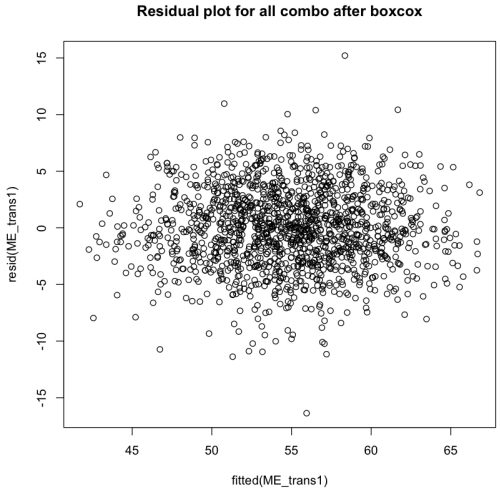
Report:

Introduction: The objective of this report is to share the findings of the possible relationship between Y and G, along with any possible interactions between G and E, and G with G. The context of this is based on the Caspi et al. where he reports on finding a significant gene that is correlated with the higher likelihood of having depression after a stressful event (the environment variable). Our dataset consists of four different environments and 20 different genetic variables. With this dataset, we will create a multiple linear regression model to find significant variables that can be used to statistically predict the likelihood of depression.

Methods and Results:
Using basic R functions, I checked for any missing values of any kind and there were none, the dataset given was complete. The data set included 1453 observations (rows) with 24 independent variables columns. Four of the columns were environment variables (E1, E2, E3, E4), and 20 of them were gene variables, (G1, G2, …, G20). To get a better understanding of my data set, I used the Ggplot library to create a scatter plot matrix. I wanted to see how the environment variables related to the Y and to each other. From there I got a good idea of what variables to keep an eye on when making my model. My strategy from there to come up with a model was to use R to observe a model with all the variables including all second-order interactions, and then get rid of variables that do not contribute to the model. The way to do that was to manually observe the P-values for each coefficient and also perform stepwise regression on the model.

An initial model with no interactions and only environment variables saw an adjusted R-squared value of 0.4592. A model with interactions and all the variables saw an adjusted R squared of 0.5162. However, the summary of that model showed that many of the interactions were statistically insignificant, clearly indicating that the model has a lot of noise (many of the interaction variables provide little value to the model). I also observed the residual plot of the model. Looking the the plot, I didn't notice much of a pattern but decided to perform a boxcox transformation regardless. After applying it, I got this residual plot:



(you can find the plot in the appendix in the notebook)

As you, can see, there is no pattern, and it is elliptical.

After performing the transformation I looked to rid the model of statistically insignificant variables. Using the knitr library, I was able to perform stepwise regression to limit the model to this: (Intercept)+E3+E1:E4+E3:E4+G2:G18+G3:G18.

Using these variables, I created a new linear regression model with the following summary:

```
Call:
lm(formula = I(Y^optimalLambda) ~ (E3 + (E1:E4) + (E3:E4) + (G2:G18) +
    (G3:G18)), data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-14.309  -2.627  -0.046   2.821  15.238

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 33.295603   0.799168  41.663  < 2e-16 ***
E3           0.654793   0.124803   5.247 1.78e-07 ***
E1:E4        0.149748   0.009377  15.970  < 2e-16 ***
E3:E4        0.112209   0.013177   8.515  < 2e-16 ***
G2:G18       2.267694   0.247902   9.148  < 2e-16 ***
G18:G3       0.807799   0.249207   3.241  0.00122 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.041 on 1447 degrees of freedom
Multiple R-squared:  0.5078,    Adjusted R-squared:  0.5061
F-statistic: 298.6 on 5 and 1447 DF,  p-value: < 2.2e-16
```

(you can find the summary in the appendix in the notebook)

As you can see, all the variables are statistically significant as shown by their p value in the last column, except the last one. Although it is significant with an alpha of 0.01, it is not by 0.001. Additionally, during stepwise regression, the BIC decreased relatively less than the rest of the stepwise steps when adding in the last variable G18:G3. Therefore, I decided to remove it.

I experimented with independent variable transformations but the difference was not significant. So in order to keep the readability of the meaning of the model intact, I decided not to include it in the final model.

The **final model** was with an adjusted R square of: 0.5033.
Y^1.6 = 42.11525 + 0.89146E3+ 0.20795E1:E4 + 0.15856E3:E4 + 3.65739G2:G18 + Z

Doing a final test for significant variables, all variables were proven to be significant:

Table: Sig Coefficients

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	42.1152493	1.1063232	38.067763	0e+00
E3	0.8914593	0.1741044	5.120258	3e-07
E1:E4	0.2079469	0.0130877	15.888711	0e+00
E3:E4	0.1585580	0.0183806	8.626372	0e+00
G2:G18	3.6573878	0.3088809	11.840772	0e+00

and there slopes were also shown to be statistically significant (cell 21).

Conclusions:
So once again our final model is :
$Y^{1.6} = 42.11525 + 0.89146E3 + 0.20795E1{:}E4 + 0.15856E3{:}E4 + 3.65739G2{:}G18 + Z.$

Our model shows us that of the four environmental variables, three of them were significant factors along with the combination of two genes, G2 and G18. The gene interaction G2 and G18 are expected to have a statistically significant impact on the likelihood of depression, given different environments, specifically E1, E3, and E4, or an interaction of some of them.
The model is able to explain 50.33% of the variance, and the model was reduced from its 24 independent variables to just 4, using stepwise regression and general analysis. All variables were proven to be statistically significant as well as their slopes, the the 99% confidence intervals of all not including 0, and the residual plot shows no pattern (all observable on the last cel (33)l in the notebook in the appendix).

Begin Appendix:

In [1]:
```
getwd()
```

'/Users/saatvik'

In [2]:
```
wdir <- "/Users/saatvik/Documents/AMS315Project2WD/dataset/"
setwd(wdir)
getwd()
```

'/Users/saatvik/Documents/AMS315Project2WD/dataset'

In [3]:
```
data <- read.csv("P2_378631.csv",header=TRUE)
data
summary(data)
any(sapply(data, function(x) any(is.nan(x))))
any(is.na(data))
any(sapply(data, is.null))
length(data$Y)
```

A data.frame: 1453 × 25

Y	E1	E2	E3	E4	G1	G2	G3	G4	G5	⋯	G11
<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<int>	<int>	<int>	<int>	<int>	⋯	<int>
15.36432	5.411122	7.707438	7.013116	9.490520	1	1	0	1	1	⋯	1
15.63839	5.513429	5.337072	9.538930	8.836345	1	1	1	1	1	⋯	0
13.51293	7.342725	5.361841	5.652944	5.178952	1	1	1	0	0	⋯	1
14.54605	9.117698	7.277338	6.135449	9.039677	1	1	1	1	0	⋯	1
14.92297	7.287646	7.590836	8.548181	9.800264	1	1	0	1	1	⋯	1
14.62860	7.548133	9.323630	5.097581	9.853720	0	0	1	0	0	⋯	1
14.35788	6.080577	5.706601	8.562305	7.038417	1	1	1	1	1	⋯	0
15.28574	9.804724	5.818995	5.434153	8.579416	1	1	1	1	1	⋯	0
15.31310	8.855946	5.864762	8.630110	7.742442	1	1	0	1	1	⋯	1
15.21135	9.907748	8.218389	8.438032	7.403527	1	1	1	1	1	⋯	1
14.46303	5.592810	5.494222	5.990943	8.595854	1	1	1	1	1	⋯	1
14.55306	8.047922	8.459573	5.312573	8.448891	1	0	1	1	1	⋯	1
14.96904	7.965340	8.428293	7.338949	8.097163	1	1	1	1	0	⋯	0
15.49337	9.127582	7.204222	8.684579	5.420102	1	1	1	1	1	⋯	1
14.87599	7.702261	6.808383	5.910515	7.990819	1	0	1	1	1	⋯	1
15.46727	8.055676	8.911857	8.091160	7.674051	1	1	1	0	0	⋯	0
16.04499	9.715707	6.374066	7.035755	8.854134	1	0	1	1	0	⋯	1
15.02214	9.499454	5.458118	8.006301	7.693842	0	0	0	1	1	⋯	1
13.97148	8.122190	5.732438	5.485354	5.643273	0	1	1	1	1	⋯	1
14.13524	5.038378	6.786403	5.651344	6.500594	1	1	1	1	1	⋯	1
14.17789	6.948360	5.602073	8.962045	5.982630	0	1	1	1	1	⋯	1
12.84060	7.418231	7.154117	5.363718	7.441034	0	1	0	1	1	⋯	1
15.05025	8.457965	7.878304	8.936761	9.830805	1	1	1	1	1	⋯	0
13.64761	8.984817	9.526638	5.178666	7.923229	0	1	1	0	1	⋯	1
14.34103	9.096724	7.527755	6.578553	7.088451	1	1	0	1	1	⋯	1
14.06732	8.689942	9.913810	9.229636	5.483259	1	1	0	1	0	⋯	1
12.23837	6.635154	5.523171	8.203703	5.181435	1	0	1	1	1	⋯	1
12.27734	7.015747	9.017709	5.054329	6.711698	1	0	1	1	1	⋯	1
13.69563	8.281051	5.089355	7.250041	5.485999	1	1	0	0	1	⋯	0
11.81812	6.582991	8.544407	6.199046	5.781948	1	1	0	0	1	⋯	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋱	⋮
13.43388	6.836696	7.502678	6.878639	5.392477	1	1	1	1	0	⋯	1
14.38049	5.577445	8.409594	7.308044	7.233383	1	1	1	1	1	⋯	1

