**Online Appendix A: Question Wordings**

*BSA (2015)*

Do you have a personal Twitter account?

Yes

No

**IF Yes**

We are interested in being able to link people's answers to this survey to the ways in which they use Twitter. We would also like to know who uses Twitter. A research project about who and how people use Twitter is being conducted by a team of researchers at Cardiff University. Are you willing to tell me the name of your personal Twitter account and for this to be passed to researchers at Cardiff University, along with your answers to this survey? Your Twitter name would not be published.

Yes

No

**IF Yes**

INTERVIEWER: Please enter the respondent's Twitter name here

Open Question (Maximum of 100 characters)

*NatCen Panel (July 2017)*

Do you have a personal Twitter account?

Yes

No

**IF Yes**

As social media plays an increasing role in society, we would like to know who uses Twitter, and how people use it. We are also interested in being able to add people’s, and specifically, your answers to this survey to publicly available information from your Twitter account such as your profile information,  tweets in the past and in future, and information about how you use your account. Your Twitter information will be treated as confidential and given the same protections as your interview data. Your Twitter username, and any information that would allow you to be identified, will not be published without your explicit permission.

Are you willing to tell me your personal Twitter username and for your Twitter information to be added to your answers to this survey?

Yes

No

HELP SCREENS AVAILABLE

HELP SCREEN: What information will you collect from my Twitter account?

We will only collect information from your Twitter account that is publicly available. This will include information from your account (such as your profile description, who you follow, and who follows you), the content of your tweets (including text, images, videos and web links), and background information about your tweets (such as when you tweeted, what type of device you tweeted from, and the location the tweet was sent from).We will collect information from your past tweets (up to the last 3,000) and will update this with information from more recent tweets on a regular basis.

HELP SCREEN: What will the information be used for?

The information will be used for social research purposes only. Adding your Twitter information and your survey answers will allow researchers from universities, charities and government to better understand your experiences and opinions. For example, using extra information from your Twitter account, researchers can start to:

•             Understand who uses Twitter and how they use it

•             See what Twitter information can tell us about people, and how accurate it is

•             Know what people in the UK are saying about things we don’t ask in our survey

•             Look at additional information related to questions asked in the survey

HELP SCREEN: Who will be able to access the information?

Matched data which includes both your survey answers and Twitter information will be made available for social research purposes only. Researchers who want to use your matched Twitter and survey information must apply to access it and present a strong scientific case to ensure that the information is used responsibly and safely. Matched statistical information from your Twitter account which you cannot be identified from (e.g. how often you Tweet, or whether you follow any politicians) will have the same access controls as your other survey answers. At no point will any information that would allow you to be identified be made available to the public

HELP SCREEN: What will you do to keep my information safe?

All information we collect will be held in accordance with the Data Protection Act 1998. Because Twitter information is public data that anyone can search, it is impossible to anonymise completely. To keep your information safe, researchers will only be able to access the matched survey answers and detailed Twitter information in a secure environment set up to protect this type of data. Only approved researchers who have gone through special training may access this information, and they will have to apply to do so. Statistical information from your Twitter account which you cannot be identified from (e.g. how often you Tweet, or whether you follow any politicians) will have the same level of protection as your other survey answers.

HELP SCREEN: What if I change my mind?

This information will be collected and stored for as long as they are useful for research purposes, or until you contact us to withdraw your permission. You can do this at any time by emailing us at [panel@natcen.ac.uk](mailto:panel@natcen.ac.uk) or calling 0800 652 4569, and do not have to give a reason.

{END OF HELP SCREENS}

**IF Yes**

What is your Twitter username?

*SOFTCHECK:* "Twitter usernames must begin with an @ character, followed a maximum of 15 characters (A-Z, a-z, 0-9, underscore), no word spaces. Please check and amend.”

*IP10 (2017)*

Do you have a personal Twitter account?

Yes

No

**IF Yes**

We would like to know who uses Twitter, and how people use it. We are also interested in being able to add people's answers to this survey to publically available information from your Twitter account such as your profile information, tweet content, and information about how you use your account. Your Twitter information will be treated as confidential and given the same protections as your interview data. Your Twitter username, and any information that would allow you to be identified, will not be published without your explicit permission. Are you willing to tell me the name of your personal Twitter account and for your Twitter information to be linked with your answers to this survey?

