Appendix 1

Grigor Dimitrov

September 27, 2016

Panel

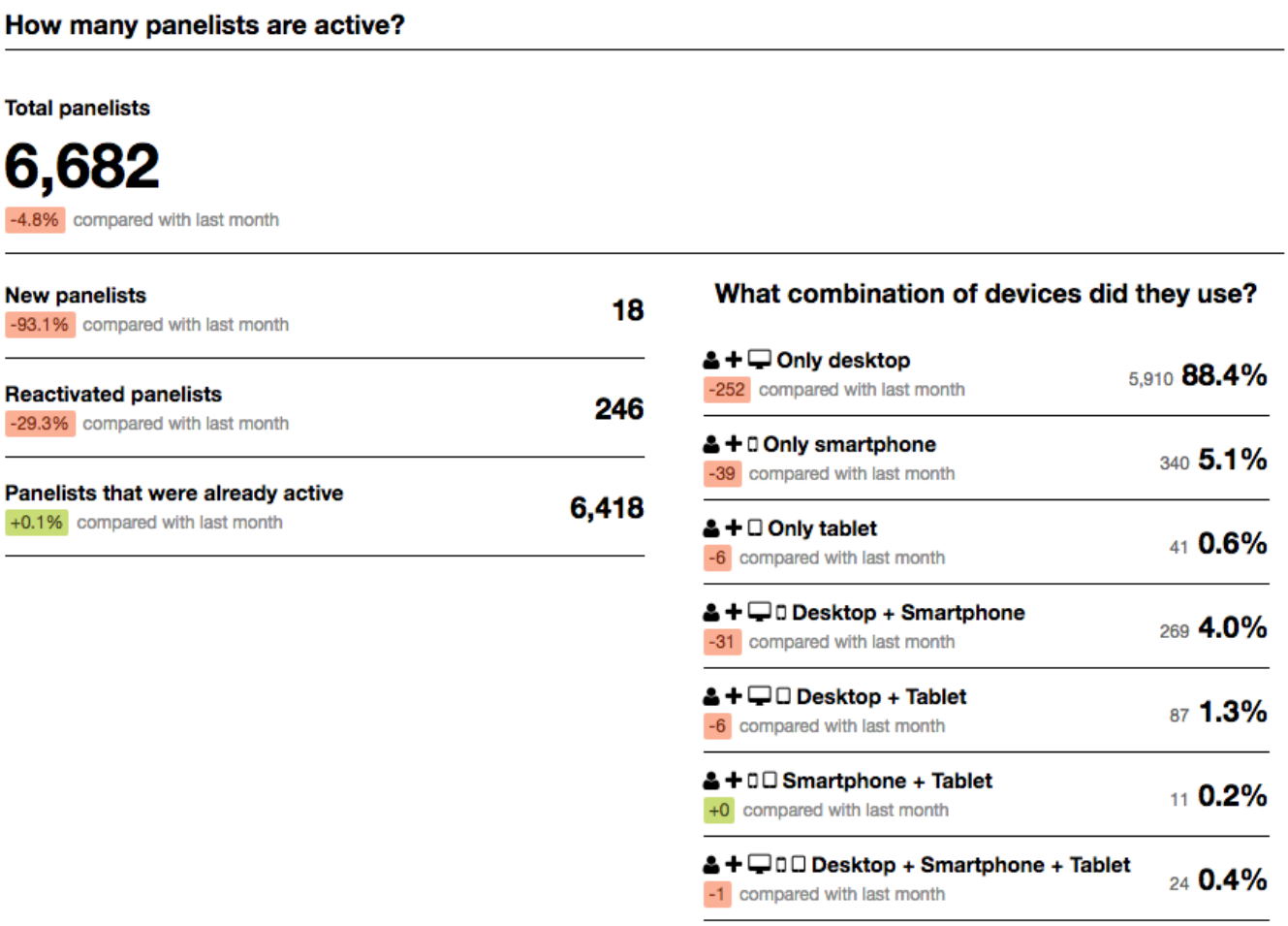

Figure 1: Panelists

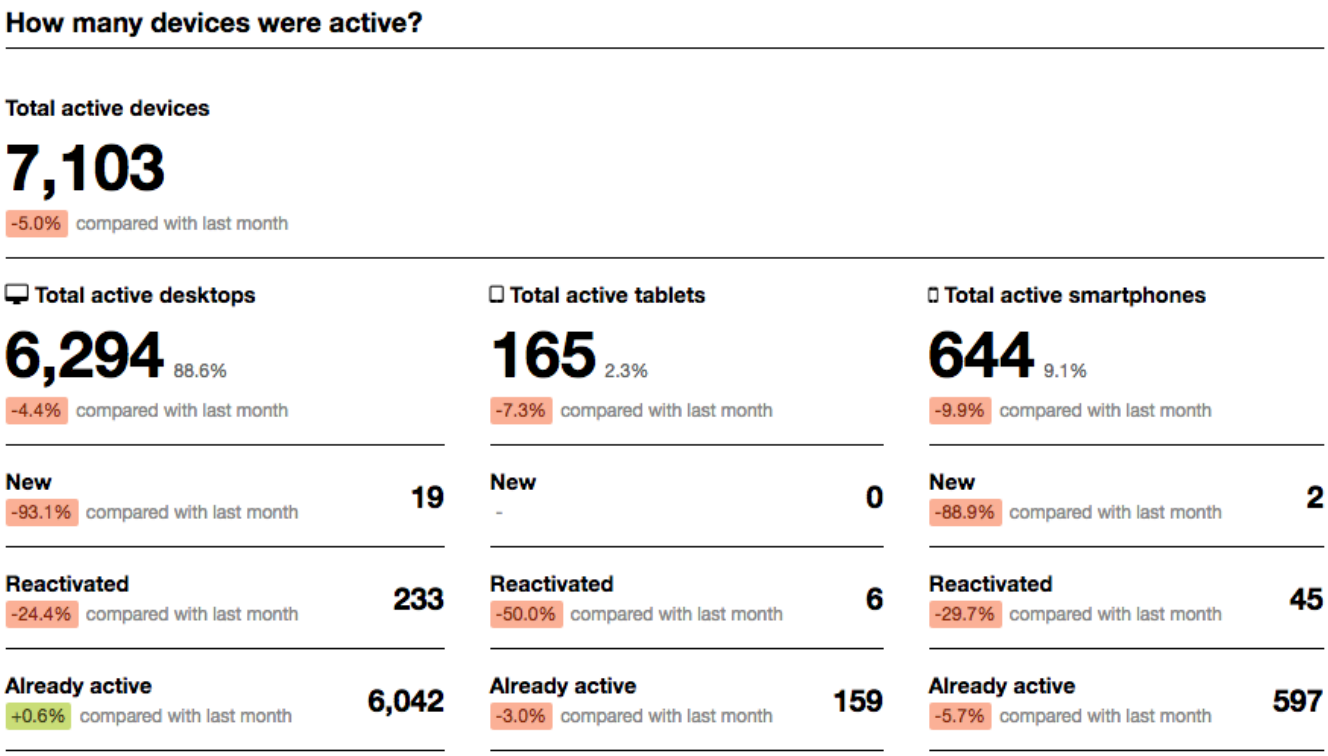

Figure 2: Devices

Data and varables exaple

```
##                used_at         host panelist_id          url
## 1  2016-01-01 00:47:06 telegraaf.nl        2112 telegraaf.nl
## 2  2016-01-01 00:48:30 telegraaf.nl        2112 telegraaf.nl
## 3  2016-01-01 00:51:28 telegraaf.nl        2112 telegraaf.nl
## 4  2016-01-01 00:53:29 telegraaf.nl        2112 telegraaf.nl
## 5  2016-01-01 01:28:52 telegraaf.nl        2112 telegraaf.nl
## 6  2016-01-01 02:01:00 telegraaf.nl        2112 telegraaf.nl
## 7  2016-01-01 02:01:12  twitter.com        2112  twitter.com
## 8  2016-01-01 02:04:07  twitter.com        2112  twitter.com
## 9  2016-01-01 02:05:07  twitter.com        2112  twitter.com
## 10 2016-01-01 02:27:31  twitter.com        2112  twitter.com
##    active_seconds browser_name mmid Class_Travel purchase
## 1              72       safari    1            0       NA
## 2             152       safari    1            0       NA
## 3               0       safari    1            0       NA
## 4               0       safari    1            0       NA
## 5              71       safari    2            0       NA
## 6              11       safari    3            0       NA
## 7             168       safari    3            0       NA
## 8              60       safari    3            0       NA
## 9               0       safari    3            0       NA
## 10             15       safari    4            0       NA
```