Y	E1	E2	E3	E4	G1	G2	G3	G4	G5	⋯	G11
<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<int>	<int>	<int>	<int>	<int>	⋯	<int>
14.61887	9.143086	8.797749	5.352015	6.246185	1	1	1	0	1	⋯	1
14.47237	6.901969	5.805156	7.760035	7.787423	1	1	0	0	1	⋯	1
15.65622	9.394050	7.601001	7.374618	9.239248	1	1	1	1	1	⋯	0
15.94584	9.429965	9.363358	8.253822	9.186511	1	1	0	0	1	⋯	1
14.07845	7.942885	8.446499	5.703365	5.217074	1	1	1	1	1	⋯	1
13.20252	6.151272	6.969178	7.320407	6.237738	1	1	1	1	1	⋯	1
14.81252	6.689357	7.041965	7.321528	7.609186	1	1	1	1	1	⋯	1
15.56714	8.532020	7.739513	8.985459	6.991195	1	1	1	1	1	⋯	0
15.82284	8.616737	7.381887	5.571866	8.255165	1	0	1	1	1	⋯	0
13.65560	8.621256	9.033298	5.882220	7.686253	1	1	1	1	1	⋯	1
14.38228	7.447046	9.741991	5.708466	6.953399	1	1	1	1	0	⋯	1
13.39103	6.493289	5.842743	6.362083	7.880209	1	1	1	0	1	⋯	0
12.06624	7.717950	8.614318	5.555147	5.590434	0	0	1	1	1	⋯	1
15.64005	6.704552	5.379315	6.031978	7.699585	1	1	1	1	1	⋯	1
13.31958	5.173569	8.698983	6.977421	6.848664	1	1	1	1	1	⋯	1
12.43706	8.435205	8.463430	5.121135	7.901393	1	0	0	1	1	⋯	0
14.03966	6.854409	5.077977	5.178702	9.266592	1	1	1	0	1	⋯	0
13.93313	8.099994	5.946394	6.444796	6.433383	1	1	0	0	1	⋯	1
14.34481	9.469406	6.487519	8.445218	9.070366	1	0	0	1	0	⋯	1
13.94865	5.456339	5.771405	7.398727	9.449237	1	0	1	0	1	⋯	1
15.72372	8.499697	9.708750	8.126885	8.342471	1	1	1	0	1	⋯	1
14.59090	9.904418	7.011754	7.993100	6.442094	1	1	1	1	0	⋯	1
13.99840	8.254188	7.134675	5.772481	7.658378	0	1	0	1	1	⋯	1
15.10657	9.553259	8.341112	5.901754	6.324230	1	1	1	1	1	⋯	1
15.16591	9.874072	5.709503	8.866929	7.609953	1	1	0	1	1	⋯	1
12.55092	6.074717	7.147842	8.587037	5.667755	0	1	1	1	1	⋯	0
16.21491	5.255842	6.459330	9.638974	8.989498	1	1	1	0	0	⋯	1
14.36388	7.069610	9.737603	7.840212	8.950186	1	1	1	1	1	⋯	0

```
       Y               E1              E2              E3        
 Min.   :10.62   Min.   :5.005   Min.   :5.002   Min.   :5.012  
 1st Qu.:13.75   1st Qu.:6.191   1st Qu.:6.301   1st Qu.:6.228  
 Median :14.45   Median :7.447   Median :7.474   Median :7.407  
 Mean   :14.41   Mean   :7.451   Mean   :7.486   Mean   :7.462  
 3rd Qu.:15.09   3rd Qu.:8.690   3rd Qu.:8.717   3rd Qu.:8.675  
 Max.   :17.56   Max.   :9.998   Max.   :9.997   Max.   :9.995  
       E4              G1              G2              G3        
 Min.   :5.007   Min.   :0.0000   Min.   :0.000   Min.   :0.0000  
 1st Qu.:6.172   1st Qu.:1.0000   1st Qu.:1.000   1st Qu.:1.0000  
 Median :7.511   Median :1.0000   Median :1.000   Median :1.0000  
 Mean   :7.499   Mean   :0.7997   Mean   :0.799   Mean   :0.8087  
 3rd Qu.:8.771   3rd Qu.:1.0000   3rd Qu.:1.000   3rd Qu.:1.0000  
 Max.   :9.992   Max.   :1.0000   Max.   :1.000   Max.   :1.0000  
       G4              G5              G6              G7        
 Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000  
 1st Qu.:1.0000   1st Qu.:1.0000   1st Qu.:1.0000   1st Qu.:1.0000  
 Median :1.0000   Median :1.0000   Median :1.0000   Median :1.0000  
 Mean   :0.7915   Mean   :0.8032   Mean   :0.7688   Mean   :0.7784  
 3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000  
 Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.0000  
       G8              G9              G10             G11       
 Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000  
 1st Qu.:1.0000   1st Qu.:1.0000   1st Qu.:1.0000   1st Qu.:1.0000  
 Median :1.0000   Median :1.0000   Median :1.0000   Median :1.0000  
 Mean   :0.8018   Mean   :0.8211   Mean   :0.7956   Mean   :0.8052  
 3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000  
 Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.0000  
      G12             G13             G14             G15       
 Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000  
 1st Qu.:1.0000   1st Qu.:1.0000   1st Qu.:1.0000   1st Qu.:1.0000  
 Median :1.0000   Median :1.0000   Median :1.0000   Median :1.0000  
 Mean   :0.7935   Mean   :0.7873   Mean   :0.8142   Mean   :0.8018  
 3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000  
 Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.0000  
      G16             G17             G18             G19       
 Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000  
 1st Qu.:1.0000   1st Qu.:1.0000   1st Qu.:1.0000   1st Qu.:1.0000  
 Median :1.0000   Median :1.0000   Median :1.0000   Median :1.0000  
 Mean   :0.7983   Mean   :0.7866   Mean   :0.7956   Mean   :0.8231  
 3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000  
 Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.0000  
      G20             
 Min.   :0.0000  
 1st Qu.:1.0000  
 Median :1.0000  
 Mean   :0.8004  
 3rd Qu.:1.0000  
 Max.   :1.0000  
FALSE
FALSE
FALSE
1453
```

In [4]:

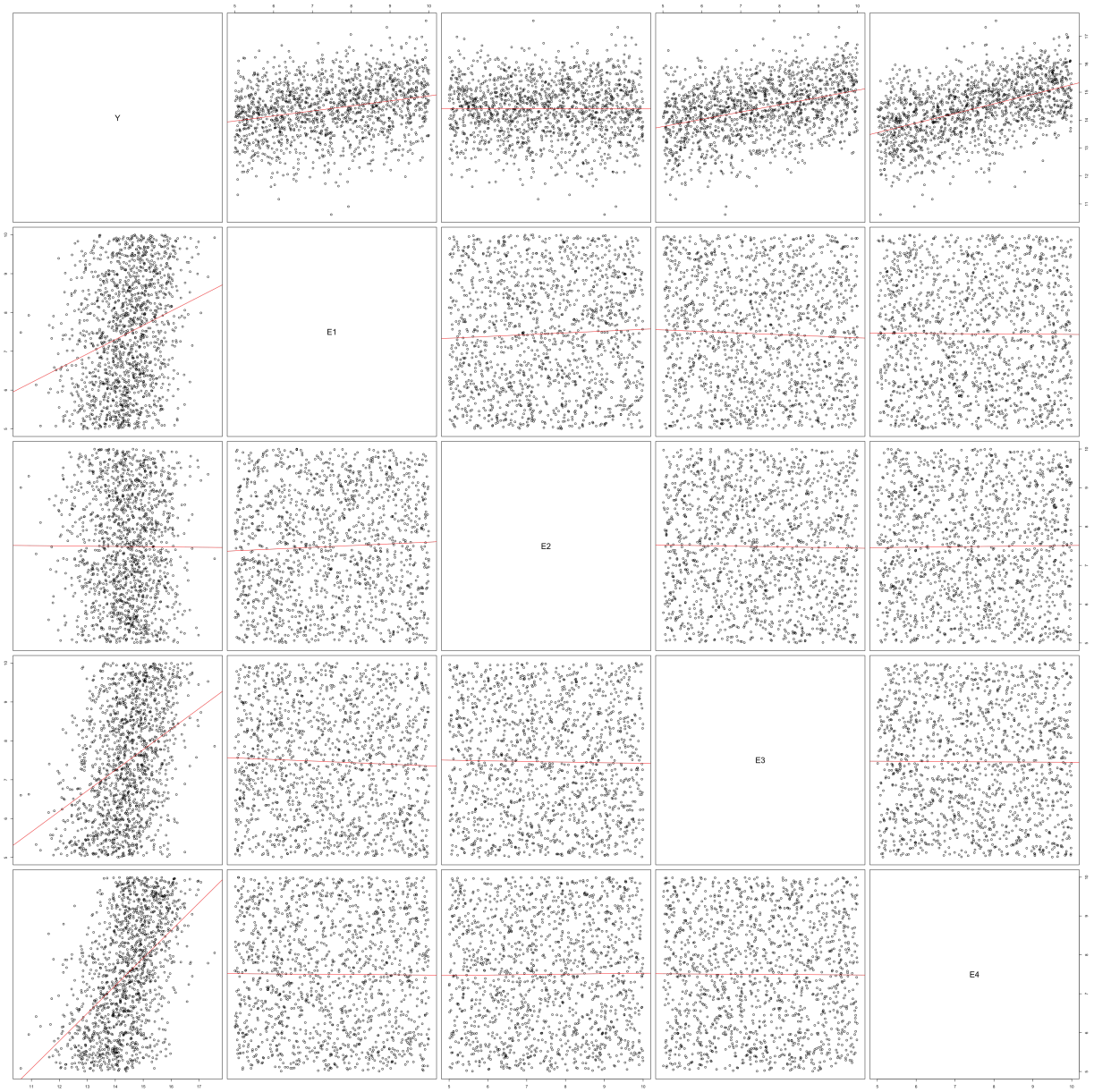
```
install.packages("car")
library(car)
```

```
The downloaded binary packages are in
        /var/folders/rh/bp1_zwxj76z6nrtwmrl525600000gn/T//Rtmpgy5Fms/downloade
d_packages
Loading required package: carData
```

In [5]:
```r
install.packages(c("ggplot2", "GGally"))
library(ggplot2)
library(GGally)
```

```
The downloaded binary packages are in
        /var/folders/rh/bp1_zwxj76z6nrtwmrl525600000gn/T//Rtmpgy5Fms/downloade
d_packages
Registered S3 method overwritten by 'GGally':
  method from
  +.gg   ggplot2
```

In [6]:
```r
dfE <- data[c("Y", "E1", "E2", "E3", "E4")]

options(repr.plot.width=30, repr.plot.height=30)
#pairs(dfE, pch = 16, cex = 1.5, col = "black")
pairs(dfE, panel = function(x, y, ...) {
  points(x, y, ...)
  abline(lm(y ~ x), col = "red")
})

options(repr.plot.width=7, repr.plot.height=7)
```



In [7]:

```
ME <- lm(Y ~ E1+E2+E3+E4, data = data)
summary(ME)
```

```
Call:
lm(formula = Y ~ E1 + E2 + E3 + E4, data = data)

Residuals:
     Min       1Q   Median       3Q      Max
-2.81288 -0.49177  0.00382  0.51844  2.36350

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  8.43932    0.20412  41.346   <2e-16 ***
E1           0.19384    0.01352  14.340   <2e-16 ***
E2          -0.01218    0.01372  -0.888    0.375
E3           0.27037    0.01357  19.931   <2e-16 ***
E4           0.34664    0.01342  25.828   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.746 on 1448 degrees of freedom
Multiple R-squared:  0.4607,    Adjusted R-squared:  0.4592
F-statistic: 309.2 on 4 and 1448 DF,  p-value: < 2.2e-16
```

