KANTAR

Panel Hacking

Jon Puleston

Private & Confidential

How do fraudsters hack the system?

1. **Fake panellists:** Respondents who are not who they say they to join panels e.g. someone from China pretending to be from USA
2. **Ghost respondents:** Hackers over-riding the transfer link process
3. **Screener hackers:**
   - **Professional respondents:** Skilled at getting through screeners
   - **Chatroom hackers:** “this is how you get through the screener”
4. **Multi-completors:** People completing surveys multiple times on multiple devices (this could be individuals or a more organised process)
5. **Bots:** Automated survey completion processes

Dealing with fake panellists: how we weed them out

Email address validation procedures:

- Verity (US) = score 1-5
- Global Z (global) = Groups A-J

IP proxy testing:

- MaxMind (used to detect credit card fraud)

Same device activity:

- Kantar Digital Fingerprint testing

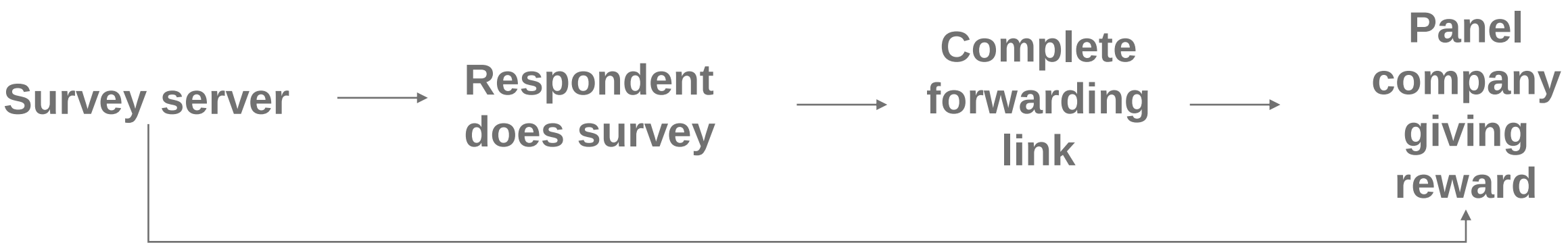
Machine learning:

- Patterns of behaviour across devices, browser types, operating systems and IP addresses e.g. identified certain older operating systems more likely to be fraudulent

Ghost respondents

Much bigger industry-wide challenge right now

Ghost respondents



Problem: we can only spot ghosts when we tally up completes between systems…

- Much more rife with sample only projects where its not so easy to cross compare
- Requires an industry co-ordinated initiative to solve

Right now panel companies are taking the hit: But these costs are getting indirectly passed back to clients

Ghost respondents: solutions?

