# Using paradata to monitor interviewers' instrument navigation behavior and inform instrument technical design: A case study from a national household survey in Ghana 

Yu-chieh (Jay) Lin , Kyle Kwaiser, Gina-Qian Cheung, \& Jennifer Kelley,<br>Chicago<br>July 25-29, 2016<br>The 2nd International Conference on Survey Methods in Multinational, Multiregional and Multicultural Contexts (3MC)

## Overview

$>$ Background and instrument
$>$ How does the instrument design affect instrument navigation by the interviewers?
$>$ How does instrument navigation affect interview length?
$>$ Some observations \& next steps

## Ghana Socioeconomic Panel Study

> Revisit panel households at 3-4 year intervals for 20 years.
$>$ Sponsored by Economic Growth Center (EGC) at Yale University
$>$ Carried out by the Institute for Statistical, Social and Economic Research (ISSER) at the University of Ghana.
> First wave (baseline) was completed on paper between October 2009 and February 2010.
> Second wave was conducted on Computer-Assisted Personal Interviewing (CAPI) between March-December 2014.
> Collaborated with Survey Research Center (SRC) at University of Michigan.
> 334 enumeration areas country-wide. Sample size of 5009 households, with approximately 18,000 individuals. Also sample size of 500 split-off households were tracked and interviewed between January-June 2015.
> Interviews are NOT digital recorded for quality monitor purpose


## Challenges from Paper to CAI

$>$ Transition from complex grid designs on paper to a computer assisted interview (CAI) instrument.
> Need to have total flexibility to jump in/out from different sections of the instrument (depending on the availability of the respondents)
$>$ Need to track real-time status of interviewing progress on multiple respondents within the same instrument.

## INSTITUTE FOR SOCIAL RESEARCH <br> SURVEY RESEARCH CENTER <br> UNIVERSITY OF MICHIGAN



## NSTITUTE FOR SOCIAL RESEARCH

## SURVEY RESEARCH CENTER



## Instrument Sections*


© 2016 by the Regents of the University of Michigan

## Consent Has Completed



## Household Roster Has Completed

| 5 Blaise 4.8 Data Entry - cilblproilghana_plworkh householdsurvey |  |  |  |
| :---: | :---: | :---: | :---: |
| Forms Answer Help |  |  |  |
| HOUSEHOLD SURVEY Person Status |  |  |  |
| Survey Status |  |  |  |
| SOConsent Forms S01A: Consent | Done | S01B3: Consent for Under 26 | Not Started |
| Rosters |  |  |  |
| S01B2: Household Roster | Done | Person Sections | Not Started |
| 504: Plot Roster | NotStarted | Plot Sections | ----n/a---- |
| S05: Non-Farm Enterprise Roster | Not Started | Non-Farm Enterprise Sections | ----n/a--- |
| Household Level Sections |  |  |  |
| SO2B: Non-Resident Spouses | Not Started | S03Ai: Animals | Not Started |
| S02A: Non-Resident Relatives | Not Started | S03Aii: Tools | Not Started |
| S10C: Social Networking | Not Started | S03Aiii: Durable Goods | Not Started |
| S10D: Information Seeking | Not Started | S03Bi: Borrowing | Not Started |
| S11: Household Consumption | Not Started | S03Bii: Lending | Not Started |
| S12: Housing Characteristics | Not Started | s03Biii: Out-Transfers | ----n/a---- |
| SO4NN: Gathering | Not Started | S03Biv: In-Transfers | ----n/a--- |
|  |  | S03Bv: Savings | Not Started |
| Not all sections have been complete | ection and c | survey. |  |

## SURVEY RESEARCH CENTER

UNIVERSITY OF MICHIGAN
Blaise 4.8 Data Entry - c:\blprojไghana_plwork\householdsurvey