In [8]:

```
ME_testNoI <- lm(Y ~ E1+E2+E3+E4+G1+G2+G3+G4+G5+G6+G7+G8+G9+G10+G11+G12+G13+G14
summary(ME_testNoI)
```

```
Call:
lm(formula = Y ~ E1 + E2 + E3 + E4 + G1 + G2 + G3 + G4 + G5 +
    G6 + G7 + G8 + G9 + G10 + G11 + G12 + G13 + G14 + G15 + G16 +
    G17 + G18 + G19 + G20, data = data)

Residuals:
     Min       1Q   Median       3Q      Max
-2.67786 -0.45085 -0.00822  0.48742  2.61488

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.813262   0.251267  31.096   <2e-16 ***
E1           0.196336   0.012862  15.265   <2e-16 ***
E2          -0.009078   0.013080  -0.694   0.4878
E3           0.265576   0.012907  20.576   <2e-16 ***
E4           0.352108   0.012726  27.669   <2e-16 ***
G1          -0.045627   0.046661  -0.978   0.3283
G2           0.399761   0.046523   8.593   <2e-16 ***
G3           0.065475   0.047564   1.377   0.1689
G4          -0.014444   0.045790  -0.315   0.7525
G5          -0.015411   0.046707  -0.330   0.7415
G6          -0.069026   0.044195  -1.562   0.1185
G7           0.014784   0.044744   0.330   0.7411
G8          -0.096178   0.046750  -2.057   0.0398 *
G9           0.051281   0.048592   1.055   0.2914
G10         -0.043765   0.046262  -0.946   0.3443
G11         -0.046808   0.047078  -0.994   0.3203
G12          0.053096   0.045923   1.156   0.2478
G13          0.044876   0.045543   0.985   0.3246
G14          0.046861   0.047816   0.980   0.3272
G15         -0.019236   0.046835  -0.411   0.6813
G16          0.038413   0.046420   0.827   0.4081
G17         -0.092722   0.045464  -2.039   0.0416 *
G18          0.452659   0.046249   9.787   <2e-16 ***
G19          0.012653   0.048676   0.260   0.7949
G20         -0.015588   0.046849  -0.333   0.7394
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7043 on 1428 degrees of freedom
Multiple R-squared:  0.526,     Adjusted R-squared:  0.518
F-statistic: 66.02 on 24 and 1428 DF,  p-value: < 2.2e-16
```

Looking at the p-values, E2 seems to be an insignificant variable, failing the t-test. Will be sure to include the rest of the environment variables.

In [9]:
```
ME_full <- lm(Y ~ (E1+E2+E3+E4+G1+G2+G3+G4+G5+G6+G7+G8+G9+G10+G11+G12+G13+G14+(
summary(ME_full)
```

```
Call:
lm(formula = Y ~ (E1 + E2 + E3 + E4 + G1 + G2 + G3 + G4 + G5 +
    G6 + G7 + G8 + G9 + G10 + G11 + G12 + G13 + G14 + G15 + G16 +
    G17 + G18 + G19 + G20)^2, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-2.9914 -0.4001 -0.0045  0.4346  2.5344

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.0496057  2.8132834   2.150  0.03173 *
E1           0.2575771  0.1845935   1.395  0.16317
E2           0.1601122  0.1887128   0.848  0.39637
E3           0.1926858  0.1809550   1.065  0.28718
E4           0.6129301  0.1897078   3.231  0.00127 **
G1           0.1870334  0.7278983   0.257  0.79726
G2           0.3104013  0.7504417   0.414  0.67923
G3           0.5131758  0.7352317   0.698  0.48533
G4           0.2134159  0.6993755   0.305  0.76031
G5          -0.7091115  0.6952454  -1.020  0.30797
G6           0.3417671  0.6847163   0.499  0.61778
G7          -0.0059013  0.6648213  -0.009  0.99292
G8          -0.2285438  0.7336056  -0.312  0.75545
G9          -1.1030343  0.7649008  -1.442  0.14956
G10         -0.0108282  0.7342760  -0.015  0.98824
G11         -0.7110442  0.7254633  -0.980  0.32723
G12         -0.5113470  0.7035158  -0.727  0.46747
G13          2.0755997  0.7313569   2.838  0.00462 **
G14         -0.4228055  0.7866183  -0.537  0.59103
G15         -0.2634945  0.7131347  -0.369  0.71183
G16          0.5079543  0.6984382   0.727  0.46721
G17         -0.8374251  0.6929435  -1.209  0.22710
G18          1.2382872  0.7394040   1.675  0.09426 .
G19          1.0856280  0.6993403   1.552  0.12085
G20         -0.0649710  0.7435310  -0.087  0.93038
E1:E2       -0.0022276  0.0102137  -0.218  0.82739
E1:E3        0.0048706  0.0100385   0.485  0.62763
E1:E4       -0.0015194  0.0101261  -0.150  0.88075
E1:G1       -0.0496312  0.0372544  -1.332  0.18305
E1:G2       -0.0530275  0.0369924  -1.433  0.15200
E1:G3       -0.0008747  0.0385277  -0.023  0.98189
E1:G4       -0.0017590  0.0365191  -0.048  0.96159
E1:G5       -0.0076800  0.0363868  -0.211  0.83287
E1:G6        0.0151604  0.0346779   0.437  0.66207
E1:G7       -0.0248852  0.0355744  -0.700  0.48437
E1:G8        0.0281119  0.0378250   0.743  0.45751
E1:G9        0.0444962  0.0380568   1.169  0.24256
E1:G10       0.0389926  0.0377181   1.034  0.30145
E1:G11       0.0252825  0.0367674   0.688  0.49182
E1:G12       0.0120954  0.0355156   0.341  0.73349
E1:G13      -0.0705156  0.0353147  -1.997  0.04608 *
E1:G14       0.0123856  0.0384786   0.322  0.74760
E1:G15       0.0249159  0.0373563   0.667  0.50492
E1:G16      -0.0464569  0.0369626  -1.257  0.20906
E1:G17       0.0481992  0.0347978   1.385  0.16628
E1:G18      -0.0238287  0.0382576  -0.623  0.53351
E1:G19      -0.0103402  0.0381663  -0.271  0.78650
E1:G20      -0.0603200  0.0368654  -1.636  0.10207
E2:E3        0.0142394  0.0101484   1.403  0.16085
```

E2:E4	0.0051126	0.0099277	0.515	0.60667	
E2:G1	-0.0456978	0.0372123	-1.228	0.21969	
E2:G2	0.0081913	0.0383822	0.213	0.83104	
E2:G3	0.0252269	0.0384373	0.656	0.51175	
E2:G4	-0.0147082	0.0368967	-0.399	0.69024	
E2:G5	-0.0350133	0.0368957	-0.949	0.34283	
E2:G6	0.0022143	0.0356390	0.062	0.95047	
E2:G7	-0.0366319	0.0359401	-1.019	0.30830	
E2:G8	0.0143083	0.0378369	0.378	0.70538	
E2:G9	0.0021426	0.0388683	0.055	0.95605	
E2:G10	-0.0528205	0.0372181	-1.419	0.15611	
E2:G11	-0.0016435	0.0370820	-0.044	0.96466	
E2:G12	-0.0006030	0.0373537	-0.016	0.98712	
E2:G13	-0.0429623	0.0351536	-1.222	0.22191	
E2:G14	-0.0398666	0.0388857	-1.025	0.30547	
E2:G15	-0.0470850	0.0362942	-1.297	0.19478	
E2:G16	0.0303063	0.0381054	0.795	0.42659	
E2:G17	-0.0045473	0.0363376	-0.125	0.90043	
E2:G18	-0.0337371	0.0370212	-0.911	0.36233	
E2:G19	-0.0561039	0.0383431	-1.463	0.14368	
E2:G20	-0.0377270	0.0373098	-1.011	0.31214	
E3:E4	-0.0071637	0.0101760	-0.704	0.48159	
E3:G1	0.0592907	0.0362629	1.635	0.10232	
E3:G2	0.0071610	0.0373197	0.192	0.84787	
E3:G3	-0.0704712	0.0368404	-1.913	0.05601	.
E3:G4	-0.0646770	0.0373714	-1.731	0.08378	.
E3:G5	0.0254054	0.0379851	0.669	0.50374	
E3:G6	-0.0506614	0.0344048	-1.473	0.14116	
E3:G7	0.0033199	0.0353873	0.094	0.92527	
E3:G8	-0.0721880	0.0372237	-1.939	0.05271	.
E3:G9	0.0274058	0.0394002	0.696	0.48684	
E3:G10	0.0222390	0.0364655	0.610	0.54207	
E3:G11	0.0691889	0.0359366	1.925	0.05444	.
E3:G12	0.0437209	0.0376194	1.162	0.24540	
E3:G13	-0.0138020	0.0356935	-0.387	0.69906	
E3:G14	-0.0296837	0.0364583	-0.814	0.41571	
E3:G15	0.0141756	0.0372901	0.380	0.70391	
E3:G16	-0.0244543	0.0364672	-0.671	0.50262	
E3:G17	0.0247436	0.0361908	0.684	0.49430	
E3:G18	0.0046152	0.0375417	0.123	0.90218	
E3:G19	-0.0240843	0.0376534	-0.640	0.52254	
E3:G20	0.0222475	0.0364460	0.610	0.54170	
E4:G1	-0.0389653	0.0369999	-1.053	0.29251	
E4:G2	-0.0100268	0.0360467	-0.278	0.78094	
E4:G3	-0.0484936	0.0384406	-1.262	0.20738	
E4:G4	0.0073448	0.0356761	0.206	0.83693	
E4:G5	0.0261535	0.0356077	0.734	0.46280	
E4:G6	-0.0206611	0.0353622	-0.584	0.55915	
E4:G7	0.0122619	0.0353855	0.347	0.72901	
E4:G8	0.0331261	0.0367666	0.901	0.36779	
E4:G9	-0.0216258	0.0391909	-0.552	0.58119	
E4:G10	0.0028321	0.0370813	0.076	0.93913	
E4:G11	-0.0505224	0.0373740	-1.352	0.17670	
E4:G12	-0.0489209	0.0367582	-1.331	0.18349	
E4:G13	-0.0316047	0.0352993	-0.895	0.37079	
E4:G14	-0.0212520	0.0382452	-0.556	0.57854	
E4:G15	0.0318894	0.0369681	0.863	0.38853	
E4:G16	-0.0826448	0.0373300	-2.214	0.02703	*
E4:G17	0.0176953	0.0360628	0.491	0.62375	
E4:G18	-0.0497616	0.0360951	-1.379	0.16828	