Server to server dual validation

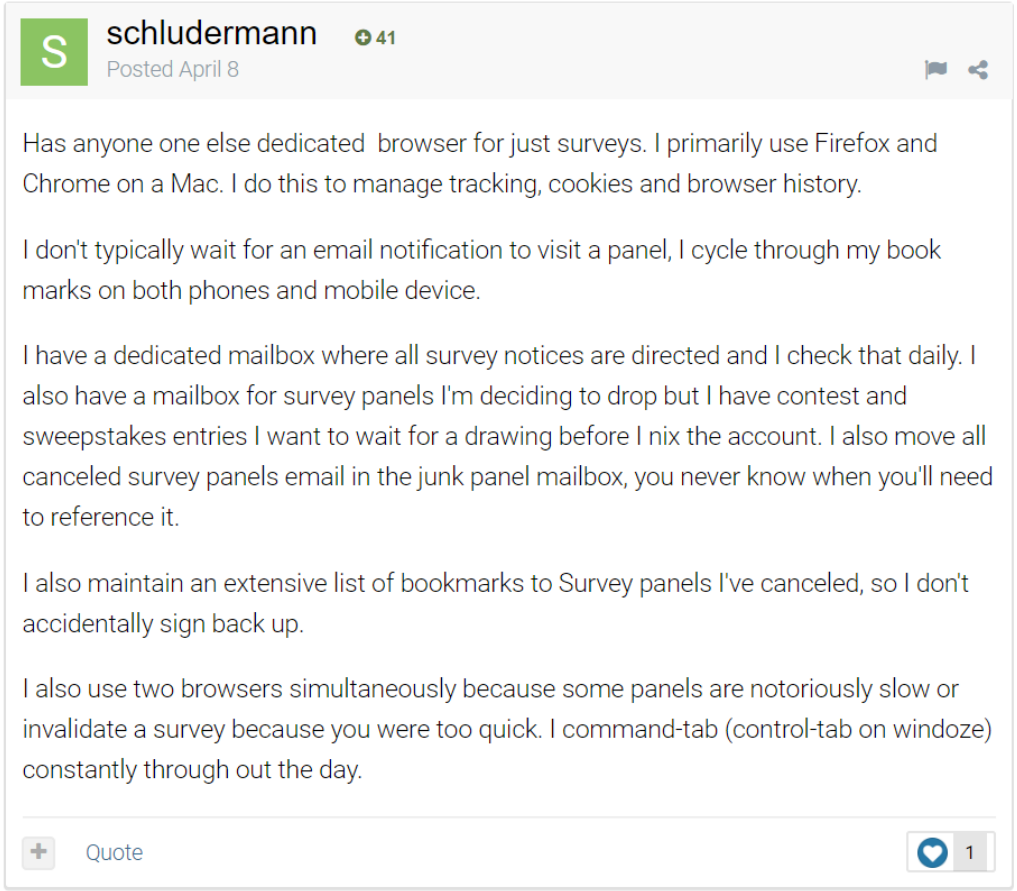
Survey server → Respondent does survey → Complete forwarding link → Panel company giving reward