## Forms Answer Help <br> HOUSEHOLD SURVEY Person Status

| Name | Background | Employment | Education | Migration | Health | Womens Hea | Mens Health | Children | Pysch/Social |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| SAMUEL KANGA SNR (SNR) | Not Started | Not Started | Not Started | Not Started | Not Started | -n/a- | Not Started | -n/a-- | Not Started |
| SAMUEL KANGAJNR (JNR) | Not Started | Not Started | Not Started | Not Started | Not Started | -n/a- | Not Started | -n/a- | Not Started |
| RAMAT DOMSON | Not Started | Not Started | Not Started | Not Started | Not Started | Not Started | --n/a-- | -n/a- | --n/a- |
| NENYA HANGAH | Not Started | Not Started | Not Started | Not Started | Not Started | -n/a-- | Not Started | Not Started | Not Started |

Blaise 4.8 Data Entry - c:\blproj\ghana_ blwork\householdsurvey
Forms Answer Help
HOUSEHOLD SURVEY
Person Status

## Survey Status

## SOConsent Forms

S01A: Consent

Rosters

| S01B2: Household Roster | Done |
| :--- | :--- |
| S04: Plot Roster | Not Started |
| S05: Non-Farm Enterprise Roster | Not Started |

Household Level Sections

| SO2B: Non-Resident Spouses | $\underline{\text { Not Started }}$ |
| :--- | :--- |
| SO2A: Non-Resident Relatives | $\underline{\text { Not Started }}$ |
| S10C: Social Networking | $\underline{\text { Not Started }}$ |
| S10D: Information Seeking | $\underline{\text { Not Started }}$ |
| S11: Household Consumption | $\underline{\text { Not Started }}$ |
| S12: Housing Characteristics | $\underline{\text { Not Started }}$ |
| SO4NN: Gathering |  |

S01B3: Consent for Under 26

| Person Sections | Not Started |
| :--- | :--- |
| Plot Sections | $---n / a------n$ |
| Non-Farm Enterprise Sections | $---n / a--$ |


| S03Ai: Animals | Not Started |
| :---: | :---: |
| S03Aii: Tools | Not Started |
| S03Aiii: Durable Goods | Not Started |
| S03Bi: Borrowing | Not Started |
| SO3Bii: Lending | Not Started |
| S03Biii: Out-Transfers | ----n/a---- |
| S03Biv: In-Transfers | ----n/a---- |

Not Started

Not all sections have been completed! Please select a section and continue the survey.

## Survey Status

## SOConsent Forms

S01A: Consent
Done
S01B3: Consent for Under 26
Not Started

## Rosters

| SO1B2: Household Roster | Don |
| :--- | :--- |
| S04. Plot Roster | Don |

504: Plot Roster Done
S05: Non-Farm Enterprise Roster
Done

Household Level Sections
SO2B: Non-Resident Spouses
SO2A: Non-Resident Relatives
Not Started
Not Started
S10C: Social Networking
Not Started
S10D: Information Seeking
S11: Household Consumption
Not Started

S12: Housing Characteristics
Not Started

S04NN: Gathering
Not Started
Not Started
S01B3: Consent for Under 26 Not Started

| Person Sections | Not Started |
| :---: | :---: |
| Plot Sections | ---n/a- |
|  |  |


| S03Ai: Animals | Not Started |
| :---: | :---: |
| SO3Aii: Tools | Not Started |
| S03Aiii: Durable Goods | Not Started |
| S03Bi: Borrowing | Not Started |
| S03Bii: Lending | Not Started |
| S03Biii: Out-Transfers | ---n/a---- |
| S03Biv: In-Transfers | ...-n/a- |
| S03Bv: Savings | Not Started |

Not all sections have been completed! Please select a section and continue the survey.