Figure 3: Desktop data

```
##                     app_name host panelist_id device_id scheme  url domain
## 1  ABN AMRO Mobiel Bankieren <NA>     1008505     10093   <NA> <NA>   <NA>
## 2  ABN AMRO Mobiel Bankieren <NA>     1008505     10093   <NA> <NA>   <NA>
## 3  ABN AMRO Mobiel Bankieren <NA>     1008505     10093   <NA> <NA>   <NA>
## 4  ABN AMRO Mobiel Bankieren <NA>     1008505     10093   <NA> <NA>   <NA>
## 5  ABN AMRO Mobiel Bankieren <NA>     1008505     10093   <NA> <NA>   <NA>
## 6  ABN AMRO Mobiel Bankieren <NA>     1008505     10093   <NA> <NA>   <NA>
## 7  ABN AMRO Mobiel Bankieren <NA>     1008505     10093   <NA> <NA>   <NA>
## 8  ABN AMRO Mobiel Bankieren <NA>     1008505     10093   <NA> <NA>   <NA>
## 9  ABN AMRO Mobiel Bankieren <NA>     1008505     10093   <NA> <NA>   <NA>
## 10 ABN AMRO Mobiel Bankieren <NA>     1008505     10093   <NA> <NA>   <NA>
##                                   app_id              used_at connection
## 1  cb46bcba-7258-4b47-8557-de3ff607b456 2016-02-22 13:43:15       wifi
## 2  cb46bcba-7258-4b47-8557-de3ff607b456 2016-02-03 18:03:34   cellular
## 3  cb46bcba-7258-4b47-8557-de3ff607b456 2016-01-23 15:38:49       wifi
## 4  cb46bcba-7258-4b47-8557-de3ff607b456 2016-01-16 15:11:21       wifi
## 5  cb46bcba-7258-4b47-8557-de3ff607b456 2016-02-03 18:43:44       wifi
## 6  cb46bcba-7258-4b47-8557-de3ff607b456 2016-01-08 14:15:29       wifi
## 7  cb46bcba-7258-4b47-8557-de3ff607b456 2016-01-24 10:20:45       wifi
## 8  cb46bcba-7258-4b47-8557-de3ff607b456 2016-02-03 17:00:56   cellular
## 9  cb46bcba-7258-4b47-8557-de3ff607b456 2016-01-22 21:34:28   cellular
## 10 cb46bcba-7258-4b47-8557-de3ff607b456 2016-01-06 07:24:50       wifi
##    duration mmid Class_Travel TravelApp
## 1        21  799            0         0
## 2        19  551            0         0
## 3        34  392            0         0
## 4        11  283            0         0
## 5        21  552            0         0
## 6        23  139            0         0
## 7        25  396            0         0
## 8         9  549            0         0
## 9        45  381            0         0
## 10       15   98            0         0
```

Figure 4: Mobile data

Analysis

The following section contains an explanation of the main techniques performed during the analysis along with their assumptions, followed by the results of the ordinary least squares diagnostic tests of the restricted model. Once the functional form of the restricted model has been selected I proceed with variable selection in order to come up with the final model. Finally, I ran the diagnostics of tests over the final model again.

Factor Analysis

Factor analysis is a widely used technique used for explaining the variance in several variables by a smaller set of latent variables. As in the current case, it is often used to consolidate several survey variables onto their "underlying" factors in order to reduce the dimensionality of the data. Factor analysis groups variables together, that is, using a lot of variables one can potentially reduce them to certain factors representing the latent underlying factors representing them by accounting the similar patterns in the variables. The intuition behind the analysis is as follows. The analysis groups together observed, correlated variables into smaller groups of unobserved (latent) variables (Yong and Pearce 2013).

In this case, I use factor analysis to to reduce the seven survey items regarding the risk and uncertainty attitude down to two constructs namely risk and uncertainty. Also to reduce the eleven-item scale of BIG5 to 5 factors representing each of on the of the five personality traits.

Regression Analysis

For testing the hypotheses of this paper, regression analysis will be utilized. The regression model or ordinary least squares (hereafter OLS) is the "cornerstone of econometrics" (Verbeek 2008). It aims at explaining a variable, y, in terms of another variable, x. In other words, using OLS researchers are able to find how will y vary as x changes, the ultimate goal being to infer the causal effect x has on y. Using such models allows to find relationships between various variables, present the effect the independent variables, xi has on the dependent variable, y in order to be able to make predictions.

The general linear regression models is represented as follows:

$$y = \beta_0 + X_1\beta_1 + X_2\beta_2 + ... + X_k\beta_k + \varepsilon$$

Where: y is the dependent variable

X_1 to X_k are the independent variables, which explain y

β_0 is the intercept, indicating the expected value of y when all the independent variables are equal to 0

β_1 to β_k are the coefficients which determine the effect x has on y

ε is the error term

Goodness of fit and model selection

The standard measures of fit include the R-squared and the adjusted R-squared, which measures the variance that is explained in the model for the independent variable by the dependent variables. The measure can be interpreted directly. For example, if the R-squared is equal to 0.45, it means that the variables included in the model explain 45% of the variation of the independent variable, y. The higher the value, the higher its predictive power. However, it should be noted that adjusted R-squared penalizes for the additional number of parameters. Thus, applying additional variables to the model, I should test if they are jointly significant in order to assess whether they are relevant or not in the model.This is typically applied by using the F-test (Wald test). Using both the R-squared, the overall F-test and applying the F-test to certain variables I can compare best which model fits the data best.

Akaike information criterion (AIC)

Model selection has been done over Akaike information criterion (AIC) introduced by Akaike (1974). AIC is a metric traditionally used for model selection. It compares the goodness of fit for a number of explanatory variables and penalizes for each additional explanatory variable.

BLUE Assumptions

There are several assumptions that need to be met when applying OLS explained in the section below. Namely Gauss-Markov assumptions for full ideal conditions for OLS. The model needs to be best linear unbiased estimator" ("BLUE") (Verbeek 2008). It is crucial for the assumptions to be met as to compute unbiased and consistent estimates that explain the variation in the dependent variable. Now, I will go through each assumption: Linear in parameters This implies that the model should have linear parameters, b, however, there can be nonlinearities in the variables, x. This assumption is met as my specified model does not include non-linearities in the parameters.

