Supplementary File: Holiday Spending Analyses

Sara J. Weston [1], Joe J. Gladstone [2], Eileen K. Graham [1], Daniel K. Mroczek [1,3], David M. Condon [1]

This document contains both the planned and exploratory analyses for the Holiday Spending project. Analyses were pre-registered at OSF. The preregistration form can be found at osf.io/ew4h5. Additional project information can be found at osf.io/7j9da.

Contents

[1] Department of Medical Social Sciences, Northwestern University, Chicago, Illinois.
[2] University College London, London, UK
[3] Department of Psychology, Northwestern University, Evanston, Illinois.

1 Data Cleaning

This section documents the methods used to clean and transform data prior to data-analysis.

First we load the data, which is stored in a .csv file. The object data.path stores the location of the data file on the machine.

```
holiday <- read.csv(paste0(data.path, "May 2017 to June 2016 - subsample.csv"),
    stringsAsFactors = FALSE)
```

We remove duplicate rows.

```
holiday <- unique(holiday)
```

The variables used throughout the analyses are listed in Table 1.

Variable	Description
userreference	Unique person ID number
transactiondate	Date of transaction (originally a string, transformed to a date)
creditdebit	Whether transaction is credits or debits the account, character class
userprecedencetagname	How the user tagged the transaction. Categories provided
amount	The total amount of money (in British pounds) in the transaction
salaryrange	self-reported income, factor class variable
derivedgender	estimated gender
big5_e	Extraversion item: "I see myself as someone who is outgoing, sociable."
big5_e_r	Extraversion item: "I see myself as someone who is reserved."
big5_a	Agreeableness item: "I see myself as someone who is generally trusting."
big5_a_r	Agreeableness item: "I see myself as someone who tends to find fault with others."
big5_c	Conscientiousness item: "I see myself as someone who does a thorough job."
big5_c_r	Conscientiousness item: "I see myself as someone who tends to be lazy."
big5_n	Neuroticism item: "I see myself as someone who gets nervous easily."
big5_n_r	Neuroticism item: "I see myself as someone who is relaxed, handles stress well."
big5_o	Openness item: "I see myself as someone who has an active imagination"
big5_o_r	Openness item: "I see myself as someone who has few artistic interests."
yearofbirth	Year the participant was born.

Table 1: Variables used in analyses

Transaction date is stored as a character variable. We make transaction date variable class Date.

```
head(holiday$transactiondate)

## [1] "2017-05-12" "2017-05-04" "2017-05-24" "2017-05-24" "2017-05-27"
## [6] "2017-04-28"

holiday$transactiondate <- as.Date(holiday$transactiondate)
```

Some participants provided data for longer than one year. We choose to remove these transactions. We acknowledge that this limits power; however, given the changing nature of the economy, we feel this is more defensible than attempting to aggregate spending patterns across different years. We ultimately remove 1.41171×10^5 transactions, which accounts for 5 % of the data.

```
length(which(holiday$transactiondate <= "2016-05-30"))

## [1] 141171

length(which(holiday$transactiondate <= "2016-05-30"))/nrow(holiday)

## [1] 0.05418219

holiday <- holiday %>% filter(transactiondate > "2016-05-30")
```

We remove“debit” categories that are actually credit card payments (i.e., already spent).

```
holiday <- subset(holiday, !(creditdebit == "Debit" &
                              userprecedencetagname %in% c("Credit card payment",
                                                           "Credit card repayment",
                                                           "Store card repayment")))
```

We remove non-income credit categories.

```
holiday <- subset(holiday, creditdebit == "Debit" |
                    (creditdebit == "Credit" &
                       userprecedencetagname %in% c("Salary (main)",
                                                    "Family benefits",
                                                    "Interest income",
                                                    "Rental income (whole property)",
                                                    "Miscelleneous income - other",
                                                    "Salary (secondary)",
                                                    "Other benefits",
                                                    "Gambling account",
                                                    "Work pension",
                                                    "State pension",
                                                    "Gift",
                                                    "Winnings",
                                                    "Mortgage release",
                                                    "Job seekers benefits",
                                                    "Rental income (room)",
                                                    "Employment - other",
                                                    "Incapacity benefits",
                                                    "Sale - other",
                                                    "Dividend",
                                                    "Investment - other",
                                                    "Tax rebate",
                                                    "Property - other",
                                                    "Child support",
                                                    "Bonus",
                                                    "Investment income - other",
                                                    "Pension - other",
                                                    "Bursary",
                                                    "Other account",
                                                    "Sharedealing account",
                                                    "Inheritance",
                                                    "Vehicle",
                                                    "Bond income",
                                                    "Overtime",
                                                    "Divorce settlement")))
```

1.1 Transformations and indices

We calculate the total amount of credit and debit for each particpant based on their transactions in each category. This measure of credit and debit (i.e., spending) is more objective than self-reported income and spending amounts, as it is based on real-world behavior. However, we also acknowledge that these data are noisy and may omit forms of spending or credit that are not linked to the app. We provide the correlations between self-report income and the estimated values for reference.

```
holiday = holiday %>%
  group_by(userreference, creditdebit) %>%
  summarize(amount = sum(amount, na.rm=T)) %>%
  ungroup() %>%
  spread(key = "creditdebit", value = "amount") %>%
  rename(yearly_spending = Debit,
         yearly_income = Credit) %>%
  dplyr::select(userreference, yearly_spending, yearly_income) %>%
  full_join(holiday)

## Joining, by = "userreference"
```

Note: After running analyses, it became clear that participants provided data across different spans of time. That is some participants provided data across a year, some provided data across less than a year and *some provided data across more than a year. Consequently, we calculate adjusted credit and spending amounts by calculating the average daily credit and debit, only from days outside the holiday season, and multiplying by 61 (the number of days in November and December) to estimate bi-monthly credit and spending.*

```
holiday <- holiday %>%
  group_by(userreference, creditdebit, transactiondate) %>%
  filter(transactiondate < "2016-11-01") %>%
  summarize(amount = sum(amount, na.rm=T)) %>%
  ungroup() %>%
  spread(key = "creditdebit", value = "amount") %>%
  group_by(userreference) %>%
  summarize(n = n(),
            total_credit = sum(Credit, na.rm=T),
            total_spend = sum(Debit, na.rm=T),
            daily_credit = total_credit/n,
            daily_spend = total_spend/n,
            est_bi_credit = daily_credit*61,
            est_bi_spend = daily_spend*61) %>%
  merge(holiday)
```

Now that we have estimates of bi-monthly credit (i.e., income), we filter out individual credit transactions.

```
holiday <- subset(holiday, creditdebit == "Debit")
```

We create a numeric version of the self-reported income variable. We do so by first assigning each response the value that is in the middle of the range they specified. We then specify the class of variable as numeric, to allow for correlations.

```
holiday$income <- dplyr::recode(holiday$salaryrange,
                              `< 10K` = "5",
                              `10K to 20K` = "15",
                              `20K to 30K` = "25",
                              `30K to 40K` = "35",
                              `40K to 50K` = "45",
                              `50K to 60K` = "55",
                              `60K to 70K` = "65",
                              `70K to 80K` = "75",
                              `> 80K` = "85",
                              .missing="missing")

holiday$income <- as.numeric(holiday$income)
```

This variable correlates 0.06 with the estimated bi-monthly income variable.

```
cor.test(holiday$income, holiday$est_bi_credit)

##
##  Pearson's product-moment correlation
##
## data:  holiday$income and holiday$est_bi_credit
## t = 75.516, df = 1815700, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.05450385 0.05740381
## sample estimates:
##        cor
## 0.05595395
```

We create a binary gender variable, male, in which 1 indicates whether the participant was identified as male, and 0 if the participant was identified as female or not identified.

```
holiday$male <- ifelse(holiday$derivedgender == "M", 1, 0)
```

1.2 Calculate personality scores

1.2.1 Average scores

We originally calculated personality scores by reverse scoring the negatively keyed variables (indiciated by "_r") and averaging responses.

```
personality.keys = list(
  extra = c("big5_e", "-big5_e_r"),
  agree = c("big5_a", "-big5_a_r"),
  con = c("big5_c", "-big5_c_r"),
  neur = c("big5_n", "-big5_n_r"),
  open = c("big5_o", "-big5_o_r"))
personality.keys <- make.keys(holiday[,grepl("big5", names(holiday))], personality.keys)
score.personality <- scoreItems(keys = personality.keys,
                                items = holiday[,grepl("big5", names(holiday))])

score.personality

## Call: scoreItems(keys = personality.keys, items = holiday[, grepl("big5",
##     names(holiday))])
##
## (Unstandardized) Alpha:
##       extra agree  con neur open
## alpha  0.75  0.28 0.57 0.64 0.32
##
## Standard errors of unstandardized Alpha:
##       extra  agree    con   neur   open
## ASE   0.001 0.0013 0.0012 0.0011 0.0013
##
## Average item correlation:
##           extra agree con neur open
## average.r   0.6  0.16 0.4 0.47 0.19
##
## Median item correlation:
## extra agree   con  neur  open
##  0.60  0.17  0.41  0.48  0.20
##
##  Guttman 6* reliability:
##          extra agree  con neur open
## Lambda.6  0.64  0.24 0.45 0.55 0.26
##
## Signal/Noise based upon av.r :
##              extra agree con neur open
## Signal/Noise   2.9  0.39 1.4  1.8 0.48
##
## Scale intercorrelations corrected for attenuation
##  raw correlations below the diagonal, alpha on the diagonal
##  corrected correlations above the diagonal:
##        extra agree    con   neur  open
## extra  0.747  0.13  0.069 -0.119  0.66
## agree  0.061  0.28  0.332 -0.384  0.47
## con    0.045  0.13  0.575 -0.473  0.13
## neur  -0.083 -0.16 -0.287  0.643 -0.13
## open   0.323  0.14  0.054 -0.058  0.32
##
##  In order to see the item by scale loadings and frequency counts of the data
##  print with the short option = FALSE

holiday <- cbind(holiday, score.personality$scores)
```

We standardize these personality variables.

```
holiday_p = holiday %>%
  group_by(userreference) %>%
  filter(row_number() == 1) %>%
  dplyr::select(userreference, extra, agree, con, neur, open)

holiday_p$z_extra = as.numeric(scale(holiday_p$extra))
holiday_p$z_agree = as.numeric(scale(holiday_p$agree))
holiday_p$z_con   = as.numeric(scale(holiday_p$con))
holiday_p$z_neur = as.numeric(scale(holiday_p$neur))
holiday_p$z_open  = as.numeric(scale(holiday_p$open))
```

```r
holiday = holiday %>%
  full_join(holiday_p)
```

1.2.2 Latent scores

Based on feedback during the review process, we also calculate latent personality traits scores. These were used in the results presented in the manuscript. However, we use this supplementary file to document all analyses performed, and so models with the average scores (instead of latent scores) are presented later in this document.