E4:G19	-0.0364268	0.0386903	-0.941	0.34665	
E4:G20	0.0308141	0.0368999	0.835	0.40385	
G1:G2	0.1617629	0.1367512	1.183	0.23709	
G1:G3	0.0498834	0.1492588	0.334	0.73828	
G1:G4	0.0719060	0.1339287	0.537	0.59144	
G1:G5	0.2582501	0.1407603	1.835	0.06681	.
G1:G6	-0.1232617	0.1304193	-0.945	0.34480	
G1:G7	0.0148682	0.1328441	0.112	0.91090	
G1:G8	-0.0089066	0.1432457	-0.062	0.95043	
G1:G9	-0.1372909	0.1392765	-0.986	0.32447	
G1:G10	0.0055128	0.1395727	0.039	0.96850	
G1:G11	-0.0565234	0.1406137	-0.402	0.68778	
G1:G12	0.2206379	0.1376657	1.603	0.10927	
G1:G13	-0.1523874	0.1347697	-1.131	0.25841	
G1:G14	0.0818523	0.1383329	0.592	0.55416	
G1:G15	0.1150346	0.1396727	0.824	0.41034	
G1:G16	-0.1301866	0.1391648	-0.935	0.34973	
G1:G17	0.0110104	0.1318619	0.083	0.93347	
G1:G18	-0.0421279	0.1429952	-0.295	0.76834	
G1:G19	-0.0899433	0.1434882	-0.627	0.53089	
G1:G20	0.1270610	0.1326126	0.958	0.33819	
G2:G3	-0.0808710	0.1421617	-0.569	0.56956	
G2:G4	-0.0608076	0.1302504	-0.467	0.64069	
G2:G5	0.1844392	0.1340299	1.376	0.16906	
G2:G6	-0.2061119	0.1302572	-1.582	0.11385	
G2:G7	0.1109537	0.1199435	0.925	0.35513	
G2:G8	-0.1766061	0.1278128	-1.382	0.16731	
G2:G9	0.2243340	0.1390149	1.614	0.10686	
G2:G10	-0.1247573	0.1285562	-0.970	0.33203	
G2:G11	-0.0114251	0.1313507	-0.087	0.93070	
G2:G12	-0.0123942	0.1270284	-0.098	0.92229	
G2:G13	-0.2723113	0.1250334	-2.178	0.02962	*
G2:G14	0.0517361	0.1400574	0.369	0.71190	
G2:G15	0.0840430	0.1453219	0.578	0.56316	
G2:G16	0.3381485	0.1469436	2.301	0.02156	*
G2:G17	0.2376652	0.1327815	1.790	0.07373	.
G2:G18	-0.1007989	0.1280715	-0.787	0.43141	
G2:G19	0.0085706	0.1411031	0.061	0.95158	
G2:G20	0.1431726	0.1265528	1.131	0.25815	
G3:G4	-0.0643293	0.1298811	-0.495	0.62049	
G3:G5	-0.1091654	0.1374532	-0.794	0.42724	
G3:G6	-0.0345398	0.1242046	-0.278	0.78099	
G3:G7	0.2435566	0.1308974	1.861	0.06305	.
G3:G8	0.1292269	0.1375868	0.939	0.34780	
G3:G9	-0.0531441	0.1400936	-0.379	0.70450	
G3:G10	-0.0539050	0.1401800	-0.385	0.70065	
G3:G11	0.2170548	0.1352487	1.605	0.10880	
G3:G12	-0.0449448	0.1280874	-0.351	0.72573	
G3:G13	-0.0541443	0.1342840	-0.403	0.68687	
G3:G14	0.0665313	0.1366607	0.487	0.62647	
G3:G15	0.0678949	0.1334672	0.509	0.61106	
G3:G16	0.1635250	0.1373596	1.190	0.23410	
G3:G17	-0.1550555	0.1345928	-1.152	0.24955	
G3:G18	0.1702085	0.1299362	1.310	0.19048	
G3:G19	-0.0302003	0.1466658	-0.206	0.83690	
G3:G20	-0.1459396	0.1316744	-1.108	0.26795	
G4:G5	0.2123119	0.1281109	1.657	0.09774	.
G4:G6	0.0423576	0.1242172	0.341	0.73317	
G4:G7	0.2349998	0.1273390	1.845	0.06523	.
G4:G8	-0.0724449	0.1303779	-0.556	0.57856	

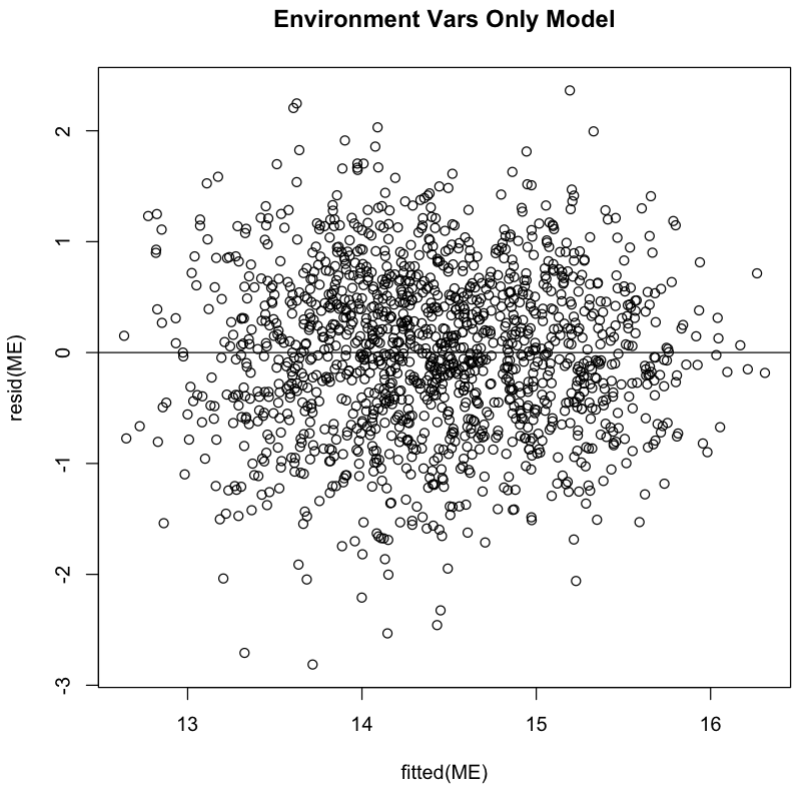
G4:G9	−0.1053211	0.1480708	−0.711	0.47705	
G4:G10	0.1157517	0.1278970	0.905	0.36563	
G4:G11	0.0567956	0.1325046	0.429	0.66827	
G4:G12	−0.1287215	0.1242387	−1.036	0.30038	
G4:G13	0.1262946	0.1260768	1.002	0.31669	
G4:G14	0.2318755	0.1383234	1.676	0.09395	.
G4:G15	−0.1259827	0.1274297	−0.989	0.32304	
G4:G16	−0.0709709	0.1291603	−0.549	0.58278	
G4:G17	−0.2858549	0.1334459	−2.142	0.03239	*
G4:G18	−0.0440718	0.1342434	−0.328	0.74275	
G4:G19	0.2275656	0.1368845	1.662	0.09669	.
G4:G20	0.0331622	0.1329114	0.250	0.80301	
G5:G6	0.0768113	0.1255696	0.612	0.54085	
G5:G7	−0.0312932	0.1297680	−0.241	0.80948	
G5:G8	0.0891937	0.1309550	0.681	0.49594	
G5:G9	−0.0624316	0.1511114	−0.413	0.67957	
G5:G10	−0.1726264	0.1338407	−1.290	0.19738	
G5:G11	0.0420644	0.1427304	0.295	0.76827	
G5:G12	−0.0059213	0.1394733	−0.042	0.96614	
G5:G13	−0.0800158	0.1289213	−0.621	0.53495	
G5:G14	0.2173991	0.1395126	1.558	0.11944	
G5:G15	0.1202218	0.1340323	0.897	0.36993	
G5:G16	−0.0423993	0.1303800	−0.325	0.74509	
G5:G17	0.0409903	0.1280539	0.320	0.74895	
G5:G18	−0.0546332	0.1343416	−0.407	0.68432	
G5:G19	0.0186951	0.1398331	0.134	0.89367	
G5:G20	0.0967255	0.1406833	0.688	0.49188	
G6:G7	−0.1285888	0.1237557	−1.039	0.29900	
G6:G8	0.1546692	0.1318990	1.173	0.24118	
G6:G9	0.0297256	0.1269337	0.234	0.81489	
G6:G10	−0.0064342	0.1287435	−0.050	0.96015	
G6:G11	0.1566873	0.1352666	1.158	0.24696	
G6:G12	0.0973209	0.1236872	0.787	0.43154	
G6:G13	−0.0866196	0.1210352	−0.716	0.47435	
G6:G14	0.0214685	0.1351028	0.159	0.87377	
G6:G15	−0.0559576	0.1259267	−0.444	0.65686	
G6:G16	−0.1792946	0.1189016	−1.508	0.13185	
G6:G17	0.2542924	0.1236017	2.057	0.03988	*
G6:G18	0.0567197	0.1267773	0.447	0.65467	
G6:G19	−0.2111247	0.1337195	−1.579	0.11464	
G6:G20	0.1692367	0.1443947	1.172	0.24142	
G7:G8	0.0385425	0.1256283	0.307	0.75905	
G7:G9	−0.0461727	0.1313946	−0.351	0.72535	
G7:G10	−0.1341862	0.1226195	−1.094	0.27404	
G7:G11	0.2115523	0.1298451	1.629	0.10353	
G7:G12	−0.1075853	0.1250294	−0.860	0.38970	
G7:G13	0.0689428	0.1242212	0.555	0.57900	
G7:G14	0.1405774	0.1293165	1.087	0.27723	
G7:G15	−0.1945104	0.1298227	−1.498	0.13434	
G7:G16	0.0321356	0.1303570	0.247	0.80532	
G7:G17	0.0815511	0.1292996	0.631	0.52835	
G7:G18	0.0727477	0.1314382	0.553	0.58005	
G7:G19	−0.0681094	0.1429744	−0.476	0.63390	
G7:G20	−0.1183924	0.1260317	−0.939	0.34773	
G8:G9	−0.1336980	0.1509695	−0.886	0.37602	
G8:G10	0.0735280	0.1274687	0.577	0.56417	
G8:G11	−0.1064452	0.1410034	−0.755	0.45046	
G8:G12	0.1363374	0.1344065	1.014	0.31062	
G8:G13	−0.1121347	0.1302729	−0.861	0.38955	
G8:G14	−0.0402110	0.1410735	−0.285	0.77567	