Encryption

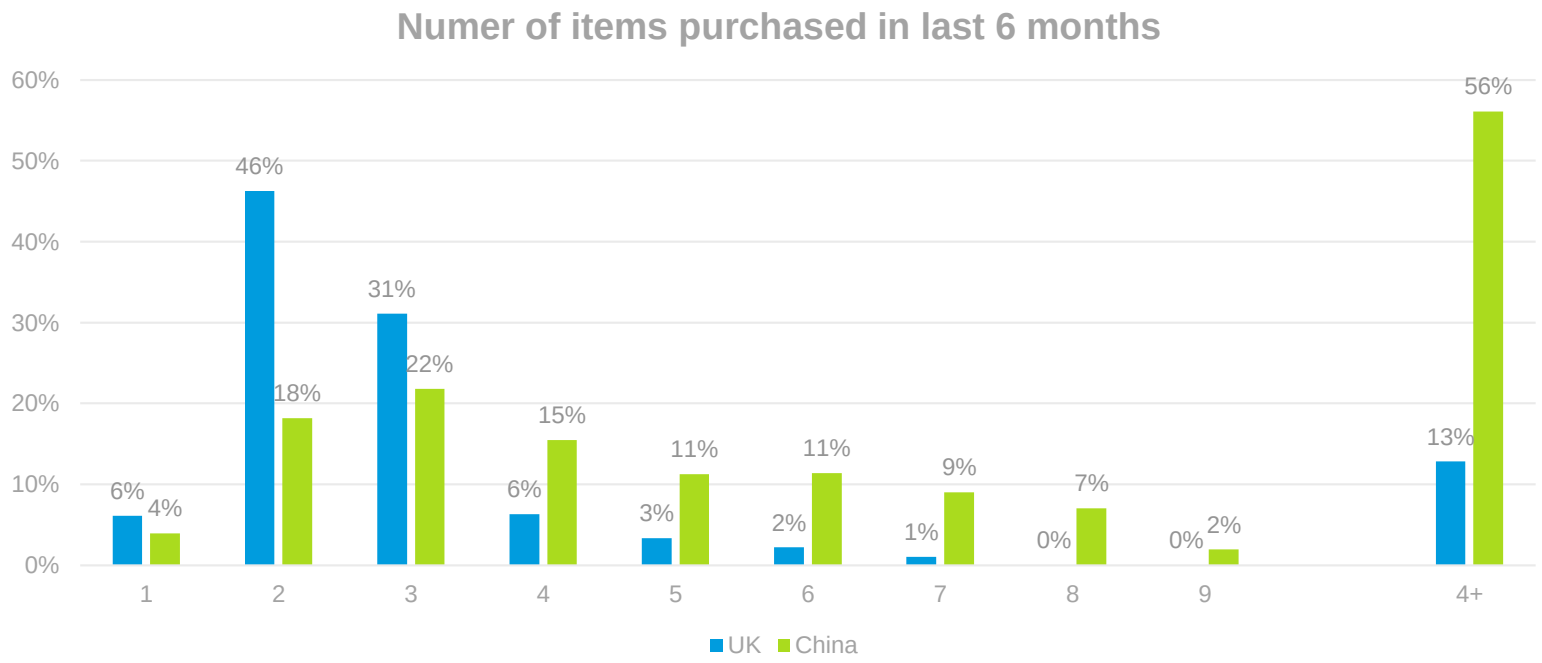
Moving battle: Shoals of ghost respondents moving from one country to the next, one survey system to the next

