Think LOCAL)*

3. Have you considered other "factors" that could be closely linked to your outcome?

4. What about contextual elements that will likely affect your outcome? (informality, levels, history of the community, politics…)

Useful Resources / Readings

▸ Living Standards Measurement Study (LSMS) people at the BANK

    ▸ Training and Survey Clinics available…

    ▸ Tons of material about survey design:
http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESE
ARCH/EXTLSMS/0,,contentMDK:21555895~menuPK:4196884~page
PK:64168445~piPK:64168309~theSitePK:3358997,00.html

▸ SDN front office may give some support to IE…

▸ Good explanation of "**Hypothesis Testing and Statistical Power of a Test**" / sampling

    ▸ http://www.indiana.edu/~statmath/stat/all/power/power.pdf

▸ And more…

▸

"*A few observation and much reasoning lead to error; many observations and a little reasoning to truth.*"

Alexis Carrel (Nobel Prize in Medicine)

GRAZIE E ARRIVEDERCI !!!