Normality

The error terms should follow a normal distribution. In large datasets, however, even if the error term does not follow a normal distribution the regression estimators are 'asymptotically normally distributed', meaning that following non-normal distribution is not crucial as the estimates will still be consistent and unbiased. The Shapiro-Wilk test can is adopted here and results presented below. The test works under null hypothesis: "the sample comes from normally distributed population" Shapiro and Wilk (1965)

Random sample

The data collection should be done randomly, meaning that the each subject should have the same probability of being selected. In this research, both in the behavioral and survey data collection parts, I can say that subjects were randomly selected for further analysis.

Multicollinearity

Multicollinearity implies that there is no perfect linear relationship between the independent (explanatory) variables as this can lead to 'unreliable regression estimated' (Verbeek 2008). For example, adding both male and female in the analysis would lead to perfect collinearity (as male + female =1) and the estimations would not work. In this example, removing one of the variables would solve the problem, however, there can be other variables that are highly correlated. Having multicollinearity would not lead to biased estimates, but to inaccurate estimates. In such a case, excluding variables from the model should be considered. There are no tests that specifically look for multicollinearity, however, there are certain indications. For instance, having two variables that are jointly significant (have big F-statistics), but independently are not significant can be a sign of multicollinearity

Homoscedasticity

Homoscedasticity implies that the variance of the error term should be the same for all values of the independent variables. If this does not hold, there is a problem with heteroscedasticity meaning that the estimates of the regression are inconsistent due to an inaccuracy of their standard errors, meaning that the t-statistics and thus the significance level of the estimates are not valid anymore. To test for homoscedasticity, I perform the Breusch-Pagan test, which hypothesizes that there is a constant variance of the error terms.

Endogeneity

The last assumption is crucial to be met as otherwise, the regression estimates are biased and inconsistent. Endogeneity implies that there is a correlation between an independent variable and the error term. There are several reasons why this assumption does not fold:

1. The model is misspecified. That is, nonlinearities are missing from the model or interaction effects are not accounted for. To account for that I perform the Ramsey-Reset test. The tests add fitted values on power and re-estimates the model. The intuition behind it is that if a nonlinear combination of independent variables can explain the dependent variable there are evidence the model is misspecified. The Ramsey-Reset test work under null hypothesis that the model has no important omitted non-linearities (Ramsey 1974)
2. Endogeneity, meaning that we are either missing important variables that explain the variance in the independent variables or we have reverse causality, that means that there can be a loop of causality between the independent and dependent variable.

Stepwise regression

The idea of stepwise regression has been introduced by Hastie and Pregibon (1992) and further improved by Ripley (2002). It is an iterative function ran over a restricted model and a set of candidate models. Each candidate model consists of a different set of explanatory variables. The function computes iteratively Akaike information criterion (AIC) values for the models comparing them to the best performing models from the previous iteration and based on the performance chooses whether to continue the loop with the new model or remain with the old one. The final output is the best performing model.

Results

Factor analysis

```
##
## Call:
## factanal(x = RAUAdata, factors = 2, rotation = "varimax")
##
## Uniquenesses:
## Q16_1 Q16_2 Q16_3 Q16_4 Q16_5 Q16_6 Q16_7
##  0.35  0.20  0.64  0.63  0.33  0.25  0.60
##
## Loadings:
##       Factor1 Factor2
## Q16_1         0.71
## Q16_2         0.88
## Q16_3         0.53
## Q16_4 0.49
## Q16_5 0.78
## Q16_6 0.84
## Q16_7 0.55
##
##                Factor1 Factor2
## SS loadings       2.11    1.90
## Proportion Var    0.30    0.27
## Cumulative Var    0.30    0.57
##
## Test of the hypothesis that 2 factors are sufficient.
## The chi square statistic is 29.15 on 8 degrees of freedom.
## The p-value is 0.000299
```

Figure 5: Factor analysis results, risk and uncertainty

```
##
## Call:
## factanal(x = BIG5data, factors = 5, rotation = "varimax")
##
## Uniquenesses:
##  Q15_1  Q15_2  Q15_3  Q15_4  Q15_5  Q15_6  Q15_7  Q15_8  Q15_9 Q15_10
##   0.70   0.49   0.68   0.52   0.60   0.61   0.70   0.21   0.63   0.80
## Q15_11
##   0.48
##
## Loadings:
##        Factor1 Factor2 Factor3 Factor4 Factor5
## Q15_1                          -0.52
## Q15_2                           0.69
## Q15_3                                   0.49
## Q15_4                                  -0.63
## Q15_5                   0.61
## Q15_6                   0.60
## Q15_7  -0.40
## Q15_8   0.89
## Q15_9           0.60
## Q15_10
## Q15_11          0.69
##
##                Factor1 Factor2 Factor3 Factor4 Factor5
## SS loadings       0.99    0.97    0.97    0.91    0.74
## Proportion Var    0.09    0.09    0.09    0.08    0.07
## Cumulative Var    0.09    0.18    0.27    0.35    0.42
##
## Test of the hypothesis that 5 factors are sufficient.
## The chi square statistic is 4.15 on 10 degrees of freedom.
## The p-value is 0.94
```