To estimate latent factor scores, we construct a single model in which each latent trait is indicated by two items. Latent variables are allowed to correlate. The model resulted in some negative variances; when this happened, we set those residual variances to 0.

```r
#reverse score personality variables
holiday[,rownames(personality.keys)] = reverse.code(keys = rowSums(personality.keys),
                                                    items = holiday[,rownames(personality.keys)])
#select one observation for each person
holiday_p = holiday %>%
  group_by(userreference) %>%
  filter(row_number() == 1)
```

```r
latent.p = '
latent.e =~ big5_e + big5_e_r

latent.a =~ big5_a + big5_a_r

latent.c =~ big5_c + big5_c_r

latent.n =~ big5_n + big5_n_r

latent.o =~ big5_o + big5_o_r

big5_a ~~ 0*big5_a
big5_n_r ~~ 0*big5_n_r
'

cfa.p = cfa(model = latent.p, data = holiday_p, REML=TRUE, missing = "pairwise")
summary(cfa.p, fit.measures = T)

## lavaan (0.5-23.1097) converged normally after  80 iterations
##
##                                                   Used       Total
##   Number of observations                          1981        2168
##
##   Estimator                                         ML
##   Minimum Function Test Statistic              473.041
##   Degrees of freedom                                27
##   P-value (Chi-square)                           0.000
##
## Model test baseline model:
##
##   Minimum Function Test Statistic             3095.371
##   Degrees of freedom                                45
##   P-value                                        0.000
##
## User model versus baseline model:
##
##   Comparative Fit Index (CFI)                    0.854
##   Tucker-Lewis Index (TLI)                       0.756
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)             -33487.122
##   Loglikelihood unrestricted model (H1)     -33250.602
##
##   Number of free parameters                         28
##   Akaike (AIC)                               67030.244
##   Bayesian (BIC)                             67186.802
##   Sample-size adjusted Bayesian (BIC)        67097.845
```

```
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.091
##   90 Percent Confidence Interval          0.084  0.099
##   P-value RMSEA <= 0.05                          0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.060
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   latent.e =~
##     big5_e            1.000
##     big5_e_r          0.637    0.048   13.169    0.000
##   latent.a =~
##     big5_a            1.000
##     big5_a_r          0.258    0.029    8.886    0.000
##   latent.c =~
##     big5_c            1.000
##     big5_c_r          0.571    0.066    8.591    0.000
##   latent.n =~
##     big5_n            1.000
##     big5_n_r          1.741    0.071   24.629    0.000
##   latent.o =~
##     big5_o            1.000
##     big5_o_r          0.708    0.087    8.097    0.000
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   latent.e ~~
##     latent.a          0.415    0.043    9.716    0.000
##     latent.c          0.129    0.044    2.968    0.003
##     latent.n         -0.038    0.027   -1.417    0.156
##     latent.o          0.551    0.038   14.480    0.000
##   latent.a ~~
##     latent.c          0.128    0.034    3.727    0.000
##     latent.n         -0.085    0.021   -3.987    0.000
##     latent.o          0.250    0.028    8.880    0.000
##   latent.c ~~
##     latent.n         -0.343    0.027  -12.666    0.000
##     latent.o          0.128    0.029    4.477    0.000
##   latent.n ~~
##     latent.o         -0.054    0.018   -3.075    0.002
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .big5_a            0.000
##    .big5_n_r          0.000
##    .big5_e            0.034    0.162    0.211    0.833
##    .big5_e_r          1.652    0.084   19.615    0.000
##    .big5_a_r          2.444    0.078   31.472    0.000
##    .big5_c            0.297    0.142    2.084    0.037
##    .big5_c_r          1.746    0.072   24.168    0.000
##    .big5_n            1.939    0.062   31.472    0.000
##    .big5_o            0.679    0.060   11.373    0.000
##    .big5_o_r          1.948    0.068   28.689    0.000
##     latent.e          2.321    0.179   12.990    0.000
##     latent.a          1.460    0.046   31.472    0.000
##     latent.c          1.306    0.151    8.670    0.000
##     latent.n          0.594    0.052   11.468    0.000
##     latent.o          0.424    0.062    6.826    0.000

#estimate latent variable scores for each person
latent.scores = as.data.frame(predict(cfa.p))
latent.scores$userreference = holiday_p$userreference
holiday = latent.scores %>%
  full_join(holiday)
```

The correlations between the average personality scores and latent variable scores are presented in Table

	extra	agree	con	neur	open	latent.e	latent.a	latent.c	latent.n	latent.o
extra	1.00	0.06	0.05	-0.08	0.32	0.89	0.16	0.04	0.04	0.65
agree	0.06	1.00	0.13	-0.16	0.14	0.10	0.69	0.14	-0.19	0.26
con	0.05	0.13	1.00	-0.29	0.05	0.03	0.08	0.86	-0.33	0.16
neur	-0.08	-0.16	-0.29	1.00	-0.06	-0.08	-0.02	-0.36	0.83	-0.16
open	0.32	0.14	0.05	-0.06	1.00	0.34	0.21	0.06	0.00	0.77
latent.e	0.89	0.10	0.03	-0.08	0.34	1.00	0.22	0.07	-0.01	0.73
latent.a	0.16	0.69	0.08	-0.02	0.21	0.22	1.00	0.14	-0.11	0.44
latent.c	0.04	0.14	0.86	-0.36	0.06	0.07	0.14	1.00	-0.44	0.24
latent.n	0.04	-0.19	-0.33	0.83	0.00	-0.01	-0.11	-0.44	1.00	-0.13
latent.o	0.65	0.26	0.16	-0.16	0.77	0.73	0.44	0.24	-0.13	1.00

Table 2: Correlations between scale scores and latent personality variables

2.

```
kable(cor(holiday[,c("extra", "agree", "con", "neur", "open",
               "latent.e", "latent.a", "latent.c",
               "latent.n", "latent.o")], use = "pairwise"),
      digits = 2, booktabs = TRUE, escape = FALSE,
      caption = "Correlations between scale scores and latent personality variables \\label{tab:pers}")
  kable_styling()
```

1.3 Wide-form data set

This data summarizes information for each participant, e.g., holiday spending, total spending, etc., so that each participant has one exactly row.

First we caculate each person's total spending the holiday season.

```
holiday_wide <- holiday %>%
  group_by(userreference) %>%
  filter(transactiondate >= "2016-11-01" & transactiondate <= "2016-12-31") %>%
  summarize(NovDec_spending = sum(amount, na.rm=T))
```

Next, we merge this with the first row of data for each person and the holiday_calc data, which contains calculated yearly income and yearly spending.

```
holiday_wide <- holiday %>%
  group_by(userreference) %>%
  filter(row_number() == 1) %>%
  merge(holiday_wide)
```

In this data set, we standardized yearly spending, estimate age in 1960, and calculate the ratio of holiday spending to total spending. We also calcualte the ratio of yearly spending to yearly income (Credit), in the event of multicollinearity issues.

```
holiday_wide = holiday_wide %>%
  mutate(
    # standardized yearly spending
       z_yearlyspending = as.numeric(scale(yearly_spending)),
    # calculate age from year of birth and 2016, then standardize
       age = 2016-yearofbirth,
       z_age = as.numeric(scale(age)),
    # standardize estimated bi-monthly income and spending
       z_est_bi_credit = as.numeric(scale(est_bi_credit)),
       z_est_bi_spend = as.numeric(scale(est_bi_spend)))
```

We also log-transform spending and income variables.

```
holiday_wide = holiday_wide %>%
  mutate(log_NovDec = log(NovDec_spending),
         log_credit = log(est_bi_credit + 1),
         log_spend = log(est_bi_spend + 1))
```

1.4 Holiday season data set

This data set contains daily spending amounts for each person and allows us to track trajectories of spending. To create this, we sum spending each day for each person and merge this with the person-level predictors. We also create two new variables: the number of days since November 1 and the ratio of spending in that day to the total amount spent during the holiday season.

```
holiday_novdec <- holiday %>%
  filter(transactiondate >= "2016-11-01" & transactiondate <= "2016-12-31") %>%
  group_by(userreference, transactiondate) %>%
  summarize(amount = sum(amount, na.rm=T)) %>%
  merge(subset(holiday_wide, select=c(userreference, age, z_age, male,
                                      est_bi_credit, est_bi_spend,
                                      z_est_bi_credit, z_est_bi_spend,
                                      NovDec_spending, z_yearlyspending,
                                      extra, agree, con, neur, open,
                                      z_extra, z_agree, z_con, z_neur, z_open,
                                      latent.e, latent.a, latent.c, latent.n, latent.o))) %>%
  mutate(days = as.numeric(transactiondate - as.Date("2016-11-01")),
         daily_ratio = amount/NovDec_spending)
```

Again, we calculate the log-transformed spending and income variables.

```
holiday_novdec = holiday_novdec %>%
  mutate(log_amount = log(amount + 1),
         log_credit = log(est_bi_credit + 1),
         log_spend = log(est_bi_spend + 1))
```

We save our two data sets.

```
save(holiday_wide, holiday_novdec, file = "cleaned_holiday.Rdata")
```

Category	Number	Percent
F	257	0.12
M	946	0.45
U	915	0.43

Table 3: Gender breakdown in sample

2 Descriptive statistics

2.1 Frequencies and summary statistics

This section describes the cleaned data set. Table 3 displays the frequency and percentage of each gender category (F - female, M - male, U - unknown). As a reminder, these genders are not self-reported, but are dervied from the data by the company who created the money management app.

```
gender.tab = data.frame(N = table(holiday_wide$derivedgender),
                        percent = table(holiday_wide$derivedgender)/nrow(holiday_wide))
kable(gender.tab[,-3], booktabs = TRUE, escape = FALSE, digits = 2,
      col.names = c("Category", "Number", "Percent"),
      caption = "Gender breakdown in sample \\label{tab:gender}") %>%
  kable_styling()
```

Table 4 displays the descriptive statistics for the variables used in the published analyses.

```
holiday_wide %>%
  select(age, male, n,
         est_bi_credit, est_bi_spend,
         NovDec_spending,
         latent.e, latent.a, latent.c, latent.n, latent.o) %>%
  rename(Age = age,
         Male = male,
         `Number of transactions` = n,
         `Estimated Bimonthly Income` = est_bi_credit,
         `Estimated Bimonthly Spending` = est_bi_spend,
         `Holiday Spending` = NovDec_spending,
         `Extraversion (latent score)` = latent.e,
         `Agreeableness (latent score)` = latent.a,
         `Conscientiousness (latent score)` = latent.c,
         `Neuroticism (latent score)` = latent.n,
         `Openness (latent score)` = latent.o) %>%

  describe() %>%
  dplyr::select(-vars, -trimmed, -mad, -range) %>%
  kable(., booktabs = TRUE, escape = FALSE, digits = 2,
        caption = "Descriptive statistics \\label{tab:describe}") %>%
  kable_styling(font_size = 10)
```

We also print the interquartile range of the spending and credit variables.

```
IQR(holiday_wide$NovDec_spending)

## [1] 12158.54
```

	n	mean	sd	median	min	max	skew	kurtosis	se
Age	2118	37.45	11.87	35.00	14.00	86.00	1.02	0.82	0.26
Male	2118	0.45	0.50	0.00	0.00	1.00	0.21	-1.95	0.01
Number of transactions	2118	95.65	37.84	103.00	1.00	154.00	-0.81	-0.05	0.82
Estimated Bimonthly Income	2118	1751.18	5735.49	156.85	0.00	116931.46	11.03	173.34	124.63
Estimated Bimonthly Spending	2118	23028.46	36824.73	13727.48	0.00	873559.07	9.17	154.58	800.16
Holiday Spending	2118	17527.43	69388.78	8758.45	1.82	2827642.79	32.72	1280.74	1507.74
Extraversion (latent score)	1935	0.00	1.52	-0.15	-3.16	2.82	-0.20	-0.81	0.03
Agreeableness (latent score)	1935	0.00	1.21	-0.27	-4.27	1.73	-0.78	0.54	0.03
Conscientiousness (latent score)	1935	0.01	1.04	0.04	-3.65	1.60	-0.86	0.61	0.02
Neuroticism (latent score)	1935	0.00	0.77	0.08	-1.06	2.38	0.72	0.02	0.02
Openness (latent score)	1935	0.00	0.49	0.02	-2.01	1.18	-0.41	0.36	0.01

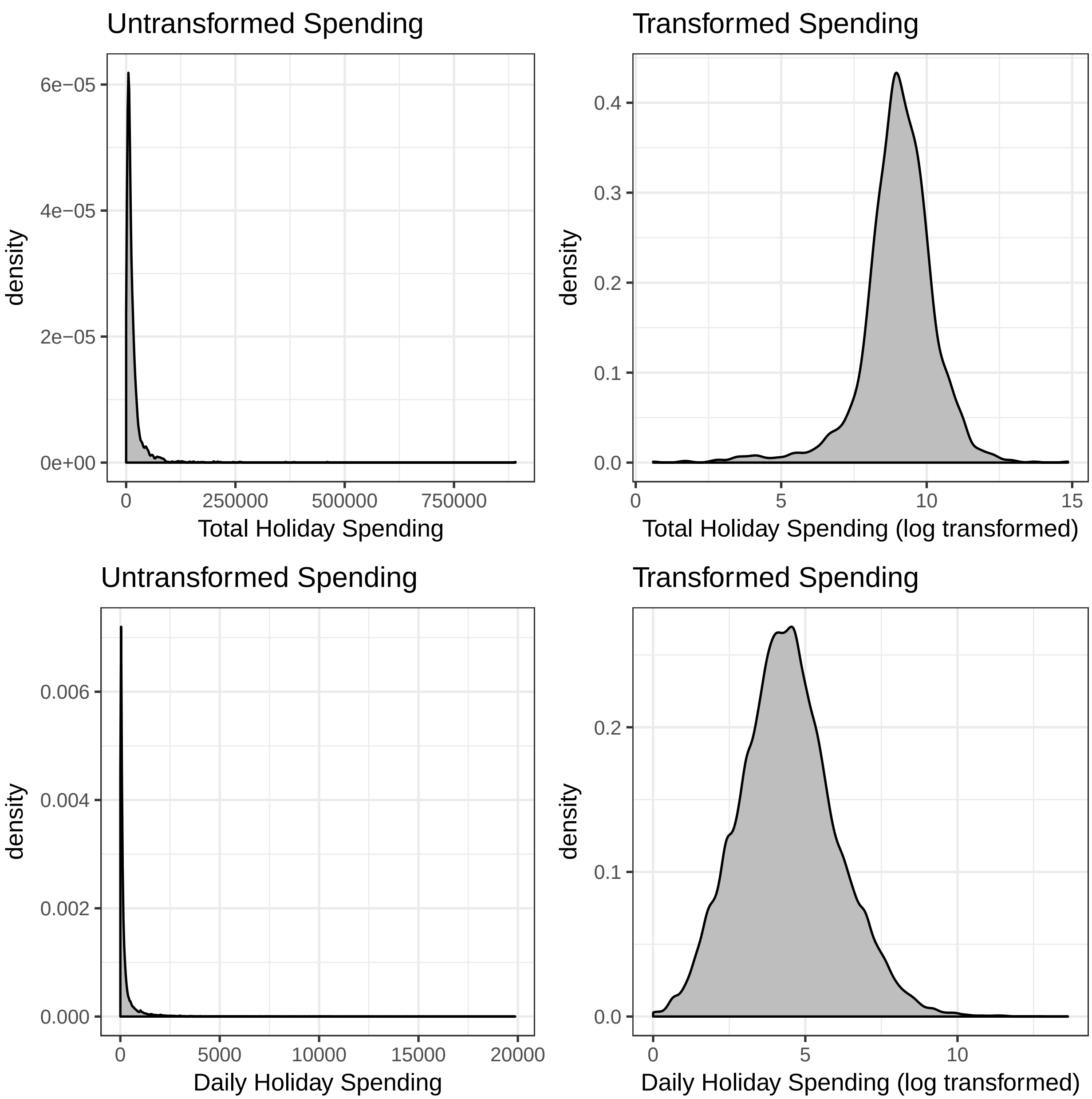
Table 4: Descriptive statistics

```
IQR(holiday_wide$est_bi_credit)

## [1] 1357.639

IQR(holiday_wide$est_bi_spend)

## [1] 15614.29
```