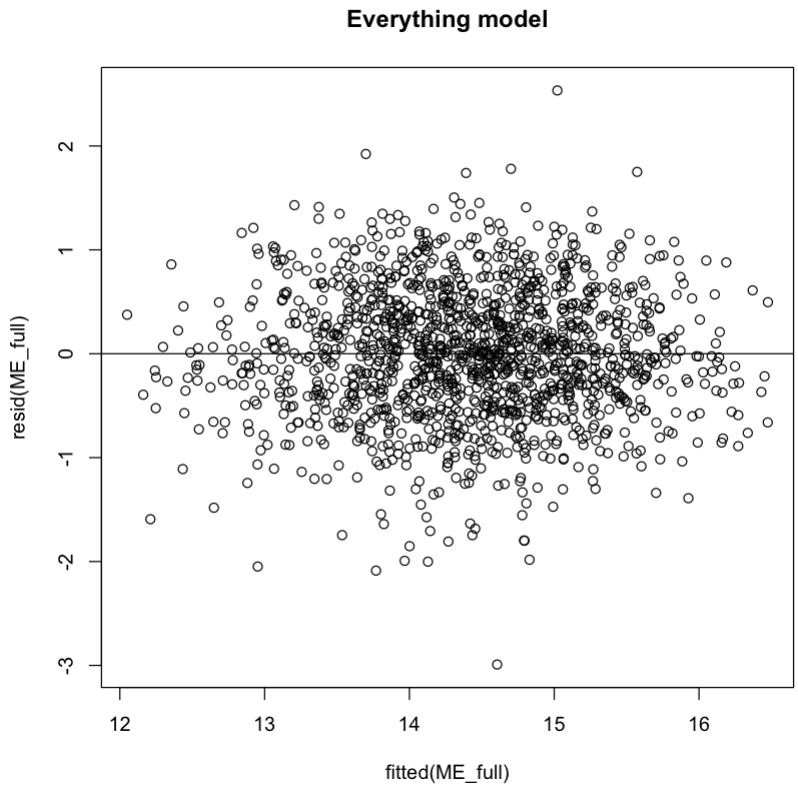
G8:G15	-0.0763964	0.1458893	-0.524	0.60062	
G8:G16	0.1330467	0.1283523	1.037	0.30015	
G8:G17	0.1001296	0.1415965	0.707	0.47962	
G8:G18	-0.0968160	0.1328898	-0.729	0.46643	
G8:G19	0.1987624	0.1448359	1.372	0.17023	
G8:G20	-0.1477027	0.1381262	-1.069	0.28514	
G9:G10	0.2719184	0.1367028	1.989	0.04692	*
G9:G11	0.1584239	0.1496219	1.059	0.28990	
G9:G12	-0.0022194	0.1519801	-0.015	0.98835	
G9:G13	0.0576537	0.1412739	0.408	0.68328	
G9:G14	-0.0327582	0.1444234	-0.227	0.82060	
G9:G15	0.1346351	0.1392653	0.967	0.33387	
G9:G16	0.2464301	0.1337677	1.842	0.06570	.
G9:G17	-0.0455835	0.1412112	-0.323	0.74690	
G9:G18	-0.0740252	0.1379332	-0.537	0.59160	
G9:G19	0.2354990	0.1433611	1.643	0.10072	
G9:G20	0.3028250	0.1443139	2.098	0.03609	*
G10:G11	-0.0023327	0.1335811	-0.017	0.98607	
G10:G12	0.0592914	0.1265342	0.469	0.63946	
G10:G13	-0.1209925	0.1391420	-0.870	0.38472	
G10:G14	0.0793911	0.1552465	0.511	0.60918	
G10:G15	0.1022116	0.1341407	0.762	0.44623	
G10:G16	-0.0994325	0.1307123	-0.761	0.44699	
G10:G17	-0.0168001	0.1260158	-0.133	0.89397	
G10:G18	-0.0187919	0.1308233	-0.144	0.88581	
G10:G19	-0.0596952	0.1316271	-0.454	0.65026	
G10:G20	0.0047724	0.1289254	0.037	0.97048	
G11:G12	-0.2461619	0.1355207	-1.816	0.06957	.
G11:G13	-0.0724326	0.1422649	-0.509	0.61075	
G11:G14	-0.0494507	0.1451654	-0.341	0.73343	
G11:G15	-0.0410313	0.1370041	-0.299	0.76462	
G11:G16	0.1789891	0.1304353	1.372	0.17025	
G11:G17	0.1234251	0.1399156	0.882	0.37789	
G11:G18	0.0414112	0.1283284	0.323	0.74698	
G11:G19	0.0570027	0.1401100	0.407	0.68420	
G11:G20	-0.2199641	0.1359804	-1.618	0.10602	
G12:G13	-0.0113044	0.1344607	-0.084	0.93301	
G12:G14	0.4325731	0.1324809	3.265	0.00113	**
G12:G15	0.1415254	0.1324169	1.069	0.28539	
G12:G16	-0.1226055	0.1342663	-0.913	0.36135	
G12:G17	-0.0116634	0.1321479	-0.088	0.92969	
G12:G18	0.3128830	0.1367063	2.289	0.02228	*
G12:G19	-0.1505707	0.1471865	-1.023	0.30653	
G12:G20	0.1035674	0.1345168	0.770	0.44150	
G13:G14	-0.1642233	0.1419375	-1.157	0.24751	
G13:G15	0.0072764	0.1291001	0.056	0.95506	
G13:G16	0.0495712	0.1304521	0.380	0.70402	
G13:G17	-0.1765527	0.1248656	-1.414	0.15765	
G13:G18	0.0820921	0.1421422	0.578	0.56369	
G13:G19	0.0305493	0.1458646	0.209	0.83415	
G13:G20	-0.1692813	0.1356775	-1.248	0.21240	
G14:G15	0.0259601	0.1400385	0.185	0.85296	
G14:G16	0.2441464	0.1384671	1.763	0.07813	.
G14:G17	-0.0900781	0.1336371	-0.674	0.50041	
G14:G18	0.0982980	0.1494282	0.658	0.51078	
G14:G19	-0.1209276	0.1373277	-0.881	0.37873	
G14:G20	0.0957457	0.1450895	0.660	0.50945	
G15:G16	-0.1907498	0.1426476	-1.337	0.18142	
G15:G17	0.0844462	0.1349839	0.626	0.53170	
G15:G18	0.1353059	0.1324569	1.022	0.30723	

```
G15:G19     -0.2132261  0.1403208  -1.520  0.12890
G15:G20     -0.0643329  0.1484666  -0.433  0.66487
G16:G17      0.1526723  0.1320782   1.156  0.24795
G16:G18     -0.3534338  0.1390798  -2.541  0.01118 *
G16:G19      0.1264456  0.1420881   0.890  0.37370
G16:G20      0.0380140  0.1373753   0.277  0.78205
G17:G18     -0.1899988  0.1379936  -1.377  0.16882
G17:G19     -0.0454304  0.1348895  -0.337  0.73633
G17:G20      0.0242489  0.1339754   0.181  0.85640
G18:G19     -0.1347717  0.1444695  -0.933  0.35108
G18:G20      0.1514790  0.1370580   1.105  0.26930
G19:G20      0.0338001  0.1401443   0.241  0.80946
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7056 on 1152 degrees of freedom
Multiple R-squared:  0.6162,    Adjusted R-squared:  0.5162
F-statistic: 6.164 on 300 and 1152 DF,  p-value: < 2.2e-16
```

Accounting for interactions, a lot of the variance changed. Unable to test if an variable is needed. Going to need to cut down on the interactions using stepwise regression

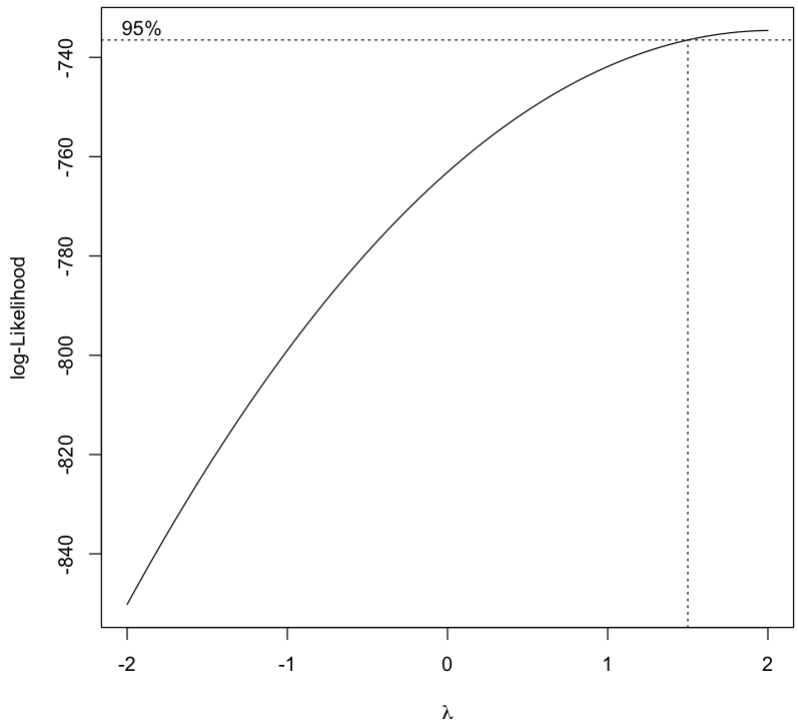
In [10]:
```
plot(fitted(ME), resid(ME), abline(0,0), main = "Environment Vars Only Model")
plot(fitted(ME_full), resid(ME_full), abline(0,0), main = "Everything model")
```





In [11]:
```
library(MASS)
transformation <- boxcox(ME_full)
optimalLambda <- 1.5
optimalLambda
```