Figure 6: Factor analysis results, BIG5

OLS

Table 1: Desktop data, restricted model

	Dependent variable:				
	log(MM)	log(Domains)	log(PV)	log(Time)	log(Length)
	(1)	(2)	(3)	(4)	(5)
Risk:seek	−0.16	−0.25**	−0.42***	−0.47***	−0.28**
	(0.12)	(0.12)	(0.13)	(0.13)	(0.12)
Uncertainty:seek	−0.15	−0.26**	−0.31***	−0.32***	−0.21*
	(0.11)	(0.11)	(0.12)	(0.12)	(0.11)
Days	0.01***	0.01***	0.004***	0.004***	0.01***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)
D MM	1.17***				
	(0.10)				
D Domains		1.08***			
		(0.10)			
D PV			0.88***		
			(0.11)		
D Time				0.86***	
				(0.11)	
D Length					1.34***
					(0.10)
D Purchase	0.50***	0.62***	0.86***	0.87***	0.48***
	(0.11)	(0.11)	(0.11)	(0.11)	(0.10)
Risk x Uncertainty	0.27	0.32	0.66**	0.83**	0.48
	(0.30)	(0.31)	(0.32)	(0.32)	(0.29)
Constant	3.12***	3.59***	5.66***	8.94***	10.99***
	(0.27)	(0.27)	(0.29)	(0.29)	(0.26)
Observations	426	426	426	426	426
R^2	0.36	0.35	0.30	0.31	0.41
Adjusted R^2	0.35	0.34	0.29	0.30	0.41
Residual Std. Error (df = 419)	0.97	0.98	1.02	1.03	0.93
F Statistic (df = 6; 419)	39.89***	37.21***	29.69***	31.07***	49.37***
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 2: Mobile data, restricted model

	Dependent variable:				
	log(MM)	log(Domains)	log(PV)	log(Time)	log(Length)
	(1)	(2)	(3)	(4)	(5)
Risk:seek	0.36	0.29	0.42	0.96	1.26
	(0.39)	(0.31)	(0.42)	(0.91)	(0.92)
Uncertainty:seek	0.84**	0.66**	0.54	1.49*	2.26**
	(0.38)	(0.30)	(0.41)	(0.89)	(0.89)
Days	0.01**	0.01***	0.004	0.02***	0.02***
	(0.003)	(0.002)	(0.003)	(0.01)	(0.01)
D MM	1.50***				
	(0.36)				
D Domains		2.28***			
		(0.29)			
D PV			2.55***		
			(0.40)		
D Time				3.09***	
				(0.82)	
D Length					3.13***
					(0.82)
Risk x Uncertainty	−2.50	−2.07	−2.07	−7.86**	−8.74**
	(1.58)	(1.25)	(1.72)	(3.71)	(3.75)
Constant	0.10	0.03	0.37	1.60	1.58
	(0.46)	(0.36)	(0.50)	(1.08)	(1.08)
Observations	101	101	101	101	101
R^2	0.36	0.55	0.42	0.37	0.36
Adjusted R^2	0.33	0.52	0.38	0.33	0.33
Residual Std. Error (df = 95)	1.50	1.19	1.64	3.52	3.56
F Statistic (df = 5; 95)	10.84***	22.94***	13.49***	10.99***	10.72***
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 3: Combined data, restricted model

	Dependent variable:				
	log(MM)	log(Domains)	log(PV)	log(Time)	log(Length)
	(1)	(2)	(3)	(4)	(5)
Risk:seek	−0.20	−0.20	−0.35***	−0.38***	−0.19
	(0.13)	(0.13)	(0.13)	(0.14)	(0.15)
Uncertainty:seek	−0.13	−0.15	−0.23*	−0.19	−0.17
	(0.12)	(0.12)	(0.13)	(0.13)	(0.14)
Days	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
D MM	0.91***				
	(0.11)				
D Domains		1.01***			
		(0.10)			
D PV			0.82***		
			(0.11)		
D Time				0.77***	
				(0.12)	
D Length					0.98***
					(0.12)
D Purchase	0.59***	0.62***	0.90***	0.85***	0.51***
	(0.11)	(0.11)	(0.12)	(0.12)	(0.13)
Risk x Uncertainty	0.25	0.27	0.58*	0.58	0.48
	(0.32)	(0.31)	(0.34)	(0.35)	(0.37)
Constant	2.67***	3.16***	5.08***	8.10***	10.01***
	(0.29)	(0.28)	(0.30)	(0.31)	(0.32)
Observations	429	429	429	429	429
R^2	0.30	0.34	0.30	0.29	0.29
Adjusted R^2	0.29	0.33	0.29	0.28	0.28
Residual Std. Error (df = 422)	1.03	1.00	1.07	1.13	1.16
F Statistic (df = 6; 422)	30.57***	36.53***	29.52***	29.22***	29.27***
Note:					*p<0.1; **p<0.05; ***p<0.01