Screener hacking

Amateurs → Pro



How to spot a screener hack…



Kantar Honesty detection

Ask a series of low incidence rate questions

Q1 = 1
Q2 = 0
Q3 = 1
Q4 = 0
Q5 = 1
Q6 = 0
Q7 = 1
Q8 = 0
Q9 = 1
Q10 = 0

If sum Q1-Q10 > X

Potential Hacker

The trouble is hackers are getting wise to this…

Another ways – looking for formula answers

Q1 = 1
Q2 = 0
Q3 = 1
Q4 = 6
Q5 = 1
Q6 = 0

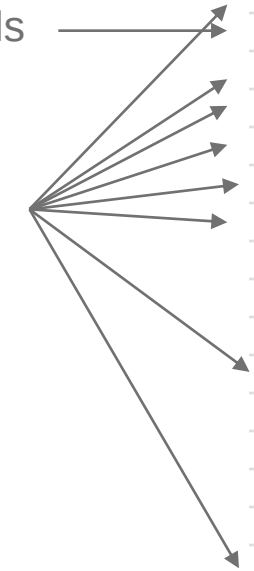
Covert to number

1010610 = 2+

Potential hack

Random Individuals

Chatroom hack: Don't score 9 for 3rd option

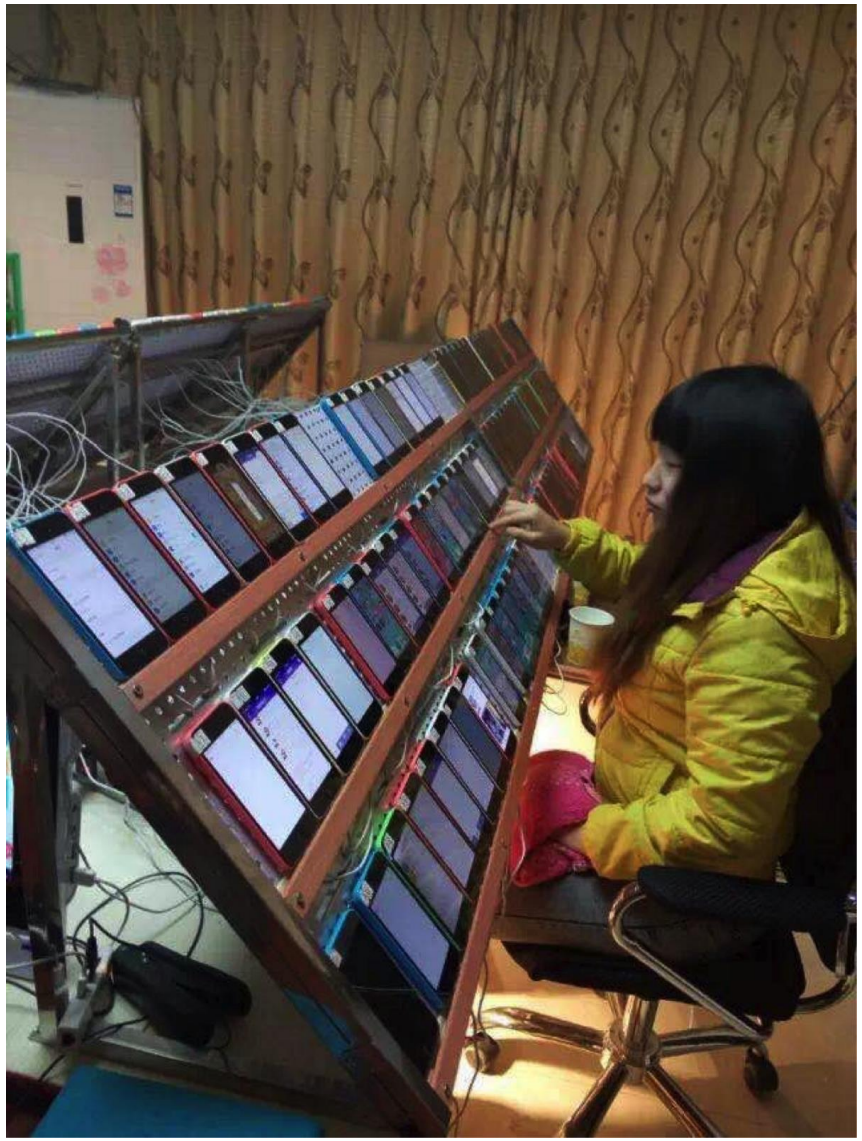


China Suspicious answer paterns	Number			
998999999999999	99	Suspicious	698	7%
999999999999999	91	total	**10153**	
776777777777777	80			
997999999999999	69			
886888888888888	66			
887888888888888	40			
775777777777777	28			
888888888888888	19			
777777777777777	19			
666666666666666	9			
665666666666666	7			
777777777777	6			
111111111198999	6			
111111111198899	5			
999999999999	4			
776777777777788	4			
99899999999	3			
999999899999999	3			
999989999999999	3			
998777799999999	3			
998888899999999	3			
997888888888888	3			
887888888899988	3			
887999999999999	3			
887888877788888	3			
776888877788877	3			
999888888888899	2			
99899999	2			
998999988888899	2			
996777777777777	2			
998999999999	2			
997788899999999	2			
999899999999999	2			
997989899999999	2			

Difficult to write code to dynamically spot this activity in a live survey

Multi-completion processes:

Double registrations → Click farms



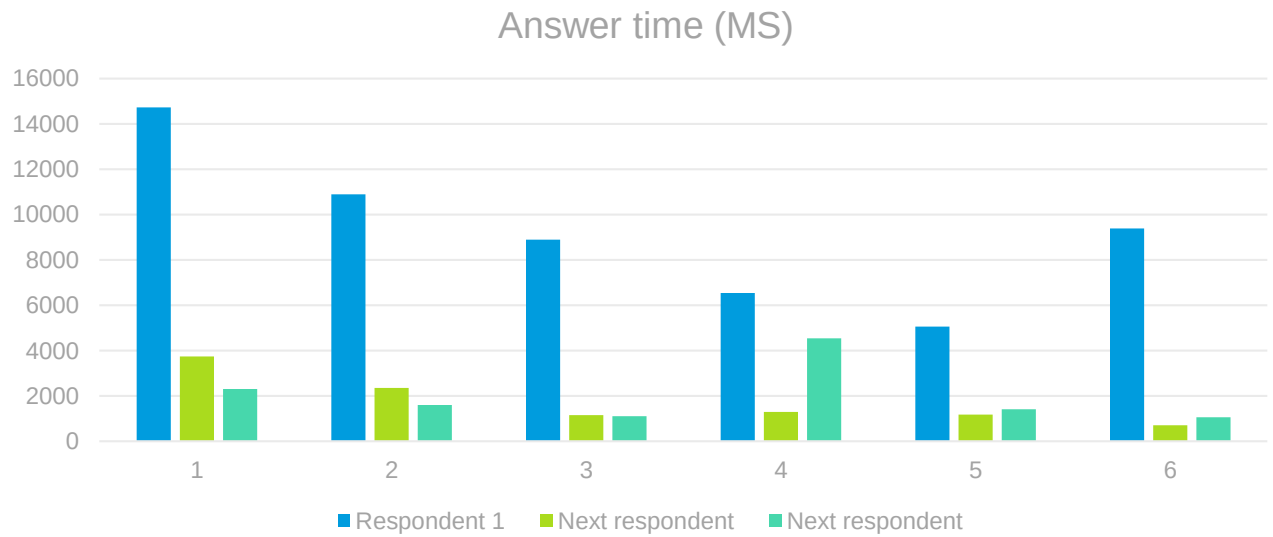
Traditional methods of spotting

IP + computer fingerprint matching→ a room full of devices each with unique id

Screening out speeder → now looking for the opposite – regular longer gaps between answering questions an indicator