## Parallel Block Instrument



Instrument types:
(1) Personal + Household
(2) Personal + Household + Plot
(3) Personal + Household + Enterprise
(4) Personal + Household + Plot + Enterprise



## Parallel Block Instrument



Instrument types:
(1) Personal + Household
(2) Personal + Household + Plot
(3) Personal + Household + Enterprise
(4) Personal + Household + Plot + Enterprise


Up to 8 Sections Per Plot

Up to 4 Sections Per Enterprise

NSTITUTE FOR SOCIAL RESEARCH

## SURVEY RESEARCH CENTER

UNIVERSITY OF MICHIGAN
Blaise 4.8 Data Entry - c:|blprojighana_plworkl householdsurvey

## Forms Answer Help <br> Survey Status



## SOConsent Forms

| S01A: Consent | Done | S01B3: Consent for Under 26 | Not Started |
| :---: | :---: | :---: | :---: |
| Rosters |  |  |  |
| S01B2: Household Roster | Done | Person Sections | Not Started |
| S04: Plot Roster | Done | Plot Sections | ---n/a---- |
| S05: Non-Farm Enterprise Roster | Done | Non-Farm Enterprise Sections | Not Started |
| Household Level Sections |  |  |  |
| SO2B: Non-Resident Spouses | Not Started | S03Ai: Animals | Not Started |
| SO2A: Non-Resident Relatives | Not Started | S03Aii: Tools | Not Started |
| S10C: Social Networking | Not Started | S03Aiii: Durable Goods | Not Started |
| S10D: Information Seeking | Not Started | S03Bi: Borrowing | Not Started |
| S11: Household Consumption | Not Started | SO3Bii: Lending | Not Started |
| S12: Housing Characteristics | Not Started | S03Biii: Out-Transfers | ---n/a---- |
| SO4NN: Gathering | Not Started | S03Biv: In-Transfers | ---n/a---- |
|  |  | SO3Bv: Savings | Not Started |

Not all sections have been completed! Please select a section and continue the survey.
Case Complete $\quad$ ॥

## Parallel Block Instrument



Instrument types:



## "Master" Dashooaro

| Forms Answer Help |  |  |  |
| :---: | :---: | :---: | :---: |
| HOUSEHOLD SURVEY Person Status | Enterprise Status Agriculture |  |  |
| Survey Status |  |  |  |
| SOConsent Forms |  |  |  |
| S01A: Consent | Done | S01B3: Consent for Under 26 | Not Started |
| Rosters |  |  |  |
| S01B2: Household Roster | Done | Person Sections | Not Started |
| S04: Plot Roster | Done | Plot Sections | Not Started |
| S05: Non-Farm Enterprise Roster | Done | Non-Farm Enterprise Sections | Not Started |
| Household Level Sections |  |  |  |
| S02B: Non-Resident Spouses | Not Started | S03Ai: Animals | Not Started |
| S02A: Non-Resident Relatives | Not Started | S03Aii: Tools | Not Started |
| S10C: Social Networking | Not Started | S03Aiii: Durable Goods | Not Started |
| S10D: Information Seeking | Not Started | S03Bi: Borrowing | Not Started |
| S11: Household Consumption | Not Started | S03Bii: Lending | Not Started |
| S12: Housing Characteristics | Not Started | S03Biii: Out-Transfers | ----n/a---- |
| S04NN: Gathering | Not Started | S03Biv: In-Transfers | ----n/a---- |
|  |  | S03Bv: Savings | Not Started |

Not all sections have been completed! Please select a section and continue the survey.

## "Benefits" of this design

> Iwers have a high-level of autonomy with respect to interview navigation.
$>$ Iwers are able to:
$>$ switch respondents easily.
$>$ jump to any section of questionnaire quickly.
> Development of a questionnaire "Dashboard" to show the status of all the questionnaire sections and all the respondents within the household.