OLS Assumptions/ Restricted model

Full model

Table 4: Restricted models tests

data_name	Residual_Standard_Err	F.Stat	NumDF	FDenDF	R.Sq	Adj.R.Sq	Shapiro_Wilk_Stat	Shapiro_Wilk_P.val	Reset_Stat	Reset_P.val	BP_Stat	BP_P.val
DESKTOP:REST:log(MM)	0.97	39.89	6	419	0.36	0.35	1.00	0.94	2.03	0.13	4.65	0.59
DESKTOP:REST:log(TD)	0.98	37.21	6	419	0.35	0.34	1.00	0.24	1.20	0.30	10.48	0.11
DESKTOP:REST:log(TDPV)	1.02	29.69	6	419	0.30	0.29	0.99	0.16	0.79	0.45	6.82	0.34
DESKTOP:REST:log(TT)	1.03	31.07	6	419	0.31	0.30	0.98	0.0001	1.31	0.27	10.88	0.09
DESKTOP:REST:log(TL)	0.93	49.37	6	419	0.41	0.41	0.99	0.01	1.97	0.14	7.77	0.26
MOBILE:REST:log(MM)	1.50	10.84	5	95	0.36	0.33	0.96	0.002	0.89	0.42	2.90	0.72
MOBILE:REST:log(TD)	1.19	22.94	5	95	0.55	0.52	0.97	0.01	2.40	0.10	3.37	0.64
MOBILE:REST:log(TDPV)	1.64	13.49	5	95	0.42	0.38	0.91	0.0000	2.04	0.14	1.84	0.87
MOBILE:REST:log(TT)	3.52	10.99	5	95	0.37	0.33	0.98	0.06	0.03	0.97	16.62	0.01
MOBILE:REST:log(TL)	3.56	10.72	5	95	0.36	0.33	0.97	0.02	0.45	0.64	14.80	0.01
COMBINED:REST:log(MM)	1.03	30.57	6	422	0.30	0.29	0.99	0.02	10.19	0.0000	11.52	0.07
COMBINED:REST:log(TD)	1.00	36.53	6	422	0.34	0.33	0.99	0.001	8.68	0.0002	11.67	0.07
COMBINED:REST:log(TDPV)	1.07	29.52	6	422	0.30	0.29	0.98	0.0000	9.92	0.0001	19.72	0.003
COMBINED:REST:log(TT)	1.13	29.22	6	422	0.29	0.28	0.93	0	12.93	0.0000	26.66	0.0002
COMBINED:REST:log(TL)	1.16	29.27	6	422	0.29	0.28	0.85	0	3.94	0.02	30.10	0.0000

Note:
Based on the results of Reset-test, all of the models containing COMBINED:REST have been rejected