2.2 Log transformation of income and spending

Recall that spending and income variables were log-transformed. Here we plot the raw and transformed values, to demonstrate the need for transformation.

```
#pdf("figure_density.pdf", height = 4.27, width = 6.26)
p1 <- ggplot(subset(holiday_wide, NovDec_spending < 1000000), aes(x = NovDec_spending)) +
  geom_density(fill = "grey") +
  scale_x_continuous("Total Holiday Spending") +
  ggtitle("Untransformed Spending", ) +
  theme_bw()
p2 <- ggplot(holiday_wide, aes(x = log_NovDec)) +
  geom_density(fill = "grey") +
  scale_x_continuous("Total Holiday Spending (log transformed)") +
  ggtitle("Transformed Spending") +
  theme_bw()
p3 <- ggplot(subset(holiday_novdec, amount < 20000), aes(x = amount)) +
  geom_density(fill = "grey") +
  scale_x_continuous("Daily Holiday Spending") +
  ggtitle("Untransformed Spending") +
  theme_bw()
p4 <- ggplot(holiday_novdec, aes(x = log_amount)) +
  geom_density(fill = "grey") +
  scale_x_continuous("Daily Holiday Spending (log transformed)") +
  ggtitle("Transformed Spending") +
  theme_bw()

grid.arrange(p1, p2, p3, p4, ncol = 2)
```




```
#dev.off()
```

2.3 Correlations between personality variables and spending

We examine the correlations between the latent personality trait scores and tranformed holiday spending. These raw associations were reported in the manuscript. To do so, we run a correlation test and save the result to an object. These results are formatted and printed in the paragraph below. We also create a correlation table, which is saved to a word document. This is copied into the manuscript.

```
cor.e <- cor.test(holiday_wide$log_NovDec, holiday_wide$latent.e, use="pairwise")
cor.a <- cor.test(holiday_wide$log_NovDec, holiday_wide$latent.a, use="pairwise")
cor.c <- cor.test(holiday_wide$log_NovDec, holiday_wide$latent.c, use="pairwise")
```

```
cor.n <- cor.test(holiday_wide$log_NovDec, holiday_wide$latent.n, use="pairwise")
cor.o <- cor.test(holiday_wide$log_NovDec, holiday_wide$latent.o, use="pairwise")


cor.table <- holiday_wide %>%
  select(age, male, log_credit, log_spend, log_NovDec,
         latent.e, latent.a, latent.c, latent.n, latent.o) %>%
  rename(income = log_credit,
         spending = log_spend,
         holiday = log_NovDec,
         extra = latent.e,
         agree = latent.a,
         neur = latent.c,
         con = latent.n,
         open = latent.o) %>%
  apa.cor.table(filename = "cortable.doc")
```

Total holiday spending (log-transformed) was positively associated with extraversion ($r = .03$, 95% CI $[-.02, .07]$, $t(1,933) = 1.12$, $p = .262$), conscientiousness ($r = .10$, 95% CI $[.06, .15]$, $t(1,933) = 4.48$, $p < .001$) and negatively associated with neuroticism ($r = -.10$, 95% CI $[-.14, -.05]$, $t(1,933) = -4.20$, $p < .001$). Holiday spending was not found to be associated with agreeableness ($r = -.06$, 95% CI $[-.10, -.01]$, $t(1,933) = -2.58$, $p = .010$) or openness ($r = .00$, 95% CI $[-.04, .05]$, $t(1,933) = 0.03$, $p = .977$).

We standardize the latent variable scores for easier interpretation of results.

```
holiday_wide = holiday_wide %>%
  mutate(z.latent.e = as.numeric(scale(latent.e)),
         z.latent.a = as.numeric(scale(latent.a)),
         z.latent.c = as.numeric(scale(latent.c)),
         z.latent.n = as.numeric(scale(latent.n)),
         z.latent.o = as.numeric(scale(latent.o)))

#add these scores to the long-form data set

holiday_novdec = holiday_wide %>%
  dplyr::select(userreference, z.latent.e, z.latent.a, z.latent.c, z.latent.n, z.latent.o) %>%
  full_join(holiday_novdec)
```

3 Published Results

The results of these analyses were published in the manuscript.

For these analyses, we log-transformed spending and income. We also used estimates of credit and spending covariates based on days outside the holiday season. Latent personality trait scores were estimated from a CFA model. The results of this model are displayed in Table 5.

```
pr.lm <- lm(log(NovDec_spending) ~ z_age + male +
              log_credit + log_spend +
              z.latent.e + z.latent.a + z.latent.c + z.latent.n + z.latent.o,
             data = holiday_wide)
```

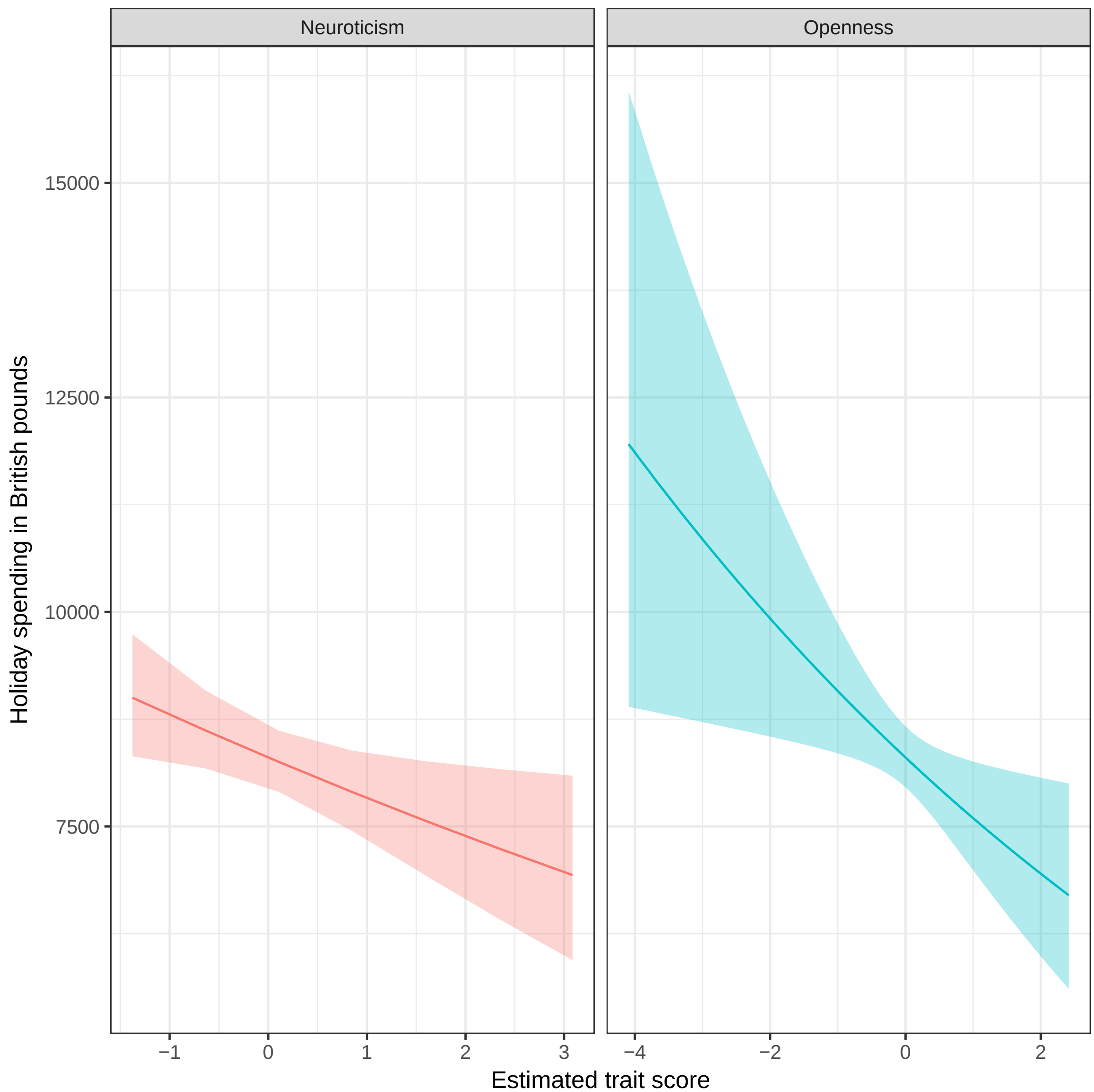
```
stargazer(pr.lm,
          title = "Do personality traits predict holiday spending? (published results)",
          intercept.bottom = F, intercept.top=T,
          ci = TRUE, digits=2, header = FALSE, label = "pr.tab",
          star.cutoffs = c(.05,.01,.001))
```

We create the plot for neuroticism and openness here.

```
n.est = plot_model(pr.lm, type = "pred", terms = "z.latent.n")
o.est = plot_model(pr.lm, type = "pred", terms = "z.latent.o")

n.est = n.est$data
n.est$trait = "Neuroticism"
o.est = o.est$data
o.est$trait = "Openness"

n.est %>%
  full_join(o.est) %>%
  ggplot(aes(x,exp(predicted))) +
  geom_ribbon(aes(ymin=exp(conf.low), ymax=exp(conf.high),
                  fill = trait),
              alpha = .3) +
  geom_line(aes(color = trait)) +
  scale_x_continuous("Estimated trait score") +
  scale_y_continuous("Holiday spending in British pounds") +
  guides(fill = FALSE, color = FALSE) +
  facet_wrap(~trait, scales = "free_x") +
  theme_bw()

## Joining, by = c("x", "predicted", "conf.low", "conf.high", "group", "trait")
```

	Dependent variable:
	log(NovDec_spending)
Constant	1.73***
	(1.29, 2.17)
z_age	0.01
	(−0.04, 0.05)
male	0.12**
	(0.03, 0.21)
log_credit	0.05***
	(0.04, 0.07)
log_spend	0.73***
	(0.69, 0.78)
z.latent.e	0.04
	(−0.02, 0.11)
z.latent.a	−0.03
	(−0.08, 0.02)
z.latent.c	0.05*
	(0.004, 0.10)
z.latent.n	−0.06*
	(−0.11, −0.01)
z.latent.o	−0.09*
	(−0.16, −0.02)
Observations	1,935
R^2	0.43
Adjusted R^2	0.43
Residual Std. Error	0.96 (df = 1925)
F Statistic	162.45*** (df = 9; 1925)
Note:	*p<0.05; **p<0.01; ***p<0.001

Table 5: Do personality traits predict holiday spending? (published results)




```
ggsave("estimatedspending.pdf", height = 4, width = 7)
```