## Questions for the design...

$>$ How does instrument design affect instrument navigation?
> Instrument parallel blocks: four instruments (household, personal, agriculture, enterprise) with multiple sections/blocks within each instrument.
$>$ How does instrument navigation affect interview length?
$>$ Order of interview initiation
> Movements between blocks
By using keystroke data (Paradata)

Timestamps
urs:Minutes:Seconds:Thousands of a second
"1/17/2012 9:00:06:304 AM","Enter Form:1","Key:3975053020 " « Sample ID "1/17/2012 9:00:06:304 AM","Metafile name:C:\blproj\HRS2012\work\HRS12.bmi" "1/17/2012 9:00:06:304 AM","Metafile timestamp:Friday, January 06, 2012 1:08:04 PM"
"1/17/2012 9:00:06:304 AM","WinUserName:14554015" « "1/17/2012 9:00:06:304 AM","DictionaryVersionInfo:0.0.0.0"

## Start IW

Audit trail file
"1/17/2012 9:00:12:702 AM","Enter Field:SecA.StartInterview.A007TRAlive_A","Status:Normal","Value:" "1/17/2012 9:00:13:965 AM","(KEY:)1[ENTR]" $\longleftarrow$ Time of first keystroke
"1/17/2012 9:00:14:276 AM","Action:Store Field Data","Field:SecA.StartInterview.A007TRAlive_A"
"1/17/2012 9:00:14:328 AM","Leave Field:SecA.StartInterview.A007TRAlive_A","Cause:Next Field",
"Status:Normal","Value:1"
"1/17/2012 9:02:51:681 AM","Enter Field:SecJ.WORKSTATUS.JO05MCurrEmpStatus[1]","Status:Normal","Value:" "1/17/2012 9:02:55:971 AM","(KEY:)15[BACK][BACK]5[ENTR]"
"1/17/2012 9:03:03:209 AM","Action:Store Field Data","Field:Sed.WORKSTATUS.JO05MCurrEmpStatus[1]"

Question with chang answer

Field","Status:Normal","Value:5"

```
"1/17/2012 9:13:24:923 AM","Enter Field:IWComplete","Status:Normal","Value:"
```

"1/17/2012 9:13:28:480 AM","(KEY:)1[ENTR]"
"1/17/2012 9:13:29:650 AM","Action:Store Field Data","Field:IWComplete"
"1/17/2012 9:13:29:728 AM","Leave Field:IWComplete","Cause:Next Field","Status:Normal","Value:1"
"1/17/2012 9:13:30:056 AM","Leave Field:IWComplete","Cause:Exit","Status:Normal","Value:1"
"1/17/2012 9:13:30:056 AM","Leave Form:1","Key:3975053020 "


## Most Common Block Moves All Types

> Edge Weight (number of times a move occurred) >= 500
> Movement within sections dominates
> Exceptions are rosters and Personal to Household


| ? Blaise 4.8 Data Entry - c:lblprojlghana_plwork\householdsurvey |  |  |  |  |  |  |  |  | -回 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Forms Answer Help |  |  |  |  |  |  |  |  |  |
| HOUSEHOLD SURVEY | Person Status Enterpris | Se Status Ag | riculture |  |  |  |  |  |  |
| Name | Background | Employment | Education | Migration | Health | Womens Healt | th Mens Health | Children | Pysch/Social |
| ADAM K (AK) | Started | Done | Done | Done | Started | --n/a- | Not Started | --n/a-- | Not Started |
| AMINAA (MINA) | Not Started | Not Started | Not Started | Not Started | Not Started | Not Started | ---n/a-- | --n/a-- | Not Started |
| ABDUL A | Not Started | Not Started | Not Started | Not Started | Not Started | --n/a- | Not Started | --n/a- | Not Started |
| TANLIDOW A | Not Started | Not Started | Not Started | Not Started | Not Started | --n/a- | --n/a-- | Not Started | --n/a-- |
| YUSSIF A | Done | --n/a-- | Not Started | --n/a-- | Not Started | --n/a- | --n/a-- | Not Started | --n/a- |
| LAlLATU A | Done | --n/a-- | --n/a-- | --n/a-- | Not Started | --n/a-- | --n/a-- | Not Started | --n/a-- |