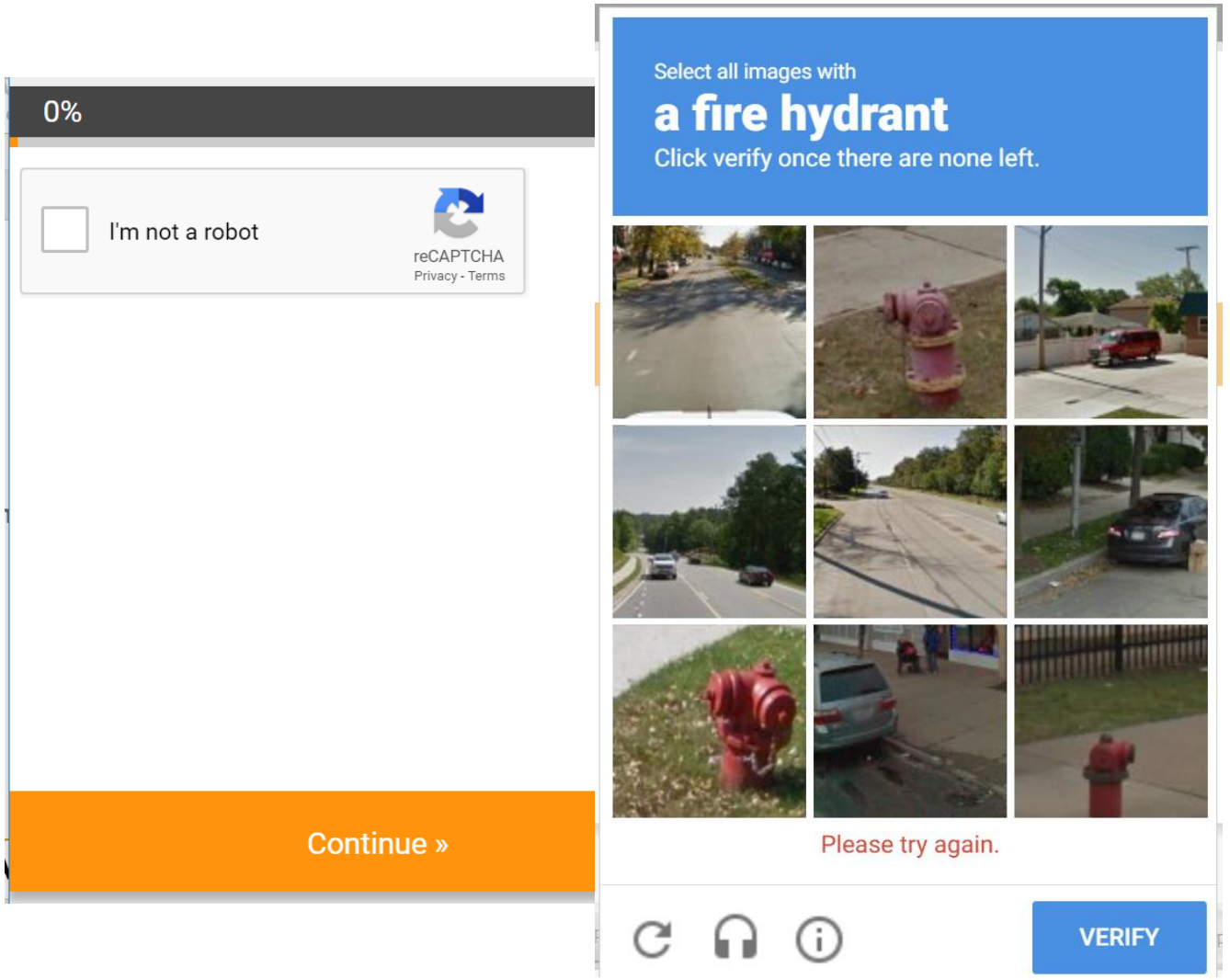
New methods

1. **Looking for cluster of similar answers:** within the same time frame respondents giving similar (not identical) answers
2. **Analysing open ended answers:** looking for similar answers with the same time frame
3. **Consistent answer time gaps:** the time it takes to answer each question across a survey
4. **Long, short, short, short patterns:** the time it takes to answer each question across a survey



Exploring machine learning techniques to spot these types of patterns

Bots and automated responses



The future…

Algorithmic consistency checks: bespoke designed or powered by machine learning

More sophisticated answer time based algorithms: time gap analysis

Using verbatim question analysis: One of the easiest way to spot hacking activity

Fraud screening moving to the end of the survey: to make it less clear where it has been spotted making it much harder to hack the process

Honest respondent validation: hard to spot a hacker, easy to pick out real & honest respondents

Differentating between fraud & satisficing

What gets commonly overclaimed

Purchase activity - ownership of prestige consumer items

Education levels

Car ownership

Baby related questions

Premium media consumption activity e.g. newspaper readership

What gets commonly under claimed

Smoking

Finance related stuff

Medical related

Any activity that is not socially desirable

Watching reality TV

Fraud or bored?

Case study

Differentiating between fraud; satisficing, bored respondents, and simply badly asked questions delivering odd results