Yes

No

HELP SCREENS AVAILABLE

HELP SCREEN: What information will you collect from my Twitter account?   
We will only collect information from your Twitter account that is publically available. This will include information from your account (such as your profile description, who you follow, and who follows you), the content of your tweets (including text, images, videos and web links), and background information about your tweets (such as when you tweeted, what type of device you tweeted from, and the location the tweet was sent from). We will collect information from your past tweets (up to the last 3,000) and will update this with information from more recent tweets on a regular basis. This information will be collected and stored for as long as they are useful for research purposes, or until you contact us to withdraw your permission. You can do this at any time, and do not have to give a reason.

HELP SCREEN: What will the information be used for?   
The information will be used for social research purposes only. Adding your Twitter information and your survey answers will allow researchers from universities, charities and government to better understand your experiences and opinions. For example, using extra information from your Twitter account, researchers can start to:   
\* Understand who uses Twitter and how they use it   
\* See what Twitter information can tell us about people, and how accurate it is   
\* Know what people in the UK are saying about things we don't ask in our survey   
\* Look at additional information related to questions asked in the survey

HELP SCREEN: Who will be able to access the information?   
Researchers who want to use matched Twitter and survey information must apply to access it and present a strong scientific case to ensure that the information is used responsibly and safely. Matched statistical information from your Twitter account which you cannot be identified from (e.g. how often you Tweet, or whether you follow any politicians) will have the same access controls as your other survey answers.

HELP SCREEN: What will you do to keep my information safe?   
Matched statistical information from your Twitter account which you cannot be identified from (e.g. how often you Tweet, or whether you follow any politicians) will have the same level of protection as your other survey answers.

**IF Yes**

What is your Twitter username (e.g. @usociety)?

*Soft check: Twitter username does not being with '@' or contains spaces* **“**Please check and amend. Twitter usernames should begin with an @ character and should not contain any spaces.”

**Online Appendix B: SAS Code Used**

libname ip "D:\IP\";

libname twit "D:\twitter\";

\*IP10 Data;

**data** j1;

set ip.j\_indresp\_ip;

x = UNIFORM(**123456**);

\*randomly assigning 3 missing cases to urban/rural;

if j\_urban\_dv = **.** and x>=**.5** then urban = **1**;

else if j\_urban\_dv = **.** and x<**.5** then urban = **0**;

else if j\_urban\_dv = **1** then urban = **1**;

else if j\_urban\_dv = **2** then urban =**0**;

else urban = **.**;

if j\_hiqual\_dv = **1** then colprof = **1**;

else if j\_hiqual\_dv = -**9** then colprof = **.**;

else colprof = **0**;

if j\_hiqual\_dv = **2** or j\_hiqual\_dv = **3** then profA = **1**;

else if j\_hiqual\_dv = -**9** then profA = **.**;

else profA = **0**;

if j\_hiqual\_dv >**3** then otheduc = **1**;

else if j\_hiqual\_dv = -**9** then otheduc = **.**;

else otheduc = **0**;

if j\_indmode = **3** then web= **1**;

else web = **0**;

if j\_sex = **1** then female = **0**;

else if j\_sex =**2** then female = **1**;

if j\_netpuse = **1** then net =**1**;

else if j\_netpuse>**1** then net =**0**;

else net =**.**;

if j\_twithave <**0** then twithave =**.**;

else if j\_twithave = **2** then twithave = **0**;

else twithave =**1**;

if j\_twitlink in (-**1** -**2** **2**) then twitlink = **0**;

else if j\_twitlink = **1** then twitlink =**1**;

else twitlink =**.**;

if j\_pdvage <**0** then age =**.**;

else age = j\_pdvage;

if j\_jbsemp in (-**1** -**2** **1** **2**) then emp=**1**;

else emp =**0**;

**run**;

\*IP10 Data to match BSA IVs;