4 Longitudinal Analyses

These analyses were briefly discussed in the published manuscript. Here, we provide greater detail and additional plots for interpretation. Resuts of the longitudinal model are displayed in Table 6.

```
# estimation parameters
nlopt <- function(par, fn, lower, upper, control) {
  .nloptr <<- res <- nloptr(par, fn, lb = lower, ub = upper,
                            opts = list
                            (algorithm = "NLOPT_LN_BOBYQA",
                                        print_level = 1,
                                        maxeval = 1000, xtol_abs = 1e-6,
                                        ftol_abs = 1e-6))
  list(par = res$solution,
       fval = res$objective,
       conv = if (res$status > 0) 0 else res$status,
       message = res$message
  )}
```

```
pr.lmer <- lmer(log(amount + 1) ~ days + z_age + male + log_credit + log_spend +
                 z.latent.e + z.latent.a + z.latent.c + z.latent.n + z.latent.o +
                  days*z_age + days*male + days*log_credit + days*log_spend +
                  days*z.latent.e + days*z.latent.a + days*z.latent.c +
                    days*z.latent.n + days*z.latent.o +
                  (1 + days |userreference),
                 data = holiday_novdec,
                 control=lmerControl(optimizer = "nloptwrap", calc.derivs = FALSE))
```

```
stargazer(pr.lmer,
          title = "Do personality traits predict holiday spending trajectories? (published results)",
          font.size = "footnotesize", single.row = TRUE, header = FALSE,
          label = "pr.long",
          intercept.bottom = F, intercept.top=T, ci = TRUE, digits=2, no.space = T,
          star.cutoffs = c(.05,.01,.001))
```

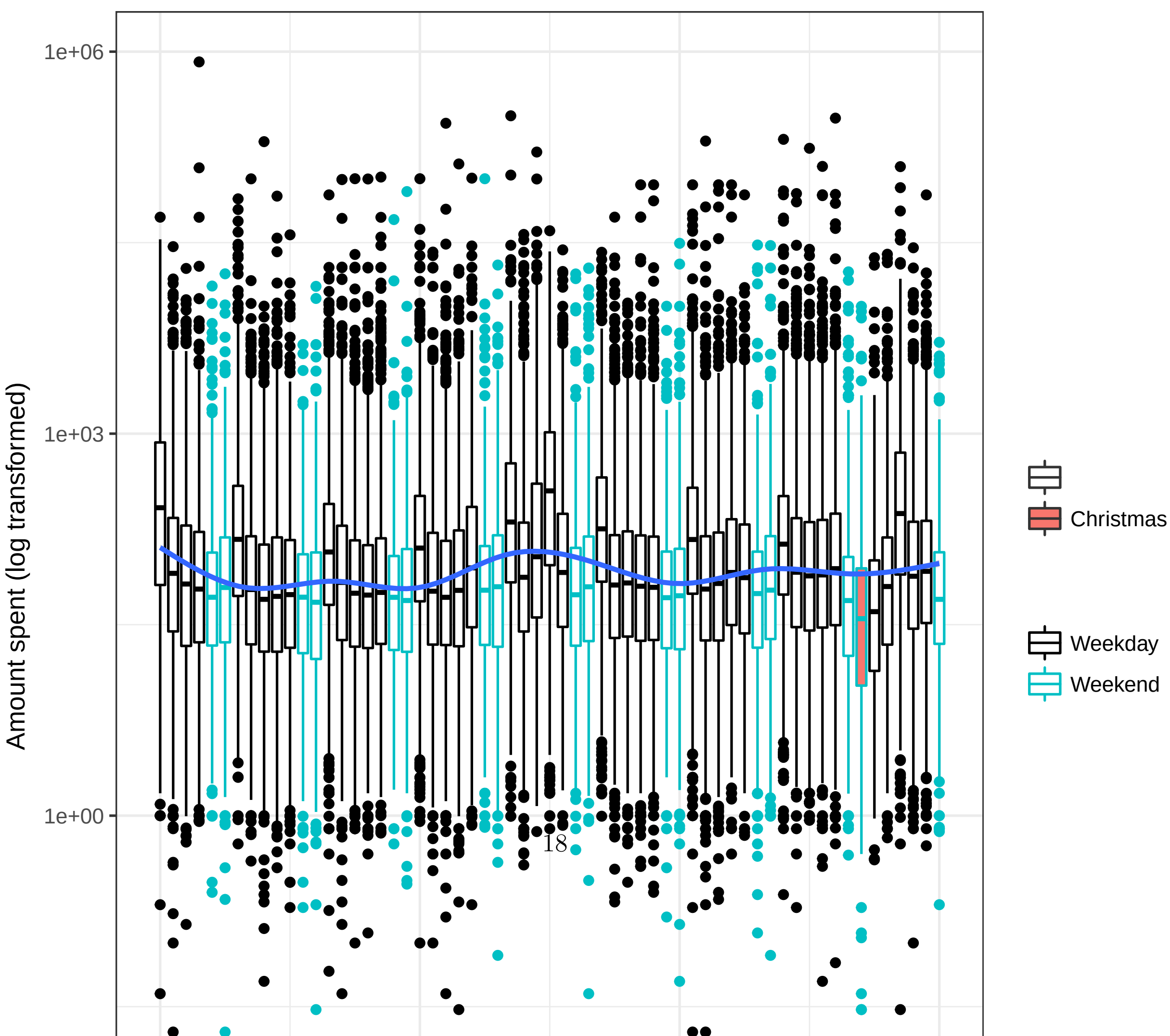
4.1 Plots

First we plot summaries of amount spent across the holiday season in Figure ??.

```
holiday_novdec %>%
  mutate(weekend = ifelse(grepl("^S", weekdays(transactiondate)), 1, 0)) %>%
  mutate(Christmas = ifelse(transactiondate == "2016-12-25", "Christmas", "")) %>%
  ggplot(aes(x = days, y = amount, color = as.factor(weekend), fill = Christmas)) +
  geom_boxplot(aes(group = cut_width(days, 1))) +
  geom_smooth(aes(x = days, y = amount), inherit.aes = F) +
  scale_y_log10("Amount spent (log transformed)") +
  scale_x_continuous("Day of holiday season") +
  scale_color_manual("", labels = c("Weekday", "Weekend"), values = c("black", "#00BFC4")) +
  scale_fill_manual("", labels = c("", "Christmas"), values = c("white", "#F8766D")) +
  theme_bw()
```

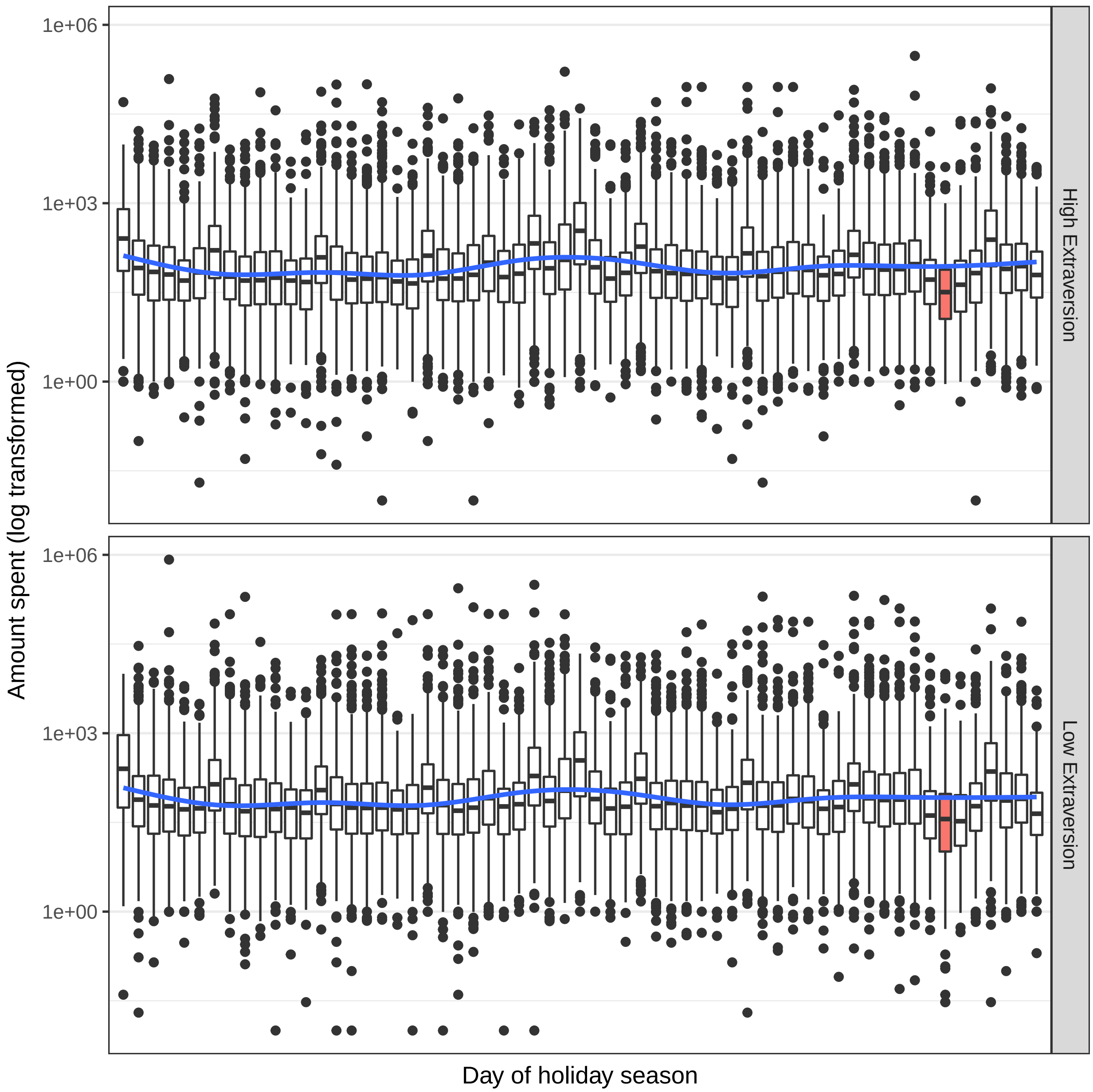
	Dependent variable:
	log(amount + 1)
Constant	−0.10 (−0.44, 0.25)
days	0.02*** (0.01, 0.03)
z_age	0.12*** (0.08, 0.15)
male	0.03 (−0.03, 0.09)
log_credit	0.03*** (0.02, 0.04)
log_spend	0.45*** (0.41, 0.49)
z.latent.e	0.02 (−0.03, 0.06)
z.latent.a	−0.02 (−0.05, 0.02)
z.latent.c	0.02 (−0.01, 0.05)
z.latent.n	−0.01 (−0.04, 0.02)
z.latent.o	−0.01 (−0.06, 0.04)
days:z_age	−0.0001 (−0.001, 0.001)
days:male	−0.0004 (−0.002, 0.001)
days:log_credit	−0.0004** (−0.001, −0.0001)
days:log_spend	−0.001*** (−0.002, −0.001)
days:z.latent.e	0.0004 (−0.001, 0.001)
days:z.latent.a	−0.0000 (−0.001, 0.001)
days:z.latent.c	0.0000 (−0.001, 0.001)
days:z.latent.n	0.0001 (−0.001, 0.001)
days:z.latent.o	0.0001 (−0.001, 0.001)
Observations	78,539
Log Likelihood	−144,210.00
Akaike Inf. Crit.	288,467.90
Bayesian Inf. Crit.	288,690.40
Note:	*p<0.05; **p<0.01; ***p<0.001

Table 6: Do personality traits predict holiday spending trajectories? (published results)