Table 5: Full model: Desktop, robust

	Dependent variable:				
	D log(MM)	D log(TD)	D log(TDPV)	D log(TT)	D log(TL)
	(1)	(2)	(3)	(4)	(5)
Risk Seeking	−0.11	−0.23*	−0.37***	−0.46***	−0.26**
	p = 0.36	p = 0.08	p = 0.005	p = 0.0002	p = 0.03
Uncertainty Seeking	−0.13	−0.25**	−0.34***	−0.37***	−0.17
	p = 0.26	p = 0.04	p = 0.004	p = 0.002	p = 0.12
Risk(x)Uncertainty: Seeking	0.09	0.10	0.39	0.76**	0.30
	p = 0.73	p = 0.69	p = 0.16	p = 0.02	p = 0.22
Days active	0.01***	0.01***	0.01**	0.01**	0.005***
	p = 0.001	p = 0.002	p = 0.02	p = 0.04	p = 0.01
D Purchase	0.53***	0.63***	0.84***	0.88***	0.50***
	p = 0.0000	p = 0.0000	p = 0.00	p = 0.00	p = 0.0000
D Micromoments	1.13***				
	p = 0.00				
D Domains		1.03***			
		p = 0.00			
D PageViews			0.83***		
			p = 0.00		
D Time				0.80***	
				p = 0.00	
D MM Length					1.36***
					p = 0.00
S1. x2 travel related purchase		0.21*	0.21*	0.29**	
		p = 0.06	p = 0.08	p = 0.02	
S1. x3+ travel related purchase		0.15	0.20	0.28**	
		p = 0.21	p = 0.12	p = 0.02	
S3. Not the only decision maker	−0.20*	−0.28***	−0.33***	−0.26**	−0.26***
	p = 0.06	p = 0.01	p = 0.002	p = 0.02	p = 0.005
D1. Age	1.01	2.58**	−0.52		
	p = 0.33	p = 0.02	p = 0.63		
D1. Age Sq	1.87*	2.14**	2.01**		
	p = 0.06	p = 0.02	p = 0.04		
Q1. Country: Asia (Base Europe)		−0.51	−0.30	0.08	0.33*
		p = 0.20	p = 0.45	p = 0.84	p = 0.08
Q1. Country: North America (Base Europe)		0.65	0.65	0.57	0.38**
		p = 0.34	p = 0.37	p = 0.44	p = 0.02
Q1. Country: South America (Base Europe)		−0.01	−0.82	−1.13*	0.25
		p = 0.98	p = 0.18	p = 0.07	p = 0.29
Q1. Country: Australia (Base Europe)		−0.64	−0.65	−0.36	−0.63
		p = 0.21	p = 0.25	p = 0.60	p = 0.14
Q1. Country: Africa (Base Europe)		−0.05	0.12	0.35	−0.14
		p = 0.83	p = 0.59	p = 0.19	p = 0.63
Q4. Planning horizon (Weeks)			0.01		
			p = 0.18		
Q5.3. Used tourist information office			0.43*		0.39
			p = 0.07		p = 0.16
Q5.5. Used travel agent	−0.28**	−0.38**	−0.37**	−0.31**	
	p = 0.04	p = 0.02	p = 0.02	p = 0.05	
Q6.1. Purchased online: Transport	−0.21**	−0.24**	−0.29***	−0.37***	−0.17*
	p = 0.05	p = 0.02	p = 0.005	p = 0.0003	p = 0.09
Q7. Trip longer than 3 nights	0.18**	0.20**	0.14	0.20**	0.19**
	p = 0.05	p = 0.04	p = 0.17	p = 0.05	p = 0.05
Q8. Visited before	−0.16	−0.17*			
	p = 0.14	p = 0.10			
Q11.1. Children (N)	−0.27*	−0.25*	−0.34**	−0.23	
	p = 0.07	p = 0.09	p = 0.03	p = 0.11	
Q11.3. Stayed at Friends/Relatives (N)	0.27*	0.36**	0.37**		0.38**
	p = 0.10	p = 0.02	p = 0.03		p = 0.02
Q11.5. Group Trip (N)	−0.43*				−0.41**
	p = 0.07				p = 0.04
Openness to experience (low)			−0.32*	−0.33**	
			p = 0.06	p = 0.04	
Consciousness (low)				−0.20*	
				p = 0.06	
Agreeableness (low)			−0.40	−0.36	
			p = 0.22	p = 0.21	
Asia x Travel Agent		1.16***	0.79*	0.54	
		p = 0.01	p = 0.08	p = 0.23	
North America x Travel Agent		−0.31	−0.46	−0.37	
		p = 0.65	p = 0.53	p = 0.62	
South America x Travel Agent		0.32	1.12*	1.74***	
		p = 0.61	p = 0.09	p = 0.01	
Australia x Travel Agent		0.68	1.09*	0.55	
		p = 0.20	p = 0.09	p = 0.45	