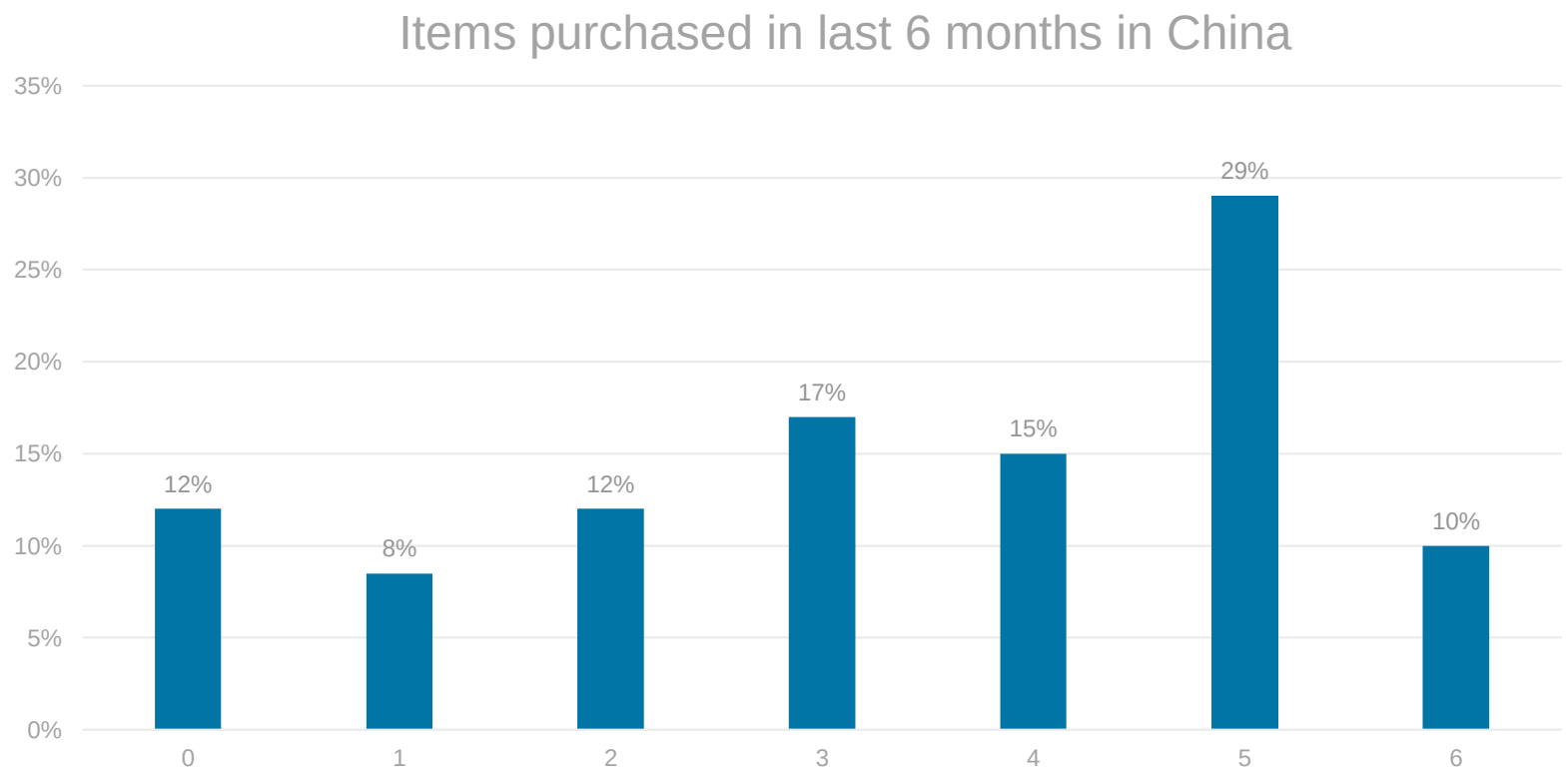
China study data: 10,000 responses

	0.6										
	Spearman	12	8	0	0	12	5	5	0	12	8
34	**Lund, Alexander (LS):** need to manually change isf not 0.6	Purchase Frequency | Bought a new car	Purchase Frequency | Bought a second hand car	Purchase Frequency | Bought fuel	Purchase Frequency | Bought motor oil	Purchase Frequency | Took out a mortgage	Purchase Frequency | Opened a bank account	Purchase Frequency | Took out an insurance policy	Purchase Frequency | Took out an investment / brokerage	Purchase Frequency | Took out a personal loan	Purchase Frequency | Signed-up for Credit card
0.562	Purchase Frequency | Bought a new car	1.000	0.695	0.244	0.592	0.664	0.591	0.610	0.513	0.649	0.639
	Purchase Frequency | Bought a second hand car	0.695	1.000	0.020	0.459	0.705	0.582	0.527	0.481	0.661	0.600
	Purchase Frequency | Bought fuel	0.244	0.020	1.000	0.499	0.083	0.119	0.281	0.213	0.109	0.154
	Purchase Frequency | Bought motor oil	0.592	0.459	0.499	1.000	0.465	0.450	0.515	0.438	0.463	0.491
	Purchase Frequency | Took out a mortgage	0.664	0.705	0.083	0.465	1.000	0.645	0.623	0.584	0.805	0.702
	Purchase Frequency | Opened a bank account	0.591	0.582	0.119	0.450	0.645	1.000	0.602	0.551	0.648	0.709
	Purchase Frequency | Took out an insurance policy	0.610	0.527	0.281	0.515	0.623	0.602	1.000	0.593	0.617	0.633
	Purchase Frequency | Took out an investment / brokerage	0.513	0.481	0.213	0.438	0.584	0.551	0.593	1.000	0.589	0.592
	Purchase Frequency | Took out a personal loan	0.649	0.661	0.109	0.463	0.805	0.648	0.617	0.589	1.000	0.706
	Purchase Frequency | Signed-up for Credit card	0.639	0.600	0.154	0.491	0.702	0.709	0.633	0.592	0.706	1.000