**data** j2;

set ip.j\_hhresp\_ip;

if j\_fihhmngrs\_dv <=**1300** then hhinc=**1**;

else if **1301**<=j\_fihhmngrs\_dv<=**2400** then hhinc =**2**;

else if **2400.000000000000001**<=j\_fihhmngrs\_dv<=**4200** then hhinc =**3**;

else if **4201**<=j\_fihhmngrs\_dv then hhinc =**4**;

keep j\_fihhmngrs\_dv j\_hidp hhinc;

**run**;

**proc** **sort** data = j1;

by j\_hidp;

**run**;

**proc** **sort** data = j2;

by j\_hidp;

**run**;

**data** j;

merge j1 j2;

by j\_hidp;

**run**;

\*Merging all IP waves for feed forward data;

**data** a;

set ip.a\_indresp\_ip;

keep pidp a\_remail\_code;

**run**;

**data** b;

set ip.b\_indresp\_ip;

keep pidp b\_remail\_code;

**run**;

**data** c;

set ip.c\_indresp\_ip;

keep pidp c\_remail\_code;

**run**;

**data** d;

set ip.d\_indresp\_ip;

keep pidp d\_remail\_code;

**run**;

**data** e;

set ip.e\_indresp\_ip;

keep pidp e\_remail\_code;

**run**;

**data** f;

set ip.f\_indresp\_ip;

keep pidp f\_remail\_code;

**run**;

**data** g;

set ip.g\_indresp\_ip;

keep pidp g\_remail\_code;

**run**;

**data** h;

set ip.h\_indresp\_ip;

keep pidp h\_remail\_code;

**run**;

**data** i;

set ip.i\_indresp\_ip;

keep pidp i\_remail\_code ;

**run**;

**proc** **sort** data =a;

by pidp;

**run**;

**proc** **sort** data =b;

by pidp;

**run**;

**proc** **sort** data =c;

by pidp;

**run**;

**proc** **sort** data =d;

by pidp;

**run**;

**proc** **sort** data =e;

by pidp;

**run**;

**proc** **sort** data =f;

by pidp;

**run**;

**proc** **sort** data =g;

by pidp;

**run**;

**proc** **sort** data =i;

by pidp;

**run**;

**proc** **sort** data =j;

by pidp;

**run**;

**data** aj;

merge a b c d e f g h i j;

by pidp;

**run**;

**data** aj;

set aj;

if a\_remail\_code =**1** or b\_remail\_code =**1** or c\_remail\_code =**1** or d\_remail\_code =**1** or e\_remail\_code =**1** or f\_remail\_code =**1** or g\_remail\_code =**1** or h\_remail\_code =**1** or i\_remail\_code =**1** or j\_remail\_code =**1**

then email =**1**;

else email =**0**;

**run**;

**data** twit.twit;

set aj;

where j\_hidp ne **.**;

**run**;

**data** twit.twit;

set twit.twit;

if colprof = **1** then educ1 = **1**;

else if profa =**1** then educ1 =**2**;

else if otheduc = **1** then educ1=**3**;

else educ1 =**.**;

**run**;

\*NatCen Panel data;

**data** twit.ncp1;

set twit.ncp;

if Cur\_HHIncome <**0** then hhinc = **.**;

else if **1**<=Cur\_HHIncome<=**6** then hhinc = **1**;

else if **7**<=Cur\_HHIncome<=**11** then hhinc = **2**;

else if **12**<=Cur\_HHIncome<=**16** then hhinc = **3**;

else if **17**<=Cur\_HHIncome<=**20** then hhinc = **4**;

if Cur\_HEdQual <**0** then colprof = **.**;

else if Cur\_HEdQual =**1** then colprof =**1**;

else colprof = **0**;

if Cur\_HEdQual <**0** then profA = **.**;

else if Cur\_HEdQual =**2** then profA =**1**;

else profA = **0**;

if profA = **1** or colprof =**1** then otheduc = **0**;

else if profA=**0** and colprof=**0** then otheduc = **1**;

else otheduc = **.**;

if Cur\_Age<**0** then age =**.**;

else age = Cur\_Age;

if FF\_Sex <**0** then female =**.**;

else if FF\_Sex =**2** then female =**1**;

else female =**0**;

if Cur\_EconAct <**0** then emp = **.**;

else if Cur\_EconAct in (**3** **4**) then emp = **1**;