Constant	3.91***	3.89***	5.36***	9.29***	10.69***
	p = 0.00	p = 0.00	p = 0.00	p = 0.00	p = 0.00

Note: *p<0.1; **p<0.05; ***p<0.01

OLS Assumptions / Full model

Table 6: Full model: Mobile, robust

	Dependent variable:				
	M log(MM)	M log(TD)	M log(TDPV)	M log(TT)	M log(TL)
	(1)	(2)	(3)	(4)	(5)
Risk Seeking	0.16	0.13	0.53	1.32	1.09
	p = 0.66	p = 0.61	p = 0.22	p = 0.17	p = 0.21
Uncertainty Seeking	0.72**	0.73**	0.57	1.98*	2.05**
	p = 0.04	p = 0.02	p = 0.11	p = 0.06	p = 0.02
Risk(x)Uncertainty: Seeking	−1.76***	−1.80***	−0.75	−7.89***	−7.37***
	p = 0.003	p = 0.0000	p = 0.34	p = 0.00	p = 0.0000
Days active	0.01***	0.01***	0.004	0.02***	0.02***
	p = 0.001	p = 0.0001	p = 0.12	p = 0.0000	p = 0.0001
D Micromoments	1.03***				
	p = 0.004				
D Domains		1.77***			
		p = 0.0000			
D PageViews			2.44***		
			p = 0.0000		
D Time				2.06***	
				p = 0.01	
D MM Length					2.20***
					p = 0.003
S2. Primary purchase: Flight	−1.09***	−0.55**	−1.02***	−1.22*	−1.06
	p = 0.0002	p = 0.03	p = 0.01	p = 0.07	p = 0.12
S3. Not the only decision maker	−0.70**	−0.40*	−1.03***	−2.19***	−2.51***
	p = 0.02	p = 0.10	p = 0.0005	p = 0.001	p = 0.0003
D1. Age			−3.51**		
			p = 0.02		
D1. Age Sq			−0.97		
			p = 0.55		
Q1. Country: Asia (Base Europe)			1.99***		
			p = 0.01		
Q1. Country: North America (Base Europe)			0.33		
			p = 0.50		
Q1. Country: South America (Base Europe)			0.40		
			p = 0.52		
Q1. Country: Australia (Base Europe)			0.90		
			p = 0.30		
Q1. Country: Africa (Base Europe)	0.04**	0.03**	0.04*	0.06**	0.05
	p = 0.02	p = 0.04	p = 0.06	p = 0.05	p = 0.14
Q4. Planning horizon (Weeks)			0.80**		
			p = 0.03		
Q5.2. Used advice of friends or relatives				−2.86**	
				p = 0.04	
Q5.3. Used tourist information office	−0.49	−0.39	−0.95**	−1.29	−1.60*
	p = 0.19	p = 0.18	p = 0.02	p = 0.19	p = 0.10
Q5.4. Used travel magazines			1.37*		
			p = 0.09		
Q5.5. Used travel agent		0.48*		2.47***	2.05***
		p = 0.08		p = 0.001	p = 0.004
Q6.1. Purchased online: Transport	0.62		0.90**		
	p = 0.11		p = 0.04		
Q6.3. Purchased online: Entertainment	−0.89***	−0.62***	−0.53	−1.87***	−1.73***
	p = 0.001	p = 0.01	p = 0.13	p = 0.004	p = 0.01
Q7. Trip longer than 3 nights	0.59**		0.53*	1.45**	1.77***
	p = 0.02		p = 0.07	p = 0.02	p = 0.004
Q8. Visited before					0.58***
					p = 0.01
Q11.2. Visited Friends/Relatives (N)			1.23***		
			p = 0.01		
Q11.3. Stayed at Friends/Relatives (N)	1.08***	0.93***		3.49***	3.11***
	p = 0.004	p = 0.005		p = 0.0002	p = 0.002
Q11.4. Stayed at Hotel/Motel/AirBNB (N)	0.91*	0.61	0.88	2.21**	2.01**
	p = 0.07	p = 0.12	p = 0.12	p = 0.05	p = 0.05
Consciousness (low)	0.60**	0.53**			
	p = 0.04	p = 0.03			
Extraversion (low)			0.88**		
			p = 0.02		
Agreeableness (low)	−1.17*	−1.26***	−1.86***	−3.56*	−3.08
	p = 0.08	p = 0.005	p = 0.01	p = 0.06	p = 0.13
Neuroticism (low)		0.67***	0.74**	2.88***	3.27***
		p = 0.002	p = 0.03	p = 0.004	p = 0.0001
North America x Travel Agent			−2.20**		
			p = 0.03		
Constant	−0.13	−0.98	−2.64*	−0.98	−4.42**
	p = 0.86	p = 0.13	p = 0.06	p = 0.66	p = 0.03