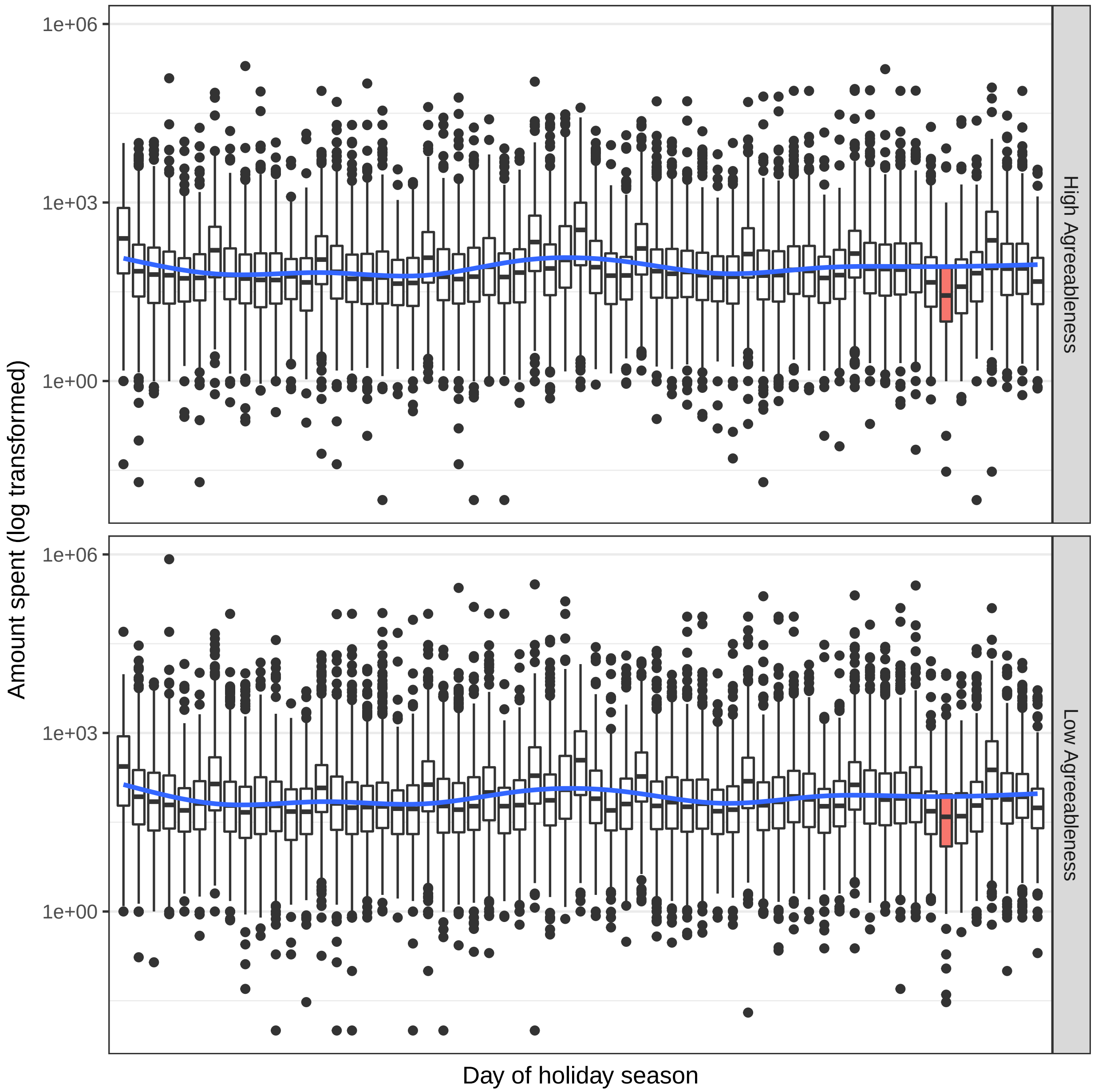


```r
ggsave(filename = "spending_boxplots.pdf", width = 10, height = 7.5)
```

```r
holiday_novdec %>%
  filter(!is.na(z.latent.e)) %>%
  mutate(extra.group = ifelse(z.latent.e >0, "High Extraversion", "Low Extraversion")) %>%
  mutate(Christmas = ifelse(transactiondate == "2016-12-25", "Christmas", "")) %>%
  ggplot(aes(x = days, y = amount, fill = Christmas)) +
  geom_boxplot(aes(group = cut_width(days, 1))) +
  geom_smooth(se = F) +
  scale_y_log10("Amount spent (log transformed)") +
  scale_x_discrete("Day of holiday season") +
  scale_fill_manual(values = c("white", "#F8766D")) +
  theme_bw()+
  theme(legend.position = "none")+
  facet_grid(extra.group ~ .)
```

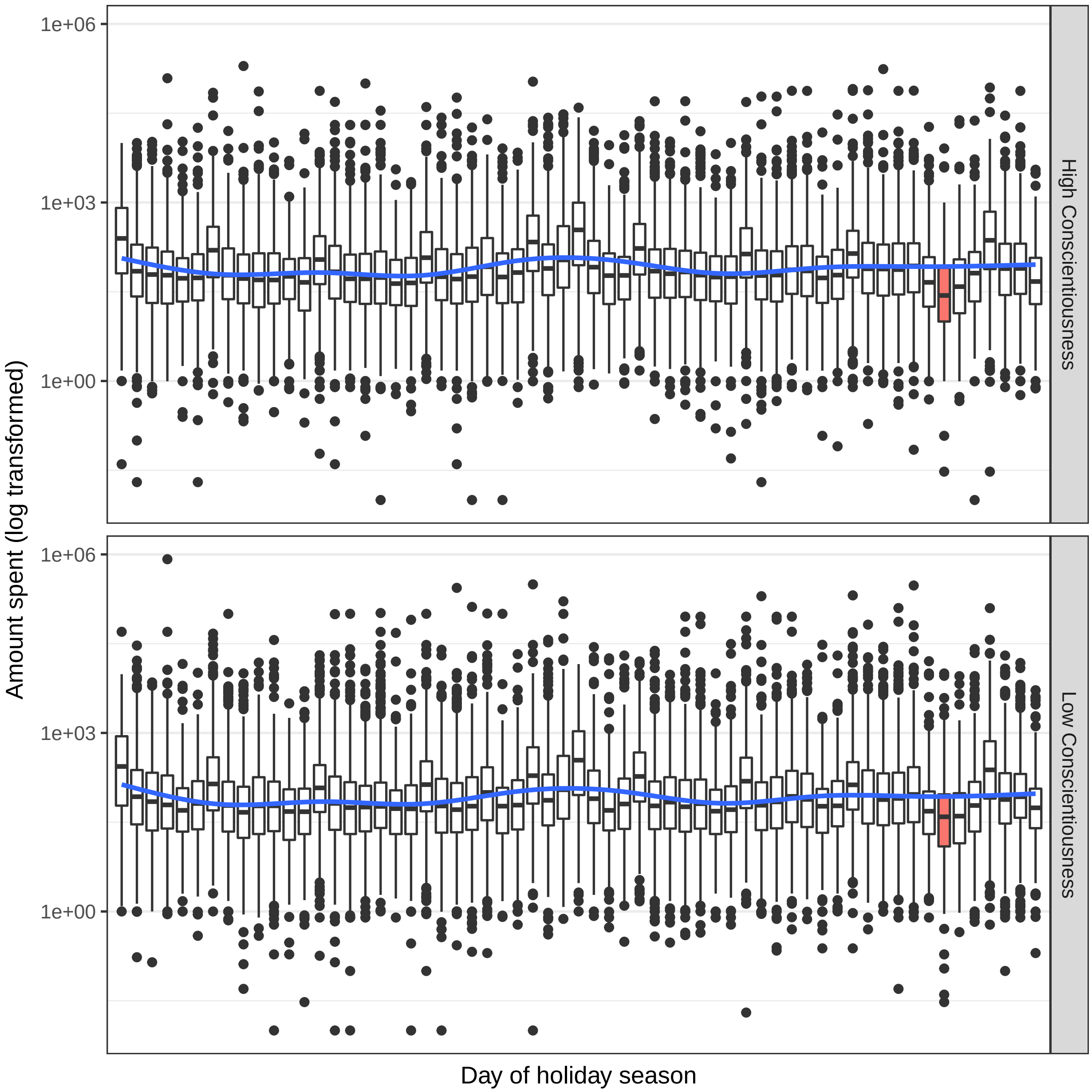


```
holiday_novdec %>%
  filter(!is.na(z.latent.a)) %>%
  mutate(agree.group = ifelse(z.latent.a >0, "High Agreeableness", "Low Agreeableness")) %>%
  mutate(Christmas = ifelse(transactiondate == "2016-12-25", "Christmas", "")) %>%
  ggplot(aes(x = days, y = amount, fill = Christmas)) +
  geom_boxplot(aes(group = cut_width(days, 1))) +
  geom_smooth(se = F) +
  scale_y_log10("Amount spent (log transformed)") +
  scale_x_discrete("Day of holiday season") +
  scale_fill_manual(values = c("white", "#F8766D")) +
  theme_bw()+
  theme(legend.position = "none")+
  facet_grid(agree.group ~ .)
```



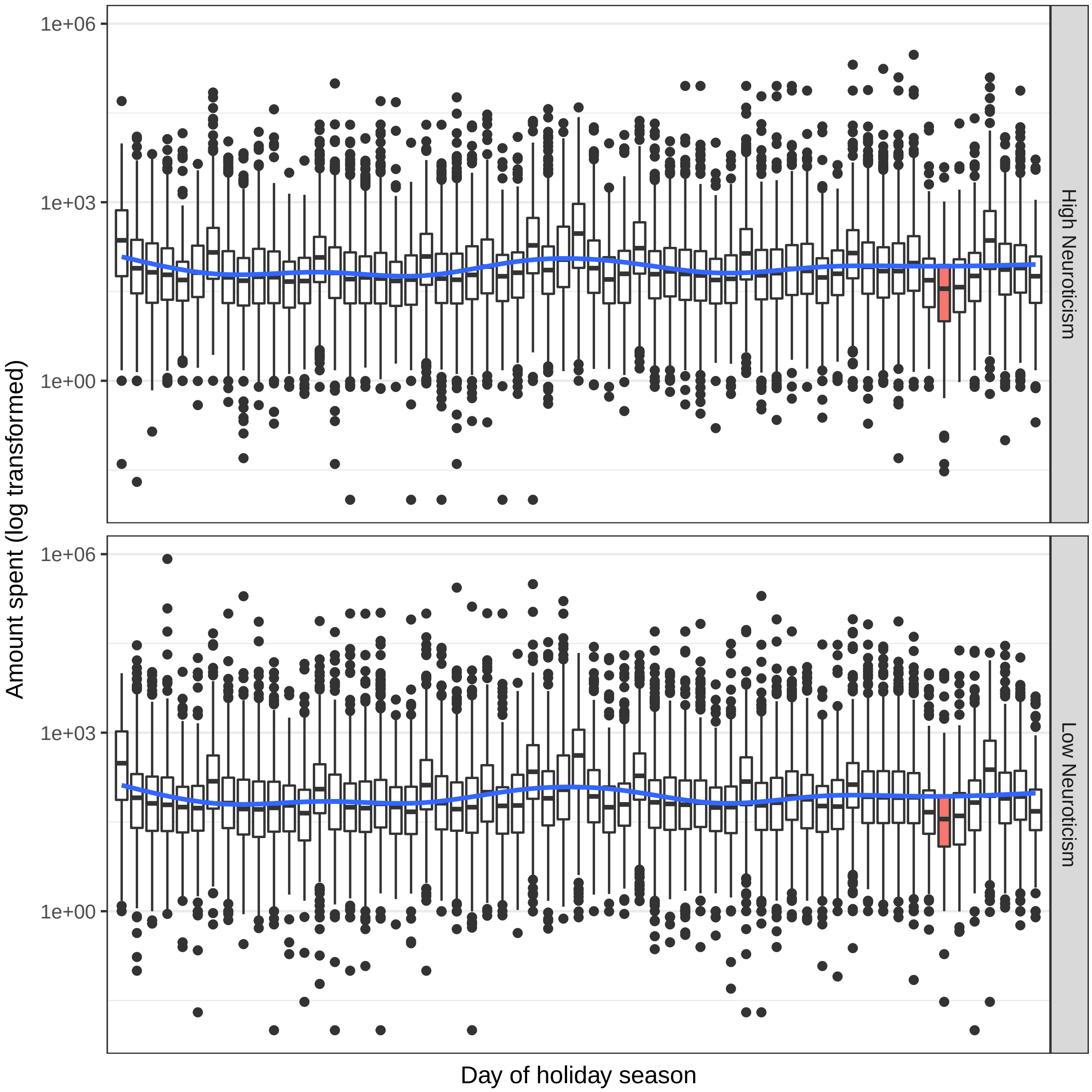

```
holiday_novdec %>%
  filter(!is.na(z.latent.c)) %>%
```

```
  mutate(con.group = ifelse(z.latent.a >0, "High Conscientiousness", "Low Conscientiousness")) %>%
  mutate(Christmas = ifelse(transactiondate == "2016-12-25", "Christmas", "")) %>%
  ggplot(aes(x = days, y = amount, fill = Christmas)) +
  geom_boxplot(aes(group = cut_width(days, 1))) +
  geom_smooth(se = F) +
  scale_y_log10("Amount spent (log transformed)") +
  scale_x_discrete("Day of holiday season") +
  scale_fill_manual(values = c("white", "#F8766D")) +
  theme_bw()+
  theme(legend.position = "none")+
  facet_grid(con.group ~ .)
```