1.5



In [12]:

```
ME_trans1 <- lm(I(Y^optimalLambda) ~ (E1+E2+E3+E4+G1+G2+G3+G4+G5+G6+G7+G8+G9+G
summary(ME_trans1)
```

```
Call:
lm(formula = I(Y^optimalLambda) ~ (E1 + E2 + E3 + E4 + G1 + G2 +
    G3 + G4 + G5 + G6 + G7 + G8 + G9 + G10 + G11 + G12 + G13 +
    G14 + G15 + G16 + G17 + G18 + G19 + G20)^2, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-16.368  -2.297  -0.054   2.437  15.207

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.819369  15.939765   0.616  0.53800
E1           1.253122   1.045887   1.198  0.23111
E2           0.940846   1.069227   0.880  0.37908
E3           0.877523   1.025272   0.856  0.39223
E4           3.228641   1.074865   3.004  0.00272 **
G1           1.465743   4.124194   0.355  0.72235
G2           1.454589   4.251923   0.342  0.73234
G3           2.833600   4.165745   0.680  0.49651
G4           1.131593   3.962588   0.286  0.77526
G5          -3.827437   3.939187  -0.972  0.33144
G6           2.297536   3.879530   0.592  0.55382
G7           0.270993   3.766807   0.072  0.94266
G8          -1.269964   4.156531  -0.306  0.76001
G9          -6.305553   4.333847  -1.455  0.14595
G10         -0.072989   4.160330  -0.018  0.98601
G11         -3.986913   4.110398  -0.970  0.33227
G12         -3.000012   3.986046  -0.753  0.45183
G13         11.929500   4.143791   2.879  0.00406 **
G14         -2.448786   4.456896  -0.549  0.58281
G15         -1.573120   4.040546  -0.389  0.69710
G16          2.798778   3.957277   0.707  0.47956
G17         -4.942796   3.926145  -1.259  0.20831
G18          6.409765   4.189384   1.530  0.12629
G19          6.251864   3.962388   1.578  0.11489
G20         -0.514974   4.212768  -0.122  0.90273
E1:E2       -0.010937   0.057870  -0.189  0.85013
E1:E3        0.036100   0.056877   0.635  0.52575
E1:E4        0.002609   0.057374   0.045  0.96374
E1:G1       -0.298982   0.211080  -1.416  0.15692
E1:G2       -0.298797   0.209595  -1.426  0.15426
E1:G3       -0.002752   0.218294  -0.013  0.98994
E1:G4        0.002320   0.206913   0.011  0.99106
E1:G5       -0.054351   0.206164  -0.264  0.79211
E1:G6        0.085793   0.196482   0.437  0.66245
E1:G7       -0.142570   0.201561  -0.707  0.47950
E1:G8        0.161424   0.214313   0.753  0.45147
E1:G9        0.265739   0.215626   1.232  0.21805
E1:G10       0.222877   0.213707   1.043  0.29721
E1:G11       0.144059   0.208320   0.692  0.48938
E1:G12       0.070244   0.201228   0.349  0.72709
E1:G13      -0.403895   0.200090  -2.019  0.04376 *
E1:G14       0.079018   0.218016   0.362  0.71709
E1:G15       0.147238   0.211657   0.696  0.48679
E1:G16      -0.265498   0.209426  -1.268  0.20515
E1:G17       0.286715   0.197160   1.454  0.14616
E1:G18      -0.104764   0.216764  -0.483  0.62897
E1:G19      -0.054890   0.216246  -0.254  0.79967
E1:G20      -0.338725   0.208875  -1.622  0.10515
E2:E3        0.081609   0.057500   1.419  0.15608
```

E2:E4	0.026467	0.056249	0.471	0.63806	
E2:G1	-0.266291	0.210841	-1.263	0.20685	
E2:G2	0.052239	0.217470	0.240	0.81021	
E2:G3	0.143648	0.217782	0.660	0.50964	
E2:G4	-0.087430	0.209053	-0.418	0.67586	
E2:G5	-0.195655	0.209047	-0.936	0.34950	
E2:G6	0.003312	0.201927	0.016	0.98692	
E2:G7	-0.216372	0.203633	-1.063	0.28820	
E2:G8	0.079130	0.214380	0.369	0.71211	
E2:G9	0.016594	0.220224	0.075	0.93995	
E2:G10	-0.308603	0.210874	-1.463	0.14362	
E2:G11	-0.012711	0.210102	-0.061	0.95177	
E2:G12	-0.009756	0.211642	-0.046	0.96324	
E2:G13	-0.247111	0.199177	-1.241	0.21498	
E2:G14	-0.220128	0.220322	-0.999	0.31795	
E2:G15	-0.262582	0.205639	-1.277	0.20189	
E2:G16	0.167631	0.215901	0.776	0.43766	
E2:G17	-0.012740	0.205885	-0.062	0.95067	
E2:G18	-0.199933	0.209758	-0.953	0.34071	
E2:G19	-0.326281	0.217248	-1.502	0.13340	
E2:G20	-0.215660	0.211394	-1.020	0.30786	
E3:E4	-0.023454	0.057656	-0.407	0.68424	
E3:G1	0.319058	0.205462	1.553	0.12073	
E3:G2	0.057486	0.211449	0.272	0.78577	
E3:G3	-0.400128	0.208734	-1.917	0.05549	.
E3:G4	-0.368543	0.211742	-1.741	0.08203	.
E3:G5	0.128747	0.215219	0.598	0.54982	
E3:G6	-0.294300	0.194934	-1.510	0.13138	
E3:G7	0.010784	0.200501	0.054	0.95711	
E3:G8	-0.415296	0.210906	-1.969	0.04918	*
E3:G9	0.160301	0.223238	0.718	0.47286	
E3:G10	0.126958	0.206610	0.614	0.53902	
E3:G11	0.392778	0.203613	1.929	0.05397	.
E3:G12	0.263416	0.213148	1.236	0.21677	
E3:G13	-0.088485	0.202235	-0.438	0.66180	
E3:G14	-0.169143	0.206569	-0.819	0.41306	
E3:G15	0.076668	0.211282	0.363	0.71677	
E3:G16	-0.130771	0.206619	-0.633	0.52692	
E3:G17	0.140398	0.205053	0.685	0.49368	
E3:G18	0.056236	0.212707	0.264	0.79153	
E3:G19	-0.132295	0.213340	-0.620	0.53531	
E3:G20	0.137717	0.206499	0.667	0.50496	
E4:G1	-0.222888	0.209637	-1.063	0.28791	
E4:G2	-0.023941	0.204237	-0.117	0.90670	
E4:G3	-0.266700	0.217800	-1.225	0.22101	
E4:G4	0.038736	0.202137	0.192	0.84806	
E4:G5	0.150654	0.201749	0.747	0.45537	
E4:G6	-0.123826	0.200358	-0.618	0.53668	
E4:G7	0.059561	0.200491	0.297	0.76646	
E4:G8	0.184265	0.208316	0.885	0.37659	
E4:G9	-0.128525	0.222051	-0.579	0.56283	
E4:G10	0.008170	0.210099	0.039	0.96899	
E4:G11	-0.285260	0.211757	-1.347	0.17821	
E4:G12	-0.277111	0.208268	-1.331	0.18360	
E4:G13	-0.174676	0.200002	-0.873	0.38264	
E4:G14	-0.120377	0.216693	-0.556	0.57865	
E4:G15	0.192715	0.209457	0.920	0.35773	
E4:G16	-0.460972	0.211508	-2.179	0.02950	*
E4:G17	0.101096	0.204328	0.495	0.62086	
E4:G18	-0.247922	0.204511	-1.212	0.22566	

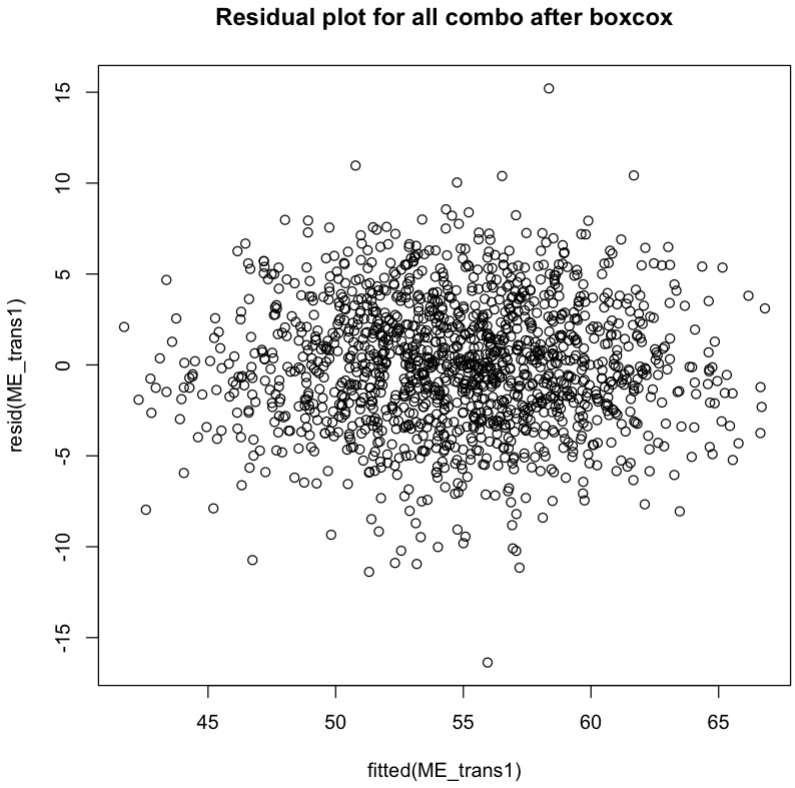
E4:G19	−0.210605	0.219215	−0.961	0.33689	
E4:G20	0.185066	0.209071	0.885	0.37624	
G1:G2	0.871310	0.774818	1.125	0.26102	
G1:G3	0.282031	0.845685	0.333	0.73882	
G1:G4	0.395271	0.758826	0.521	0.60254	
G1:G5	1.507963	0.797533	1.891	0.05890	.
G1:G6	−0.740253	0.738942	−1.002	0.31666	
G1:G7	0.076959	0.752681	0.102	0.91858	
G1:G8	−0.048630	0.811615	−0.060	0.95223	
G1:G9	−0.805313	0.789126	−1.021	0.30770	
G1:G10	0.062064	0.790804	0.078	0.93746	
G1:G11	−0.307537	0.796702	−0.386	0.69956	
G1:G12	1.244721	0.779999	1.596	0.11081	
G1:G13	−0.897086	0.763591	−1.175	0.24031	
G1:G14	0.461438	0.783779	0.589	0.55616	
G1:G15	0.648029	0.791371	0.819	0.41303	
G1:G16	−0.724386	0.788493	−0.919	0.35845	
G1:G17	0.029137	0.747115	0.039	0.96890	
G1:G18	−0.243806	0.810196	−0.301	0.76353	
G1:G19	−0.468573	0.812989	−0.576	0.56449	
G1:G20	0.696706	0.751369	0.927	0.35399	
G2:G3	−0.462666	0.805473	−0.574	0.56581	
G2:G4	−0.345367	0.737985	−0.468	0.63988	
G2:G5	1.060580	0.759399	1.397	0.16280	
G2:G6	−1.156406	0.738023	−1.567	0.11741	
G2:G7	0.586839	0.679587	0.864	0.38803	
G2:G8	−0.992358	0.724174	−1.370	0.17085	
G2:G9	1.236766	0.787644	1.570	0.11664	
G2:G10	−0.667406	0.728386	−0.916	0.35971	
G2:G11	−0.070097	0.744219	−0.094	0.92498	
G2:G12	−0.086597	0.719730	−0.120	0.90425	
G2:G13	−1.508038	0.708426	−2.129	0.03349	*
G2:G14	0.304179	0.793551	0.383	0.70156	
G2:G15	0.475698	0.823379	0.578	0.56355	
G2:G16	1.871625	0.832567	2.248	0.02476	*
G2:G17	1.294543	0.752326	1.721	0.08557	.
G2:G18	−0.542194	0.725639	−0.747	0.45510	
G2:G19	0.002118	0.799475	0.003	0.99789	
G2:G20	0.785997	0.717035	1.096	0.27323	
G3:G4	−0.314809	0.735892	−0.428	0.66888	
G3:G5	−0.651863	0.778795	−0.837	0.40276	
G3:G6	−0.171908	0.703730	−0.244	0.80706	
G3:G7	1.360547	0.741651	1.834	0.06684	.
G3:G8	0.726556	0.779552	0.932	0.35152	
G3:G9	−0.312994	0.793755	−0.394	0.69342	
G3:G10	−0.327986	0.794245	−0.413	0.67972	
G3:G11	1.228846	0.766305	1.604	0.10908	
G3:G12	−0.272116	0.725730	−0.375	0.70776	
G3:G13	−0.352412	0.760839	−0.463	0.64332	
G3:G14	0.399023	0.774305	0.515	0.60642	
G3:G15	0.401807	0.756211	0.531	0.59528	
G3:G16	0.913151	0.778265	1.173	0.24091	
G3:G17	−0.866390	0.762588	−1.136	0.25614	
G3:G18	0.981777	0.736205	1.334	0.18261	
G3:G19	−0.127836	0.830993	−0.154	0.87777	
G3:G20	−0.833914	0.746053	−1.118	0.26390	
G4:G5	1.203627	0.725863	1.658	0.09755	.
G4:G6	0.236405	0.703801	0.336	0.73701	
G4:G7	1.358237	0.721490	1.883	0.06001	.
G4:G8	−0.426511	0.738707	−0.577	0.56380	