Note: *p<0.1; **p<0.05; ***p<0.01

Table 7: Full model tests, Desktop

	Residual_Standard_Err	F.Stat	NumDF	FDenDF	R.Sq	Adj.R.Sq	Shapiro_Wilk_Stat	Shapiro_Wilk_P.val	Reset_Stat	Reset_P.val	BP_Stat	BP_P.val
log(MM)	0.942	17.762	16	409	0.410	0.387	0.997	0.655	0.789	0.455	12.209	0.729
log(TD)	0.943	11.378	26	399	0.426	0.388	0.993	0.056	2.085	0.126	23.757	0.644
log(TDPV)	0.979	12.899	19	406	0.376	0.347	0.991	0.008	0.044	0.957	29.519	0.058
log(TT)	0.982	10.797	24	401	0.393	0.356	0.977	0.00000	3.194	0.042	36.150	0.069
log(TL)	0.903	20.830	17	408	0.465	0.442	0.993	0.058	0.575	0.563	19.391	0.307

Akaike, Hirotugu. 1974. "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control* 19 (6). Ieee: 716–23.

Hastie, TJ, and D Pregibon. 1992. "Statistical Models in S, Chapter Generalized Linear Models." *Wadsworth & Brooks/Cole* 51.

Ramsey, James B. 1974. "Classical Model Selection Through Specification Error Tests." *Frontiers in Econometrics*. Academic Press New York, 13–47.

Ripley, BD. 2002. "Modern Applied Sta Tistics with S." Springer-Verlag, New York.

Shapiro, Samuel Sanford, and Martin B Wilk. 1965. "An Analysis of Variance Test for Normality (Complete Samples)." *Biometrika* 52 (3/4). JSTOR: 591–611.

Verbeek, Marno. 2008. *A Guide to Modern Econometrics*. John Wiley & Sons.

Yong, An Gie, and Sean Pearce. 2013. "A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis." *Tutorials in Quantitative Methods for Psychology* 9 (2): 79–94.

Table 8: Full model tests, Mobile

	Residual_Standard_Err	F.Stat	NumDF	FDenDF	R.Sq	Adj.R.Sq	Shapiro_Wilk_Stat	Shapiro_Wilk_P.val	Reset_Stat	Reset_P.val	BP_Stat	BP_P.val
log(MM)	1.288	8.393	14	86	0.577	0.509	0.989	0.543	3.061	0.052	17.158	0.248
log(TD)	1.037	11.351	17	83	0.699	0.638	0.982	0.188	1.682	0.192	10.712	0.871
log(TDPV)	1.373	5.848	27	73	0.684	0.567	0.991	0.761	1.816	0.170	29.234	0.505
log(TT)	2.855	8.979	16	84	0.631	0.561	0.987	0.452	3.920	0.024	21.297	0.167
log(TL)	2.909	8.691	16	84	0.623	0.552	0.988	0.520	4.248	0.018	21.391	0.164