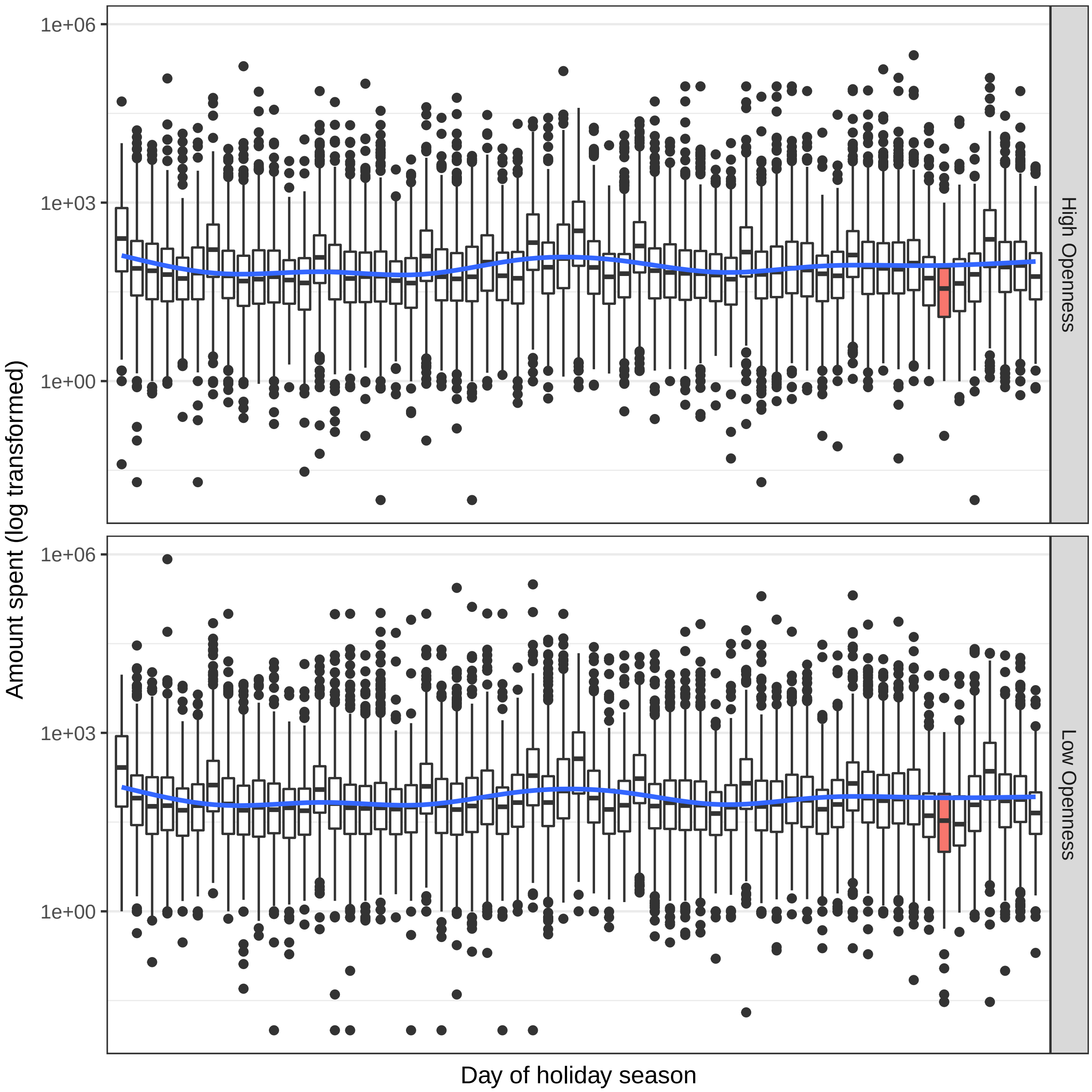

```
holiday_novdec %>%
  filter(!is.na(z.latent.n)) %>%
  mutate(neur.group = ifelse(z.latent.n >0, "High Neuroticism", "Low Neuroticism")) %>%
  mutate(Christmas = ifelse(transactiondate == "2016-12-25", "Christmas", "")) %>%
```

```
ggplot(aes(x = days, y = amount, fill = Christmas)) +
geom_boxplot(aes(group = cut_width(days, 1))) +
geom_smooth(se = F) +
scale_y_log10("Amount spent (log transformed)") +
scale_x_discrete("Day of holiday season") +
scale_fill_manual(values = c("white", "#F8766D")) +
theme_bw()+
theme(legend.position = "none")+
facet_grid(neur.group ~ .)
```




```
holiday_novdec %>%
  filter(!is.na(z.latent.o)) %>%
  mutate(open.group = ifelse(z.latent.o >0, "High Openness", "Low Openness")) %>%
  mutate(Christmas = ifelse(transactiondate == "2016-12-25", "Christmas", "")) %>%
  ggplot(aes(x = days, y = amount, fill = Christmas)) +
  geom_boxplot(aes(group = cut_width(days, 1))) +
```

```
  geom_smooth(se = F) +
  scale_y_log10("Amount spent (log transformed)") +
  scale_x_discrete("Day of holiday season") +
  scale_fill_manual(values = c("white", "#F8766D")) +
  theme_bw()+
  theme(legend.position = "none")+
  facet_grid(open.group ~ .)
```



5 Reviewed Results

The results of these analyses were submitted to the journal and peer reviewed. Parts of the analyses were adjusted, per the suggestions of the reviewers and editor (those changes are in the seciton 3). These are the original results.

For these analyses, we log-transformed spending and income. We also used estimates of credit and spending covariates based on days outside the holiday season. The results of the main analyses are shown in Table 7. The results of the longitudinal analyses are in Table ??

```
rr.lm <- lm(log(NovDec_spending) ~ z_age + male + log_credit +
                    log_spend + z_extra + z_agree + z_con + z_neur + z_open,
             data = holiday_wide)
```

```
stargazer(rr.lm,
          title = "Do personality traits predict holiday spending? (peer-reviewed)",
          intercept.bottom = F, intercept.top=T,
          label = "rr.main",
          ci = TRUE, digits=2, header = FALSE,
          star.cutoffs = c(.05,.01,.001))
```

```
ex3.lmer <- lmer(log(amount + 1) ~ days + z_age + male + log_credit + log_spend +
                  z_extra + z_agree + z_con + z_neur + z_open +
                  days*z_age + days*male + days*log_credit + days*log_spend +
                  days*z_extra + days*z_agree + days*z_con + days*z_neur + days*z_open +
                  (1 + days |userreference),
                 data = holiday_novdec,
                 control=lmerControl(optimizer = "nloptwrap", calc.derivs = FALSE))
```

```
stargazer(ex3.lmer,
          title = "Do personality traits predict holiday spending trajectories? (exploratory, new covar
          label = "rr.long",
          font.size = "footnotesize", single.row = TRUE, header = FALSE,
          intercept.bottom = F, intercept.top=T, ci = TRUE, digits=2, no.space = T,
          star.cutoffs = c(.05,.01,.001))
```

	Dependent variable:
	log(NovDec_spending)
Constant	1.81^{***}
	(1.39, 2.23)
z_age	-0.004
	(-0.05, 0.04)
male	0.11^{*}
	(0.02, 0.19)
log_credit	0.06^{***}
	(0.04, 0.07)
log_spend	0.72^{***}
	(0.68, 0.77)
z_extra	0.02
	(-0.02, 0.06)
z_agree	-0.02
	(-0.06, 0.02)
z_con	0.04^{*}
	(0.0005, 0.09)
z_neur	-0.06^{**}
	(-0.10, -0.01)
z_open	-0.08^{***}
	(-0.12, -0.03)
Observations	2,118
R^2	0.43
Adjusted R^2	0.43
Residual Std. Error	0.96 (df = 2108)
F Statistic	176.84^{***} (df = 9; 2108)
Note:	$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table 7: Do personality traits predict holiday spending? (peer-reviewed)

	Dependent variable:
	log(amount + 1)
Constant	−0.14 (−0.47, 0.18)
days	0.02*** (0.01, 0.02)
z_age	0.11*** (0.08, 0.14)
male	0.01 (−0.05, 0.07)
log_credit	0.02*** (0.01, 0.03)
log_spend	0.45*** (0.42, 0.49)
z_extra	0.02 (−0.01, 0.05)
z_agree	−0.005 (−0.03, 0.02)
z_con	0.01 (−0.02, 0.04)
z_neur	−0.02 (−0.05, 0.01)
z_open	−0.02 (−0.05, 0.01)
days:z_age	−0.0002 (−0.001, 0.001)
days:male	−0.0001 (−0.001, 0.001)
days:log_credit	−0.0004*** (−0.001, −0.0002)
days:log_spend	−0.001*** (−0.002, −0.001)
days:z_extra	0.0004 (−0.0003, 0.001)
days:z_agree	0.0000 (−0.001, 0.001)
days:z_con	0.0000 (−0.001, 0.001)
days:z_neur	0.0002 (−0.0005, 0.001)
days:z_open	0.0001 (−0.001, 0.001)
Observations	86,191
Log Likelihood	−157,926.50
Akaike Inf. Crit.	315,900.90
Bayesian Inf. Crit.	316,125.60
Note:	*p<0.05; **p<0.01; ***p<0.001

Table 8: Do personality traits predict holiday spending trajectories? (exploratory, new covariates and log transformed)

6 Pre-registered Analyses

These are the results of the analyses that were originally pre-registered. We deviated from the preregistration for a few reasons. First, we did not consider that the spending and income variables would be highly skewed. Second, we were unaware that participants did not participate in the study for equal amounts of time. These results are not included in the published manuscript.

6.1 Question 1: Does personality predict the amount of spending during the holidays?

We calculate a standardized credit variable, representing total income.

```
holiday_wide = holiday_wide %>%
  mutate(z_credit = as.numeric(scale(total_credit)))

holiday_novdec = holiday_wide %>%
  dplyr::select(userreference, z_credit) %>%
  full_join(holiday_novdec)

## Joining, by = "userreference"
```

We also created a variable to represent the ratio of spending during the holiday season to the whole year's spending for each person. This was used in beta regression models.

```
holiday_wide = holiday_wide %>%
  mutate(
```

```
    # calculate ratio of yearly spending to yearly income
    yearlyspending_ratio = yearly_spending/total_credit,
    # calculate ratio of Nov/Dec spending to total yearly spending
    NovDecspending_ratio = NovDec_spending/yearly_spending)
```

We predict total amount of spending during the holiday season by personality traits (average scores), age, gender, income (credit) and spending. We use both a linear model and a beta-regression model. Results of these models are in Table 9.

```
prereg.q1.lm <- lm(NovDec_spending ~ z_age + male + z_credit + z_yearlyspending +
                    z_extra + z_agree + z_con + z_neur + z_open,
                  data = holiday_wide)
```

```
prereg.q1.br <- betareg(NovDecspending_ratio ~ z_age + male + z_credit + z_yearlyspending +
                  z_extra + z_agree + z_con + z_neur + z_open | 1,
                data = holiday_wide)
```

```
stargazer(prereg.q1.lm, prereg.q1.br,
          title = "Do personality traits predict holiday spending? (pre-registered)",
          intercept.bottom = F, intercept.top=T,
          ci = TRUE, digits=2, header = FALSE, label = "pr.tab1",
          star.cutoffs = c(.05,.01,.001))
```

6.2 Question 2: Does personailty predict changes in spending across the holiday season?

Next we examine how personality traits may predict changes in spending. Results of this model are shown in Table 10.

```
q2.lmer <- lmer(amount ~ days + z_age + male + z_credit + z_yearlyspending +
                  z_extra + z_agree + z_con + z_neur + z_open +
                  days*z_age + days*male + days*z_credit + days*z_yearlyspending +
                  days*z_extra + days*z_agree + days*z_con + days*z_neur + days*z_open +
                  (1 + days |userreference),
                data = holiday_novdec,
                control=lmerControl(optimizer = "nloptwrap", calc.derivs = FALSE))
```

```
stargazer(q2.lmer, title = "Do personality traits predict holiday spending trajectories? (pre-registere
          font.size = "footnotesize", header = FALSE,
          intercept.bottom = F, intercept.top=T, ci = TRUE, digits=2, no.space = T,
          label = "pr.tab2",
          star.cutoffs = c(.05,.01,.001))
```

Here we plot the effect of extraversion on the slope.