else emp=**0**;

if July17\_TwitHas <**0** then twithave = **.**;

else if July17\_TwitHas = **2** then twithave = **0**;

else twithave = July17\_TwitHas;

if July17\_TwitConsent in (-**8** -**9** **2**) then twitlink =**0**;

else if July17\_TwitConsent =**1** then twitlink =**1**;

else twitlink =**.**;

if July17\_IntMode ="Web" then web =**1**;

else web =**0**;

if TwittN <**0** then twitlink15 =**.**;

else if TwittN =**2** then twitlink15 =**0**;

else twitlink15 = TwittN;

if Twitt in (-**1** -**2** **8** **9**) then twithave15 =**.**;

else if twitt = **2** then twithave15 =**0**;

else twithave15 = twitt;

**run**;

**proc** **freq** data = twit.twit;

tables web;

**run**;

**proc** **freq** data = twit.ncp1;

tables web twithave15 twitlink15 twithave15\*twithave twitlink\*twitlink15;

**run**;

**proc** **freq** data = ncp1;

table Cur\_EconAct;

**run**;

**data** twit.ncp1;

set twit.ncp1;

if colprof = **1** then educ1 = **1**;

else if profa =**1** then educ1 =**2**;

else if otheduc = **1** then educ1=**3**;

else educ1 =**.**;

if twithave15 = **1** and twithave =**0** then haveboth =-**1**;

else if twithave15 = **0** and twithave =**1** then haveboth =**1**;

else if twithave15 = **1** and twithave =**1** then haveboth =**2**;

else if twithave15 = **0** and twithave =**0** then haveboth =**0**;

else haveboth =**.**;

if twitlink15 = **1** and twitlink =**0** then linkboth =-**1**;

else if twitlink15 = **0** and twitlink =**1** then linkboth =**1**;

else if twitlink15 = **1** and twitlink =**1** then linkboth =**2**;

else if twitlink15 = **0** and twitlink =**0** then linkboth =**0**;

else linkboth =**.**;

if linkboth in (**0** **2**) then linkchange = **0**;

else if linkboth in (-**1** **1**) then linkchange = **1**;

else linkchange = **.**;

**run**;

\*\*TABLE 1;

**proc** **freq** data = twit.twit;

tables twitlink twithave;

where j\_hhorig ne **17**;

**run**;

**proc** **freq** data = twit.twit;

table twithave\*web/ chisq;

where j\_hhorig ne **17**;

**run**;

**proc** **freq** data = twit.ncp1;

table twithave\*web/ chisq;

**run**;

\*Table 2;

**proc** **freq** data = twit.twit;

table twitlink\*(web female hhinc2 emp educ1 urban net email j\_hhorig)/ chisq;

where j\_hhorig ne **17**;

**run**;

**proc** **freq** data = twit.ncp1;

table twitlink\*(web female hhinc emp educ1)/ chisq;

**run**;

**proc** **ttest** data = twit.ncp1;

class twitlink;

var age;

**run**;

**proc** **ttest** data = twit.twit;

class twitlink;

var age;

where j\_hhorig ne **17**;

**run**;

**proc** **freq** data = twit.twit;

tables web female hhinc emp colprof profa age urban net email j\_hhorig;

where twitlink ne **.**;

**run**;

\*Table 3;

**proc** **logistic** data = twit.twit;

class j\_hhorig hhinc2;

model twitlink (event ='1') = web female hhinc2 emp colprof profa age urban net email j\_hhorig;

where j\_hhorig ne **17**;

**run**;

**proc** **logistic** data = twit.ncp1;

class hhinc;

model twitlink (event ='1') =web female hhinc emp colprof profa age;

**run**;

\*Table 4;

**proc** **freq** data = twit.ncp1;

tables haveboth linkboth linkchange;

**run**;

\*Table 4;

**proc** **freq** data = twit.ncp1;

tables (haveboth linkboth linkchange)\*web/chisq;

**run**;

**proc** **freq** data = twit.ncp1;

tables linkchange\*web/fisher;

**run**;

**proc** **freq** data = twit.ncp1;

tables web\*twitlink;

where twitlink15 =**1**;

**run**;