## Most Common Block Moves All Types

> Tendency to move laterally or within the same questionnaire content
> Optional sections introduce multiple, common paths
© 2016 by the Regents of the University of Michigan


| HOUSEHOLD SURVEY Person Status | Enterprise StatusAgriculture |  |  |
| :---: | :---: | :---: | :---: |
| Survey Status |  |  |  |
| SOConsent Forms |  |  |  |
| S01A: Consent | Done | S01B3: Consent for Under 26 | Done |
| Rosters |  |  |  |
| S01B2: Household Roster | Done | Person Sections | Started |
| S04: Plot Roster | Done | Plot Sections | Done |
| S05: Non-Farm Enterprise Roster | Done | Non-Farm Enterprise Sections | Not Started |
| Household Level Sections |  |  |  |
| S02B: Non-Resident Spouses | Not Started | S03Ai: Animals | Done |
| SO2A: Non-Resident Relatives | Not Started | S03Aii: Tools | Started |
| S10C: Social Networking | Started | S03Aiii: Durable Goods | Started |
| S10D: Information Seeking | Done | S03Bi: Borrowing | Not Started |
| S11: Household Consumption | Not Started | S03Bii: Lending | Done |
| S12: Housing Characteristics | Not Started | S03Biii: Out-Transfers | ----n/a---- |
| SO4NN: Gathering | Done | S03Biv: In-Transfers | ----n/a---- |
|  |  | S03Bv: Savings | Done |

## Most Common Block Moves All Types

$>$ Tendency to work down the columns
$>$ Non-resident Relatives and Consumption introduce multiple common paths


## Moving Out of a Section

$>$ Type 4 interviews showing moves out of the Enterprise section
$>$ "fourlmportant" block has most exit moves

## Instrument Parallel Blocks

$>$ We should have some instructions about the optimal interviewing paths for the desired navigation
$>$ The parallel blocks programming needs to match with the optimal navigation design
$>$ The interviewer training needs to emphasize the design and avoid "jump around too much"
$>$ How does instrument navigation affect interview length?
$>$ Order of interview initiation
> Movements between blocks

## N = 4223 interviews

- Type 1) Personal + Household
- Type 2) Personal + Household + Plot
- Type 3) Personal + Household + Enterprise
- Type 4) Personal + Household + Plot + Enterprise


## Interview Length

- Adjusted mean 290 (+/-156) minutes
- Ranged from ~ $\mathbf{3 0}$ minutes to $\mathbf{2 2}$ hours
- Type 4 longer than all others $(P<0.05)$
- Type 1 shorter than all others ( $\mathrm{P}<0.05$ )
- Type 2 \& 3 not different




## Order of Section Entry

- Split sample into two groups:
- Completed rosters first
- Others
- T-test comparison of IW length


Household Roster

Consent under 26

Enterprise Sections

Agriculture Sections

Household
Sections

Personal Sections

Up to 14 Sections

Up to 8 Sections Per Person

Up to 4 Sections Per Enterprise

Up to 8 Sections Per Plot

- Rosters-first group showed significantly lower interview lengths
- Ave Roster First: 257.3
- Ave Others: 280.3
- $P<.0001$


## Some Observations (Roster-First)

$>$ Region (10):
> Only 2 regions; no significant differences in four regions.
$>$ HH Type (1-4):
$>$ Only HH without any or with both agriculture/enterprises sections.
$>\mathrm{HH}$ size (1-17):
$>$ Single-person HH ; no significant differences in HH size $=2$ or 3 .
$>$ Team leadership (23):
$>$ Only in 8 teams; no significant differences in 4 teams.

## Number of Block Moves <br> Per Block

- On a per household basis
- Average 52 blocks per household
- Average Block Moves per Block is 1.21
- $\operatorname{Min}=1, \mathrm{Max}=2.82$ (for all types)



## Differences by Interview

 Type- Less movement in Type 2 (Plot only, $\mathrm{P}<.05$ )
- Type 3 (Enterprise only) more movement than Type 1/2 ( P < 0.05), trending toward more than Type 4

Number of Block Moves Per Block by Interview Type

## Interview Length \& Number Block Moves Per Block

- Interview length increases with increasing movement between blocks
- IW Length $=80.6$ * MovesPerBlock $+216$
- $P=.028 ; R^{2}=0.302 ; \operatorname{Adj} R^{2}=0.255$
- Non-linear relationship?