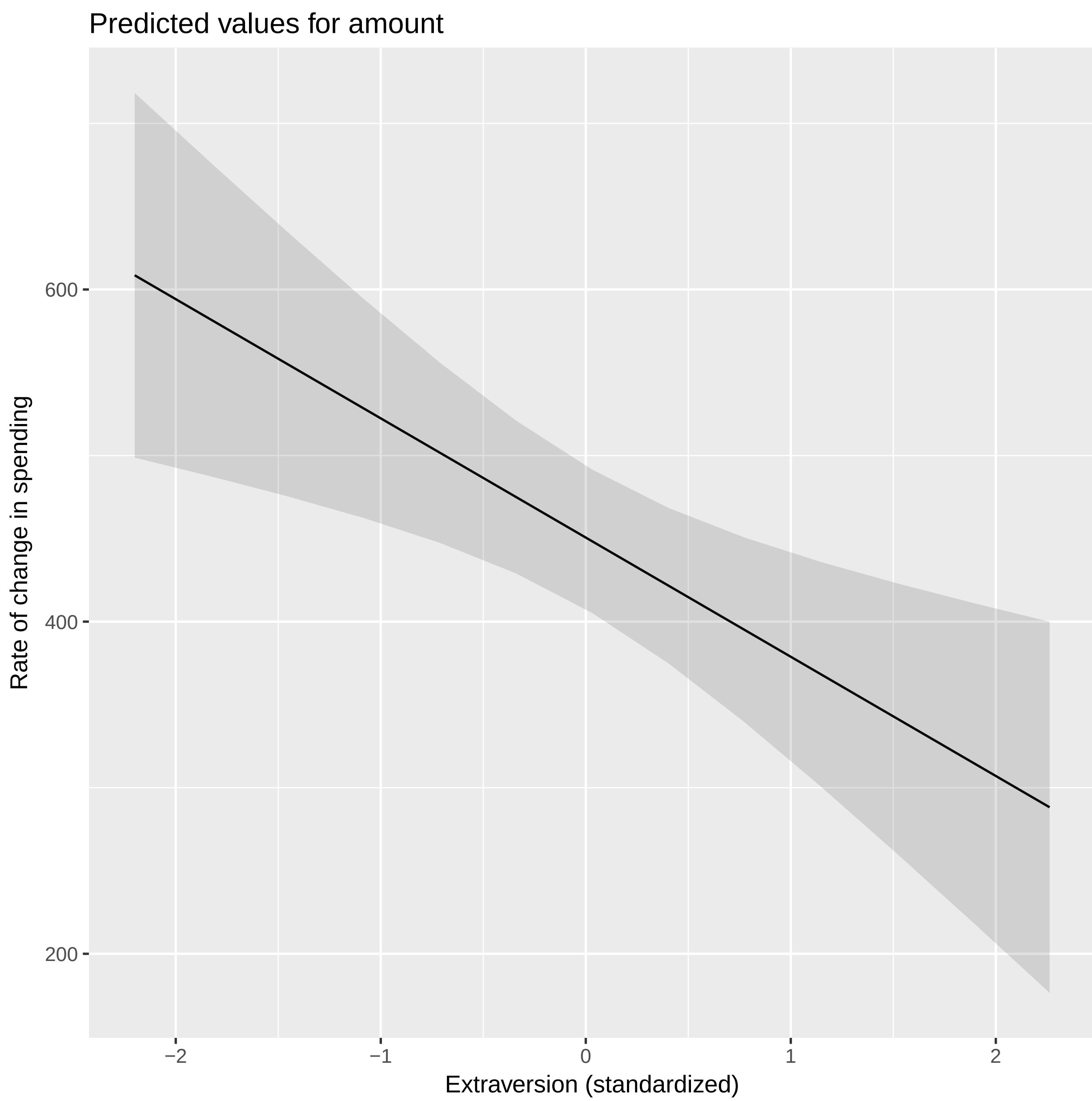
```
plot_model(q2.lmer, type = "pred", term = "z_extra",
          axis.title = c("Extraversion (standardized)", "Rate of change in spending"))
```

	Dependent variable:	
	NovDec_spending	NovDecspending_ratio
	OLS	*beta*
	(1)	(2)
Constant	18,250.86***	−1.58***
	(15,367.66, 21,134.07)	(−1.62, −1.55)
z_age	−3,032.66**	−0.03*
	(−5,234.36, −830.97)	(−0.06, −0.001)
male	−1,623.18	0.003
	(−5,985.76, 2,739.40)	(−0.05, 0.06)
z_credit	−1,512.16	0.01
	(−3,729.08, 704.76)	(−0.02, 0.04)
z_yearlyspending	49,326.78***	−0.03*
	(47,077.47, 51,576.08)	(−0.06, −0.0003)
z_extra	−3,653.67**	−0.003
	(−5,903.09, −1,404.26)	(−0.03, 0.02)
z_agree	742.32	0.01
	(−1,443.18, 2,927.82)	(−0.02, 0.04)
z_con	−185.62	0.02
	(−2,425.08, 2,053.85)	(−0.01, 0.05)
z_neur	102.44	−0.02
	(−2,170.56, 2,375.44)	(−0.05, 0.01)
z_open	92.84	−0.02
	(−2,155.30, 2,340.98)	(−0.05, 0.01)
Observations	2,118	2,118
R^2	0.48	0.01
Adjusted R^2	0.48	
Log Likelihood		2,245.94
Residual Std. Error	49,902.12 (df = 2108)	
F Statistic	220.58*** (df = 9; 2108)	
Note:		*p<0.05; **p<0.01; ***p<0.001

Table 9: Do personality traits predict holiday spending? (pre-registered)

	Dependent variable:
	amount
Constant	383.53***
	(261.28, 505.78)
days	0.57
	(−2.51, 3.65)
z_age	−105.04*
	(−199.36, −10.72)
male	−63.83
	(−246.63, 118.97)
z_credit	87.28
	(−1.31, 175.88)
z_yearlyspending	1,216.77***
	(1,126.41, 1,307.12)
z_extra	−167.71***
	(−262.56, −72.86)
z_agree	29.79
	(−62.01, 121.59)
z_con	45.80
	(−48.60, 140.21)
z_neur	−16.70
	(−112.50, 79.11)
z_open	9.89
	(−84.79, 104.57)
days:z_age	1.96
	(−0.42, 4.33)
days:male	0.39
	(−4.20, 4.98)
days:z_credit	−3.82***
	(−6.03, −1.62)
days:z_yearlyspending	−9.87***
	(−12.10, −7.64)
days:z_extra	3.28**
	(0.90, 5.66)
days:z_agree	−0.52
	(−2.82, 1.78)
days:z_con	−1.42
	(−3.79, 0.95)
days:z_neur	0.86
	(−1.55, 3.26)
days:z_open	−0.20
	(−2.58, 2.18)
Observations	86,191
Log Likelihood	−840,693.80
Akaike Inf. Crit.	1,681,436.00
Bayesian Inf. Crit.	1,681,660.00
Note:	*p<0.05; **p<0.01; ***p<0.001

Table 10: Do personality traits predict holiday spending trajectories? (pre-registered)


Predicted values for amount
600
400
200
Rate of change in spending
−2
−1
0
1
2
Extraversion (standardized)

7 Exploratory analyses

After running the pre-registered analyses, we were troubled by some of the results/issues with the data. We tested a few methods of re-analyzing the data, which are documented here. These results are not included in the published manuscript.

7.1 Recalculated covariates

After running the preregistered analyses, it became clear that not all participants had provided data for the full year. Above, we calculated adjusted credit and spending amounts by calculating the average daily credit and debit and multiplying to estimate bi-monthly credit and spending. Here we use those values as covariates, instead of the 'z_credit' and 'z_yearlyspending' variables we preregistered. Results of these models are in Tables 11 and 12.

```
ex1.lm <- lm(NovDec_spending ~ z_age + male + z_est_bi_credit + z_est_bi_spend +
              z_extra + z_agree + z_con + z_neur + z_open,
            data = holiday_wide)
```

```
ex1.br <- betareg(NovDecspending_ratio ~ z_age + male + z_est_bi_credit + z_est_bi_spend +
                   z_extra + z_agree + z_con + z_neur + z_open | 1,
                 data = holiday_wide)
```

```
stargazer(ex1.lm, ex1.br,
          title = "Do personality traits predict holiday spending? (exploratory, new covariates)",
          intercept.bottom = F, intercept.top=T, ci = TRUE, digits=2, header = FALSE, label = "ex1:tab"
          star.cutoffs = c(.05,.01,.001))
```

7.2 Research Question 2

```
ex1.lmer <- lmer(amount ~ days + z_age + male + z_est_bi_credit + z_est_bi_spend +
                  z_extra + z_agree + z_con + z_neur + z_open +
                  days*z_age + days*male + days*z_est_bi_credit + days*z_est_bi_spend +
                  days*z_extra + days*z_agree + days*z_con + days*z_neur + days*z_open +
                  (1 + days |userreference),
                 data = holiday_novdec,
                 control=lmerControl(optimizer = "nloptwrap", calc.derivs = FALSE))
```

```
stargazer(ex1.lmer, title = "Do personality traits predict holiday spending trajectories? (exploratory,
          font.size = "footnotesize", header = FALSE,label = "ex1:tab2",
          intercept.bottom = F, intercept.top=T, ci = TRUE, digits=2, no.space = T,
          star.cutoffs = c(.05,.01,.001))
```

7.3 Fractional binomial

We also thought that perhaps our ratio outcomes were better served by using a fractional binomial model, instead of beta regression. We attempted these models in Table 13.

```
ex2.lm1 <- glm(NovDecspending_ratio ~ z_age + male + z_credit + z_yearlyspending +
              z_extra + z_agree + z_con + z_neur + z_open,
            data = holiday_wide, family = quasibinomial(link=logit))
ex2.lm2 <- glm(NovDecspending_ratio ~ z_age + male + z_est_bi_credit + z_est_bi_spend +
              z_extra + z_agree + z_con + z_neur + z_open,
            data = holiday_wide, family = quasibinomial(link=logit))
```

	Dependent variable:	
	NovDec_spending	NovDecspending_ratio
	OLS	*beta*
	(1)	(2)
Constant	16,650.64***	−1.59***
	(12,749.44, 20,551.83)	(−1.63, −1.56)
z_age	3,001.81*	−0.01
	(12.83, 5,990.78)	(−0.03, 0.02)
male	1,912.35	0.01
	(−3,988.75, 7,813.45)	(−0.04, 0.07)
z_est_bi_credit	6,910.28***	−0.02
	(3,957.60, 9,862.97)	(−0.05, 0.01)
z_est_bi_spend	12,096.00***	−0.11***
	(9,095.97, 15,096.03)	(−0.15, −0.08)
z_extra	−2,437.13	0.01
	(−5,488.32, 614.07)	(−0.02, 0.03)
z_agree	−506.90	0.01
	(−3,464.71, 2,450.90)	(−0.02, 0.03)
z_con	1,982.91	0.03*
	(−1,050.31, 5,016.12)	(0.001, 0.05)
z_neur	−1,054.54	−0.01
	(−4,130.34, 2,021.27)	(−0.04, 0.01)
z_open	−577.64	−0.02
	(−3,620.05, 2,464.78)	(−0.05, 0.01)
Observations	2,118	2,118
R^2	0.06	0.04
Adjusted R^2	0.05	
Log Likelihood		2,276.27
Residual Std. Error	67,539.42 (df = 2108)	
F Statistic	14.06*** (df = 9; 2108)	
Note:		*p<0.05; **p<0.01; ***p<0.001

Table 11: Do personality traits predict holiday spending? (exploratory, new covariates)

	Dependent variable:
	amount
Constant	381.36***
	(237.53, 525.20)
days	0.84
	(−2.40, 4.08)
z_age	38.09
	(−73.01, 149.19)
male	15.35
	(−200.13, 230.84)
z_est_bi_credit	306.06***
	(196.71, 415.41)
z_est_bi_spend	387.16***
	(275.76, 498.55)
z_extra	−143.08*
	(−255.08, −31.08)
z_agree	−0.50
	(−108.70, 107.70)
z_con	93.04
	(−18.24, 204.32)
z_neur	−45.87
	(−158.70, 66.96)
z_open	−5.31
	(−116.86, 106.23)
days:z_age	1.01
	(−1.49, 3.52)
days:male	−0.25
	(−5.08, 4.58)
days:z_est_bi_credit	−5.63***
	(−8.06, −3.20)
days:z_est_bi_spend	−3.73**
	(−6.22, −1.24)
days:z_extra	3.11*
	(0.60, 5.62)
days:z_agree	−0.31
	(−2.74, 2.11)
days:z_con	−1.72
	(−4.22, 0.78)
days:z_neur	1.07
	(−1.47, 3.60)
days:z_open	−0.05
	(−2.56, 2.45)
Observations	86,191
Log Likelihood	−841,276.60
Akaike Inf. Crit.	1,682,601.00
Bayesian Inf. Crit.	1,682,826.00
Note:	*p<0.05; **p<0.01; ***p<0.001

Table 12: Do personality traits predict holiday spending trajectories? (exploratory, new covariates)

```r
ex2.glmer1 <- glmer(daily_ratio ~ days + z_age + male + z_credit + z_yearlyspending +
                  z_extra + z_agree + z_con + z_neur + z_open +
                  days*z_age + days*male + days*z_est_bi_credit + days*z_est_bi_spend +
                  days*z_extra + days*z_agree + days*z_con + days*z_neur + days*z_open +
                  (1 + days |userreference), weights = holiday_novdec$NovDec_spending,
                  family = "binomial",
                  data = holiday_novdec,
                  control=glmerControl(optimizer = "nloptwrap", calc.derivs = FALSE))

## Warning in eval(family$initialize, rho):  non-integer #successes in a binomial glm!
```

```r
stargazer(ex2.lm1, ex2.lm2, ex2.glmer1,
          title = "Do personality traits predict holiday spending trajectories? (exploratory, new covari
          font.size = "footnotesize", header = FALSE, label = "ex2:tab",
          intercept.bottom = F, intercept.top=T, ci = TRUE, digits=2, no.space = T,
          star.cutoffs = c(.05,.01,.001))
```