34% purchase questions correlate above 0.6

299 pairs of twins

High proportion of overclaiming



“Real Lives” honesty priming

我早餐吃什么

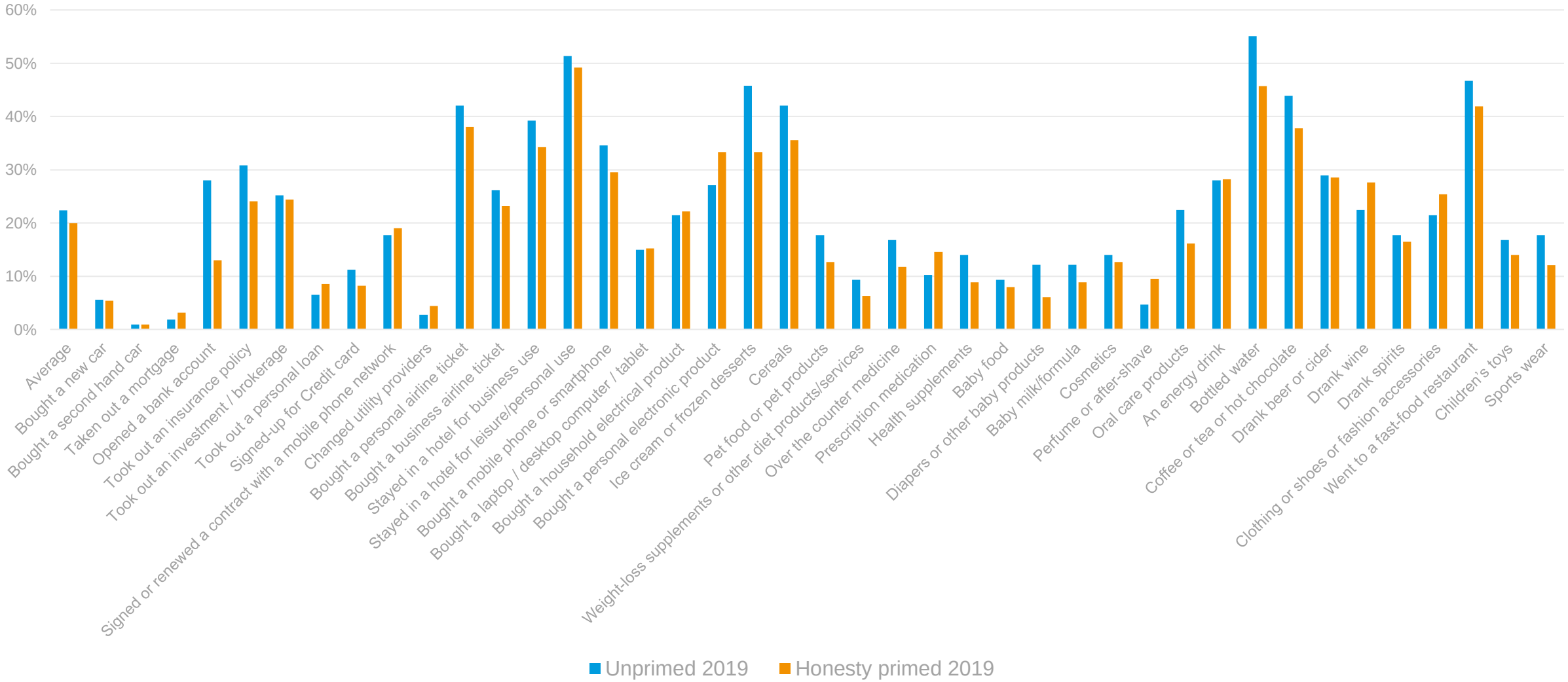
我上班采用什么出行方式

我下班后做什么

我的上一次旅行

但在本次调研中，我们关注的是人们的真实生活。

Done recently: China



Panel hacking

Mostly our problem right now – but we are all paying for it one way or another

New procedures needed to spot hackers

Co-ordinated response needed

Right now if you can’t spot a fraudulent respondent in your data more fool you

Honesty priming can help to significantly reduce the overall scale of overclaim making it easier to identify and deal with real fraudulent behaviour

Questions for you…

1. What can we do to tackle ghosting more effectively?
2. Does any one have any smart ideas on how to better pick up on fraudulent activity on the fly in live survey?
3. How to we square of the fact that in doing this we reduce panel supply and so it will push up cost for those that implement it?