G4:G9	−0.597119	0.838954	−0.712	0.47677	
G4:G10	0.641409	0.724651	0.885	0.37627	
G4:G11	0.303094	0.750757	0.404	0.68650	
G4:G12	−0.732570	0.703923	−1.041	0.29823	
G4:G13	0.706736	0.714338	0.989	0.32270	
G4:G14	1.310656	0.783726	1.672	0.09473	.
G4:G15	−0.714362	0.722003	−0.989	0.32267	
G4:G16	−0.419578	0.731809	−0.573	0.56652	
G4:G17	−1.587776	0.756091	−2.100	0.03595	*
G4:G18	−0.226977	0.760609	−0.298	0.76544	
G4:G19	1.285948	0.775573	1.658	0.09758	.
G4:G20	0.229276	0.753062	0.304	0.76083	
G5:G6	0.394154	0.711464	0.554	0.57968	
G5:G7	−0.179623	0.735252	−0.244	0.80704	
G5:G8	0.471138	0.741977	0.635	0.52557	
G5:G9	−0.381696	0.856181	−0.446	0.65582	
G5:G10	−0.930272	0.758327	−1.227	0.22017	
G5:G11	0.229619	0.808695	0.284	0.77651	
G5:G12	−0.028045	0.790241	−0.035	0.97170	
G5:G13	−0.471648	0.730455	−0.646	0.51861	
G5:G14	1.228605	0.790464	1.554	0.12039	
G5:G15	0.670869	0.759413	0.883	0.37720	
G5:G16	−0.216695	0.738719	−0.293	0.76932	
G5:G17	0.231297	0.725540	0.319	0.74994	
G5:G18	−0.296673	0.761165	−0.390	0.69679	
G5:G19	0.109572	0.792280	0.138	0.89003	
G5:G20	0.548462	0.797097	0.688	0.49154	
G6:G7	−0.714156	0.701187	−1.018	0.30866	
G6:G8	0.865253	0.747326	1.158	0.24719	
G6:G9	0.162246	0.719193	0.226	0.82156	
G6:G10	−0.036990	0.729447	−0.051	0.95957	
G6:G11	0.861391	0.766406	1.124	0.26127	
G6:G12	0.520845	0.700799	0.743	0.45750	
G6:G13	−0.478568	0.685773	−0.698	0.48541	
G6:G14	0.082221	0.765478	0.107	0.91448	
G6:G15	−0.370092	0.713487	−0.519	0.60406	
G6:G16	−1.018211	0.673684	−1.511	0.13096	
G6:G17	1.440150	0.700314	2.056	0.03997	*
G6:G18	0.279886	0.718307	0.390	0.69687	
G6:G19	−1.168987	0.757641	−1.543	0.12312	
G6:G20	0.921054	0.818125	1.126	0.26048	
G7:G8	0.205168	0.711796	0.288	0.77322	
G7:G9	−0.268140	0.744468	−0.360	0.71878	
G7:G10	−0.764779	0.694749	−1.101	0.27121	
G7:G11	1.227115	0.735688	1.668	0.09559	.
G7:G12	−0.612113	0.708403	−0.864	0.38773	
G7:G13	0.408008	0.703824	0.580	0.56223	
G7:G14	0.771985	0.732693	1.054	0.29227	
G7:G15	−1.064341	0.735562	−1.447	0.14818	
G7:G16	0.146637	0.738589	0.199	0.84266	
G7:G17	0.439340	0.732598	0.600	0.54882	
G7:G18	0.385466	0.744715	0.518	0.60484	
G7:G19	−0.399264	0.810078	−0.493	0.62220	
G7:G20	−0.685999	0.714082	−0.961	0.33692	
G8:G9	−0.739685	0.855377	−0.865	0.38736	
G8:G10	0.427789	0.722224	0.592	0.55375	
G8:G11	−0.567869	0.798911	−0.711	0.47735	
G8:G12	0.799306	0.761533	1.050	0.29412	
G8:G13	−0.634749	0.738113	−0.860	0.38999	
G8:G14	−0.246925	0.799308	−0.309	0.75744	

G8:G15	−0.382180	0.826593	−0.462	0.64391	
G8:G16	0.763768	0.727231	1.050	0.29383	
G8:G17	0.581182	0.802271	0.724	0.46895	
G8:G18	−0.570588	0.752940	−0.758	0.44872	
G8:G19	1.111780	0.820625	1.355	0.17575	
G8:G20	−0.830991	0.782609	−1.062	0.28854	
G9:G10	1.510104	0.774543	1.950	0.05146	.
G9:G11	0.917560	0.847742	1.082	0.27932	
G9:G12	−0.013077	0.861103	−0.015	0.98789	
G9:G13	0.328621	0.800443	0.411	0.68148	
G9:G14	−0.158392	0.818288	−0.194	0.84655	
G9:G15	0.752134	0.789063	0.953	0.34069	
G9:G16	1.397240	0.757914	1.844	0.06551	.
G9:G17	−0.230476	0.800088	−0.288	0.77335	
G9:G18	−0.416819	0.781515	−0.533	0.59390	
G9:G19	1.304933	0.812269	1.607	0.10843	
G9:G20	1.700447	0.817667	2.080	0.03778	*
G10:G11	−0.003783	0.756856	−0.005	0.99601	
G10:G12	0.395375	0.716930	0.551	0.58141	
G10:G13	−0.678938	0.788364	−0.861	0.38931	
G10:G14	0.455389	0.879610	0.518	0.60476	
G10:G15	0.570437	0.760027	0.751	0.45308	
G10:G16	−0.570207	0.740602	−0.770	0.44150	
G10:G17	−0.104021	0.713992	−0.146	0.88419	
G10:G18	−0.103494	0.741231	−0.140	0.88898	
G10:G19	−0.321222	0.745785	−0.431	0.66676	
G10:G20	0.041672	0.730478	0.057	0.95452	
G11:G12	−1.382729	0.767846	−1.801	0.07200	.
G11:G13	−0.429656	0.806058	−0.533	0.59411	
G11:G14	−0.291885	0.822492	−0.355	0.72275	
G11:G15	−0.259055	0.776251	−0.334	0.73865	
G11:G16	1.011401	0.739032	1.369	0.17141	
G11:G17	0.692593	0.792747	0.874	0.38248	
G11:G18	0.258940	0.727095	0.356	0.72181	
G11:G19	0.294047	0.793849	0.370	0.71115	
G11:G20	−1.264490	0.770450	−1.641	0.10102	
G12:G13	−0.062575	0.761840	−0.082	0.93455	
G12:G14	2.472552	0.750623	3.294	0.00102	**
G12:G15	0.799303	0.750260	1.065	0.28693	
G12:G16	−0.685777	0.760739	−0.901	0.36753	
G12:G17	−0.064238	0.748736	−0.086	0.93164	
G12:G18	1.745763	0.774564	2.254	0.02439	*
G12:G19	−0.863029	0.833943	−1.035	0.30094	
G12:G20	0.601307	0.762158	0.789	0.43030	
G13:G14	−0.903858	0.804203	−1.124	0.26128	
G13:G15	0.044323	0.731467	0.061	0.95169	
G13:G16	0.296579	0.739128	0.401	0.68831	
G13:G17	−1.023345	0.707475	−1.446	0.14832	
G13:G18	0.474423	0.805362	0.589	0.55592	
G13:G19	0.125578	0.826454	0.152	0.87925	
G13:G20	−0.981940	0.768734	−1.277	0.20174	
G14:G15	0.126130	0.793443	0.159	0.87372	
G14:G16	1.375356	0.784540	1.753	0.07986	.
G14:G17	−0.510535	0.757174	−0.674	0.50028	
G14:G18	0.517346	0.846644	0.611	0.54128	
G14:G19	−0.692168	0.778084	−0.890	0.37388	
G14:G20	0.542101	0.822062	0.659	0.50974	
G15:G16	−1.084074	0.808226	−1.341	0.18009	
G15:G17	0.472406	0.764805	0.618	0.53691	
G15:G18	0.736354	0.750487	0.981	0.32672	

```
G15:G19     -1.204674   0.795043  -1.515  0.12999
G15:G20     -0.369351   0.841196  -0.439  0.66069
G16:G17      0.890621   0.748341   1.190  0.23424
G16:G18     -1.971363   0.788011  -2.502  0.01250 *
G16:G19      0.695748   0.805056   0.864  0.38765
G16:G20      0.244996   0.778354   0.315  0.75300
G17:G18     -1.095332   0.781857  -1.401  0.16150
G17:G19     -0.254526   0.764270  -0.333  0.73917
G17:G20      0.197585   0.759090   0.260  0.79469
G18:G19     -0.752231   0.818549  -0.919  0.35830
G18:G20      0.863510   0.776556   1.112  0.26638
G19:G20      0.161099   0.794043   0.203  0.83926
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.998 on 1152 degrees of freedom
Multiple R-squared:  0.6165,    Adjusted R-squared:  0.5166
F-statistic: 6.173 on 300 and 1152 DF,  p-value: < 2.2e-16
```

In [13]:
```
plot(resid(ME_trans1) ~ fitted(ME_trans1), main = 'Residual plot for all combo
```



In [14]:
```
summary(ME)$adj.r.square
summary(ME_trans1)$adj.r.square
```

```
0.459197679925302
0.516641668142231
```

In [15]:
```
# install.packages('IRkernel')
# IRkernel::installspec(name = 'ir35', displayname = 'R 4.0.0')
install.packages('leaps', dependencies=TRUE, repos='http://cran.rstudio.com/')
library(leaps)
MF <- regsubsets(model.matrix(ME_trans1)[,-1], I((data$Y)^1.5),
                 nbest = 1, nvmax = 5,
```

```
                 method = 'forward', intercept = TRUE)
temp <- summary(MF)
```

```
The downloaded binary packages are in
        /var/folders/rh/bp1_zwxj76z6nrtwmrl525600000gn/T//Rtmpgy5Fms/downloade
d_packages
```