	Dependent variable:		
	NovDecspending_ratio		daily_ratio
	glm: quasibinomial *link = logit*		*generalized linear* *mixed-effects*
	(1)	(2)	(3)
Constant	−1.59***	−1.60***	−3.64***
	(−1.63, −1.55)	(−1.63, −1.56)	(−3.71, −3.56)
days			0.002
			(−0.001, 0.004)
z_age	−0.02	−0.002	0.02
	(−0.05, 0.005)	(−0.03, 0.03)	(−0.04, 0.07)
male	−0.01	−0.004	−0.11*
	(−0.07, 0.05)	(−0.06, 0.05)	(−0.22, −0.01)
z_credit	−0.0003		−0.31***
	(−0.03, 0.03)		(−0.41, −0.20)
z_yearlyspending	0.02		−0.18***
	(−0.01, 0.05)		(−0.22, −0.13)
z_est_bi_credit		0.01	0.34***
		(−0.01, 0.04)	(0.22, 0.46)
z_est_bi_spend		−0.11***	0.10**
		(−0.15, −0.07)	(0.04, 0.16)
days:z_age			−0.0001
			(−0.002, 0.002)
days:male			−0.001
			(−0.005, 0.003)
days:z_est_bi_credit			−0.002
			(−0.004, 0.0000)
days:z_est_bi_spend			−0.0003
			(−0.002, 0.002)
days:z_extra			0.001
			(−0.001, 0.003)
days:z_agree			−0.0005
			(−0.002, 0.002)
days:z_con			−0.0003
			(−0.002, 0.002)
days:z_neur			−0.0000
			(−0.002, 0.002)
days:z_open			0.001
			(−0.001, 0.003)
z_extra	−0.01	0.004	−0.06*
	(−0.03, 0.02)	(−0.03, 0.03)	(−0.12, −0.01)
z_agree	0.01	0.01	0.01
	(−0.02, 0.04)	(−0.02, 0.04)	(−0.05, 0.06)
z_con	0.02	0.03	−0.01
	(−0.01, 0.05)	(−0.004, 0.06)	(−0.07, 0.04)
z_neur	−0.01	−0.01	0.02
	(−0.04, 0.02)	(−0.04, 0.02)	(−0.04, 0.08)
z_open	−0.01	−0.01	0.01
	(−0.04, 0.02)	(−0.04, 0.02)	(−0.04, 0.07)
Observations	2,118	2,118	86,191
Log Likelihood			−48,398,255.00
Akaike Inf. Crit.			96,796,561.00
Bayesian Inf. Crit.			96,796,795.00
Note:			*p<0.05; **p<0.01; ***p<0.001

Table 13: Do personality traits predict holiday spending trajectories? (exploratory, new covariates and fractional bionmial)

8 Session Information

The analyses displayed here were run on a computer with the following settings:

```
##  setting  value
##  version  R version 3.4.4 (2018-03-15)
##  system   x86_64, darwin15.6.0
##  ui       X11
##  language (EN)
##  collate  en_US.UTF-8
##  tz       America/Chicago
##  date     2018-06-19
```

The following packages were used:

package	*	version	date	source
abind		1.4-5	2016-07-21	CRAN (R 3.4.0)
apaTables	*	1.5.1	2017-06-20	CRAN (R 3.4.1)
arm		1.9-3	2016-11-27	CRAN (R 3.4.0)
assertthat		0.2.0	2017-04-11	CRAN (R 3.4.0)
backports		1.1.2	2017-12-13	cran (@1.1.2)
base	*	3.4.4	2018-03-15	local
bayesplot		1.4.0	2017-09-12	CRAN (R 3.4.2)
betareg	*	3.1-0	2016-08-06	CRAN (R 3.4.0)
bindr		0.1.1	2018-03-13	cran (@0.1.1)
bindrcpp	*	0.2.2	2018-03-29	cran (@0.2.2)
blme		1.0-4	2015-06-14	CRAN (R 3.4.0)
broom		0.4.4	2018-03-29	cran (@0.4.4)
carData		3.0-0	2017-08-28	CRAN (R 3.4.1)
cellranger		1.1.0	2016-07-27	CRAN (R 3.4.0)
cli		1.0.0	2017-11-05	CRAN (R 3.4.2)
coda		0.19-1	2016-12-08	CRAN (R 3.4.0)
codetools		0.2-15	2016-10-05	CRAN (R 3.4.4)
coin		1.2-1	2017-07-17	CRAN (R 3.4.1)
colorspace		1.3-2	2016-12-14	CRAN (R 3.4.0)
compiler		3.4.4	2018-03-15	local
crayon		1.3.4	2017-09-16	CRAN (R 3.4.1)
datasets	*	3.4.4	2018-03-15	local
devtools	*	1.13.4	2017-11-09	CRAN (R 3.4.2)
digest		0.6.15	2018-01-28	cran (@0.6.15)
dplyr	*	0.7.5	2018-05-19	cran (@0.7.5)
DT		0.2	2016-08-09	CRAN (R 3.4.0)
effects		4.0-0	2017-09-15	CRAN (R 3.4.1)
evaluate		0.10.1	2017-06-24	CRAN (R 3.4.1)
flexmix		2.3-14	2017-04-28	CRAN (R 3.4.0)
forcats	*	0.2.0	2017-01-23	CRAN (R 3.4.0)
foreign		0.8-69	2017-06-22	CRAN (R 3.4.4)
formatR		1.5	2017-04-25	CRAN (R 3.4.0)
Formula		1.2-2	2017-07-10	CRAN (R 3.4.1)
ggeffects		0.2.2	2017-09-20	CRAN (R 3.4.2)

ggplot2	*	2.2.1	2016-12-30	CRAN (R 3.4.0)
glmmTMB	*	0.1.4	2017-10-26	CRAN (R 3.4.2)
glue		1.2.0	2017-10-29	CRAN (R 3.4.2)
graphics	*	3.4.4	2018-03-15	local
grDevices	*	3.4.4	2018-03-15	local
grid		3.4.4	2018-03-15	local
gridExtra	*	2.3	2017-09-09	CRAN (R 3.4.1)
gtable		0.2.0	2016-02-26	CRAN (R 3.4.0)
haven		1.1.0	2017-07-09	CRAN (R 3.4.0)
highr		0.6	2016-05-09	CRAN (R 3.4.0)
hms		0.3	2016-11-22	CRAN (R 3.4.0)
htmltools		0.3.6	2017-04-28	CRAN (R 3.4.0)
htmlwidgets		0.9	2017-07-10	CRAN (R 3.4.0)
httpuv		1.4.3	2018-05-10	cran (@1.4.3)
httr		1.3.1	2017-08-20	CRAN (R 3.4.1)
jsonlite		1.5	2017-06-01	CRAN (R 3.4.0)
kableExtra	*	0.7.0	2018-01-15	CRAN (R 3.4.3)
knitr	*	1.20	2018-02-20	cran (@1.20)
labeling		0.3	2014-08-23	CRAN (R 3.4.0)
later		0.7.2	2018-05-01	cran (@0.7.2)
lattice		0.20-35	2017-03-25	CRAN (R 3.4.4)
lavaan	*	0.5-23.1097	2017-02-24	CRAN (R 3.4.0)
lazyeval		0.2.1	2017-10-29	CRAN (R 3.4.2)
lme4	*	1.1-14	2017-09-27	CRAN (R 3.4.2)
lmtest		0.9-35	2017-02-11	CRAN (R 3.4.0)
lubridate		1.7.1	2017-11-03	CRAN (R 3.4.2)
magrittr		1.5	2014-11-22	CRAN (R 3.4.0)
MASS		7.3-49	2018-02-23	CRAN (R 3.4.4)
Matrix	*	1.2-12	2017-11-30	CRAN (R 3.4.4)
memoise		1.1.0	2017-04-21	CRAN (R 3.4.0)
merTools		0.3.0	2016-12-12	CRAN (R 3.4.0)
methods	*	3.4.4	2018-03-15	local
mgcv		1.8-23	2018-01-21	CRAN (R 3.4.4)
mime		0.5	2016-07-07	CRAN (R 3.4.0)
minqa		1.2.4	2014-10-09	CRAN (R 3.4.0)
mnormt		1.5-5	2016-10-15	CRAN (R 3.4.0)
modelr		0.1.1	2017-07-24	CRAN (R 3.4.1)
modeltools		0.2-21	2013-09-02	CRAN (R 3.4.0)
multcomp		1.4-8	2017-11-08	CRAN (R 3.4.2)
munsell		0.4.3	2016-02-13	CRAN (R 3.4.0)
mvtnorm		1.0-6	2017-03-02	CRAN (R 3.4.0)
nlme		3.1-131.1	2018-02-16	CRAN (R 3.4.4)
nloptr		1.0.4	2014-08-04	CRAN (R 3.4.0)
nnet		7.3-12	2016-02-02	CRAN (R 3.4.4)
papaja	*	0.1.0.9492	2017-10-14	Github (crsh/papaja@ede6845)
parallel		3.4.4	2018-03-15	local
pbivnorm		0.6.0	2015-01-23	CRAN (R 3.4.0)

pillar		1.2.3	2018-05-25	cran (@1.2.3)
pkgconfig		2.0.1	2017-03-21	CRAN (R 3.4.0)
plyr		1.8.4	2016-06-08	CRAN (R 3.4.0)
prediction		0.2.0	2017-04-19	CRAN (R 3.4.0)
promises		1.0.1	2018-04-13	cran (@1.0.1)
psych	*	1.8.4	2018-05-06	cran (@1.8.4)
purrr	*	0.2.5	2018-05-29	cran (@0.2.5)
pwr		1.2-1	2017-03-25	CRAN (R 3.4.0)
quadprog		1.5-5	2013-04-17	CRAN (R 3.4.0)
R6		2.2.2	2017-06-17	CRAN (R 3.4.0)
RColorBrewer		1.1-2	2014-12-07	CRAN (R 3.4.0)
Rcpp		0.12.17	2018-05-18	cran (@0.12.17)
readr	*	1.1.1	2017-05-16	CRAN (R 3.4.0)
readxl		1.0.0	2017-04-18	CRAN (R 3.4.0)
reshape2		1.4.3	2017-12-11	cran (@1.4.3)
rlang		0.2.1	2018-05-30	cran (@0.2.1)
rmarkdown		1.9	2018-03-01	cran (@1.9)
rprojroot		1.3-2	2018-01-03	cran (@1.3-2)
rstudioapi		0.7	2017-09-07	CRAN (R 3.4.1)
rvest		0.3.2	2016-06-17	CRAN (R 3.4.0)
sandwich		2.4-0	2017-07-26	CRAN (R 3.4.1)
scales		0.5.0	2017-08-24	CRAN (R 3.4.1)
shiny		1.1.0	2018-05-17	cran (@1.1.0)
sjlabelled		1.0.5	2017-11-09	CRAN (R 3.4.2)
sjmisc		2.6.2	2017-10-26	CRAN (R 3.4.2)
sjPlot	*	2.4.0	2017-10-19	CRAN (R 3.4.2)
sjstats		0.12.0	2017-10-16	CRAN (R 3.4.2)
snakecase		0.5.1	2017-09-20	CRAN (R 3.4.2)
splines		3.4.4	2018-03-15	local
stargazer	*	5.2	2015-07-14	CRAN (R 3.4.0)
stats	*	3.4.4	2018-03-15	local
stats4		3.4.4	2018-03-15	local
stringdist		0.9.4.6	2017-07-31	CRAN (R 3.4.1)
stringi		1.2.2	2018-05-02	cran (@1.2.2)
stringr	*	1.3.1	2018-05-10	cran (@1.3.1)
survey		3.32-1	2017-06-22	CRAN (R 3.4.1)
survival		2.41-3	2017-04-04	CRAN (R 3.4.4)
TH.data		1.0-8	2017-01-23	CRAN (R 3.4.0)
tibble	*	1.4.2	2018-01-22	cran (@1.4.2)
tidyr	*	0.8.1	2018-05-18	cran (@0.8.1)
tidyselect		0.2.4	2018-02-26	cran (@0.2.4)
tidyverse	*	1.2.1	2017-11-14	CRAN (R 3.4.2)
TMB		1.7.12	2017-12-11	CRAN (R 3.4.3)
tools		3.4.4	2018-03-15	local
utils	*	3.4.4	2018-03-15	local
viridisLite		0.2.0	2017-03-24	CRAN (R 3.4.0)
withr		2.1.0	2017-11-01	CRAN (R 3.4.2)
xml2		1.1.1	2017-01-24	CRAN (R 3.4.0)

xtable	1.8-2	2016-02-05	CRAN (R 3.4.0)
zoo	1.8-0	2017-04-12	CRAN (R 3.4.0)