Average Moves Per Block \& Interview Length


## Interview Length \& Number of Block Moves Per Section

Number interviews ranges from $\mathrm{N}=1277$ to 2
Interviews are clustered under 1.5 Block Moves Per Section

## Interview Length \& Number Block Moves Per Block By Type




Type 4) Personal + Household + Enterprise
+ Plot
IW Length $=190$ * MovesPerBlock +144
P < .0001, Adj $\mathbf{R}^{2}=.70$

## Some Observations (Block Moves)

$>$ Movement between blocks are with a cost
$>$ Interview length increases with increasing movement between blocks
> Some movements are explainable with the instrument design but others are unsure --- why

## What else we look into?

$>$ Other paradata usage examples:
$>$ By block
$>$ By Iwer
$>$ By question
$>$ By respondent
$>$ By time/lw order

## Midblock Exits (by block)



## Midblock Exits (by block by person)

Midblock exits divided by total exits (ratio_midblk_exits_by_blk) for specific, ennumerated blocks Larger bubbles indicate more exit events.


## Midblock Exits - Overall (by Iwer)

Midblock exits divided by total exits (ratio_midblk_exits_by_blk) on a per-IWER basis Data represent IWERs with > 75 total exits

© 2016 by the Regents of the University of Michigan

## Midblock Exits (by Iwer) - Consumption

Midblock exits divided by total exits (ratio_midblk_exit_iwer_blk) per IWER for the s11_consumption block Data represent IWERs with > 20 total exits


## Midblock Exits (by Iwer) - Health

Midblock exits divided by total exits (ratio_midblk_exit_iwer_blk) per IWER for the s06_health[1] block Data represent IWERs with > 20 total exits


## Midblock Exits (by Iwer) - Housing

Midblock exits divided by total exits (ratio_midblk_exit_iwer_blk) per IWER for the s12_housing block Data represent IWERs with > $\mathbf{2 0}$ total exits


## Midblock Exits (by Iwer) - HH Roster

Midblock exits divided by total exits (ratio_midblk_exit_iwer_blk) per IWER for the s01b2_householdroster block Data represent IWERs with > 20 total exits


## Midblock Exits (by Iwer) - Non Resident Relatives

Midblock exits divided by total exits (ratio_midblk_exit_iwer_blk) per IWER for the s02a_nr_relatives block Data represent IWERs with > $\mathbf{2 0}$ total exits


## QC - Quick Read

 second by the Iw Order


## QC - Response Changes

Response changes divided by total responses (ratio_resp_chg_by_blk) for specific blocks Data represent blocks with > 100 responses and a response change rate >=3\%.

© 2016 by the Regents of the University of Michigan

Average \# of Response Changes by the Iw Order


## Next Steps

$>$ Link the block movement data with interviewer level.
$>$ Identify "sensitive" QC indicators (e.g., data discrepancies).
$>$ Who answered the Iw.
$>$ Interviewers debriefing.

- Apply all the lessons we learned in this wave to next wave instrument design


## Acknowledgements

$>$ Contributors from Yale University and Institute for Statistical, Social and Economic Research (ISSER) at the University of Ghana
$>$ University of Michigan
> Project staff: Beth-Ellen Pennell, Gina-Qian Cheung, Yu-Chieh (Jay) Lin, Lisa Wood, Joel Devonshire, Jennie Williams, Xuetao (Brant) Zhang, Kyle S. Kwaiser
$>$ Analysis funding: U-M Survey Research Operations

# Thank You! 

Contact info:<br>yuchiehl@umich.edu