In [16]:
```
install.packages('knitr', dependencies=TRUE, repos='http://cran.rstudio.com/')
library(knitr)
```

```
The downloaded binary packages are in
        /var/folders/rh/bp1_zwxj76z6nrtwmrl525600000gn/T//Rtmpgy5Fms/downloade
d_packages
```

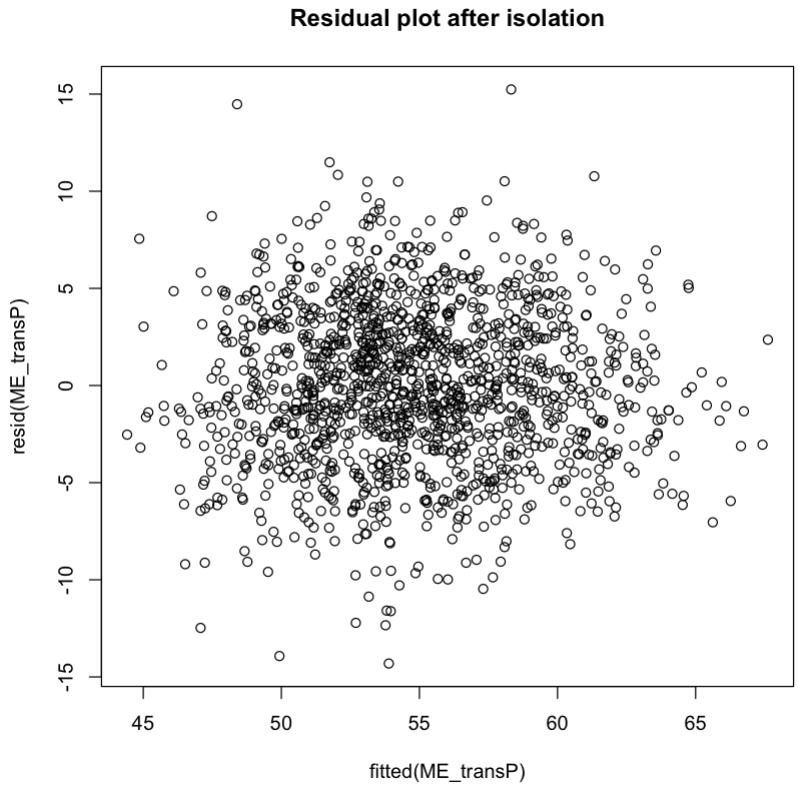
In [17]:
```
var <- colnames(model.matrix(ME_trans1))
M_test <- apply(temp$which, 1,
                function(x) paste0(var[x], collapse='+'))
kable(data.frame(cbind(model = M_test, adjR2 = temp$adjr2, BIC = temp$bic)),
     caption = 'test model summary')
```

```
Table: test model summary

|model                                    |adjR2             |BIC
|
|:----------------------------------------|:-----------------|:---------------
--|
|(Intercept)+E3:E4                        |0.367932723119669 |-653.01572806946
2 |
|(Intercept)+E1:E4+E3:E4                  |0.445102792810913 |-835.93548948907
7 |
|(Intercept)+E1:E4+E3:E4+G2:G18           |0.494246644552623 |-964.39869167400
7 |
|(Intercept)+E3+E1:E4+E3:E4+G2:G18        |0.502895079781164 |-983.18175631410
1 |
|(Intercept)+E3+E1:E4+E3:E4+G2:G18+G3:G18 |0.506137651719943 |-986.41300881394
1 |
```

In [18]:
```
ME_transP <- lm(I(Y^optimalLambda) ~ (E3+(E1:E4)+(E3:E4)+(G2:G18)+(G3:G18)), da
summary(ME_transP)
```

```
Call:
lm(formula = I(Y^optimalLambda) ~ (E3 + (E1:E4) + (E3:E4) + (G2:G18) +
    (G3:G18)), data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-14.309  -2.627  -0.046   2.821  15.238

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 33.295603   0.799168  41.663  < 2e-16 ***
E3           0.654793   0.124803   5.247 1.78e-07 ***
E1:E4        0.149748   0.009377  15.970  < 2e-16 ***
E3:E4        0.112209   0.013177   8.515  < 2e-16 ***
G2:G18       2.267694   0.247902   9.148  < 2e-16 ***
G18:G3       0.807799   0.249207   3.241  0.00122 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.041 on 1447 degrees of freedom
Multiple R-squared:  0.5078,    Adjusted R-squared:  0.5061
F-statistic: 298.6 on 5 and 1447 DF,  p-value: < 2.2e-16
```

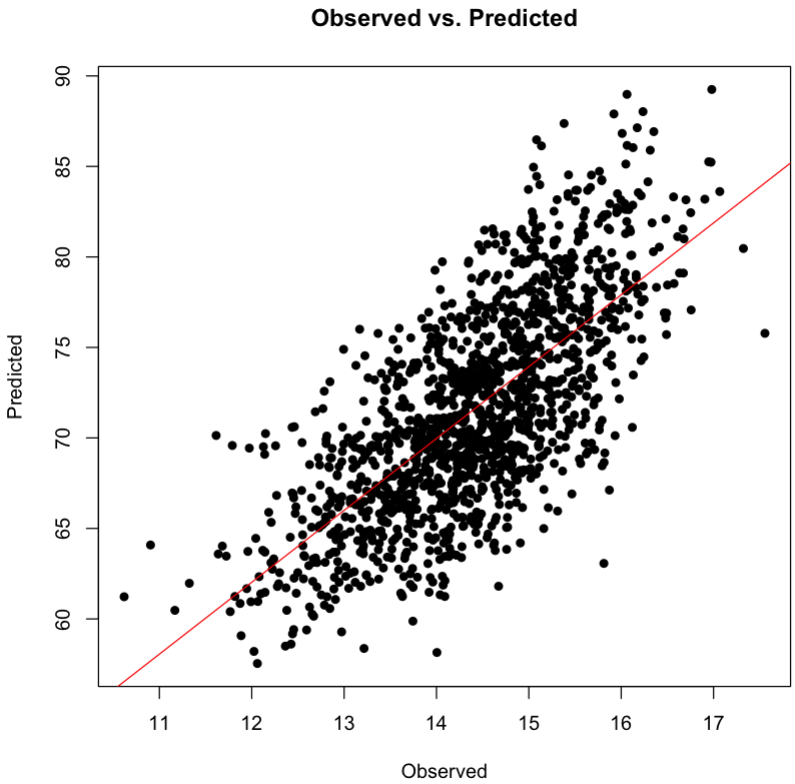
In [19]:
```
plot(resid(ME_transP) ~ fitted(ME_transP), main = 'Residual plot after isolati
```



In [20]:
```
data$E1_E4 <- data$E1 * data$E4
data$E3_E4 <- data$E3 * data$E4
data$G2_G18 <- data$G2 * data$G18
data$G3_G18 <- data$G3 * data$G18

#MEPlot <- lm(Y ~ (E3 + E1_E4 + E3_E4 + G2_G18 + G3_G18), data = data)

MEPlotNoT <- lm(Y^1.6 ~ (E3+E1:E4+E3:E4+G2:G18), data = data)
```

```
summary(MEPlotNoT)

data$predicted <- predict(MEPlotNoT)

plot(data$Y, data$predicted, main = "Observed vs. Predicted", xlab = "Observed"

abline(lm(data$predicted ~ data$Y), col = "red")

abline(0, 1, col = "blue", lty = 2)
```

```
Call:
lm(formula = Y^1.6 ~ (E3 + E1:E4 + E3:E4 + G2:G18), data = data)

Residuals:
     Min       1Q   Median       3Q      Max 
-19.5538  -3.8397  -0.0525   3.9236  22.1887 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 42.11525    1.10632  38.068  < 2e-16 ***
E3           0.89146    0.17410   5.120 3.46e-07 ***
E1:E4        0.20795    0.01309  15.889  < 2e-16 ***
E3:E4        0.15856    0.01838   8.626  < 2e-16 ***
G2:G18       3.65739    0.30888  11.841  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.641 on 1448 degrees of freedom
Multiple R-squared:  0.5047,    Adjusted R-squared:  0.5033 
F-statistic: 368.8 on 4 and 1448 DF,  p-value: < 2.2e-16
```



In [21]:
```
# data$log_E1 <- log(data$E1)
# data$log_E2 <- log(data$E2)
# data$log_E3 <- log(data$E3)
```

```r
# data$log_E4 <- log(data$E4)
# data$log_G2 <- log(data$G2)
# data$log_G18 <- log(data$G18)
# data$log_G3 <- log(data$G3)

data$sqrt_E1 <- sqrt(data$E1)
data$sqrt_E2 <- sqrt(data$E2)
data$sqrt_E3 <- sqrt(data$E3)
data$sqrt_E4 <- sqrt(data$E4)
data$sqrt_G2 <- sqrt(data$G2)
data$sqrt_G18 <- sqrt(data$G18)
data$sqrt_G3 <- sqrt(data$G3)

# data$log_E1 <- (data$E1)^1/3
# data$log_E2 <- (data$E2)^1/3
# data$log_E3 <- (data$E3)^1/3
# data$log_E4 <- (data$E4)^1/3
# data$log_G2 <- (data$G2)^1/3
# data$log_G18 <- (data$G18)^1/3
# data$log_G3 <- (data$G3)^1/3

data$sqrt_E1_E4 <- data$sqrt_E1 * data$sqrt_E4
data$sqrt_E3_E4 <- data$sqrt_E3 * data$sqrt_E4

data$sqrt_G2_G18 <- data$sqrt_G2 * data$sqrt_G18
data$sqrt_G3_G18 <- data$sqrt_G3 * data$sqrt_G18
data$Y2.6 <- (data$Y)^1.6
#E3+E1:E4+E3:E4+G2:G18
#MEPlot <- lm(Y^3.5 ~ (log(E3) + log_E1_E4 + log_E3_E4 + G2:G18), data = data)
MEPlot <- lm(Y^1.6 ~ (sqrt(E3) + sqrt_E1_E4 + sqrt_E3_E4 + G2:G18), data = data

data$predicted <- predict(MEPlot)

plot(data$Y2.6, data$predicted, main = "Observed vs. Predicted", xlab = "Obser

abline(lm(data$predicted ~ data$Y2.6), col = "red")

abline(0, 1, col = "blue", lty = 2)

# summary(MEPlot)
# temp <- summary(MEPlot)
# kable(temp$coefficients[ abs(temp$coefficients[,4]) <= 0.001, ], caption='Sig
# kable(temp$coefficients[ abs(temp$coefficients[,3]) >= 4, ], caption='Sig Coe

temp <- summary(MEPlotNoT)
kable(temp$coefficients[ abs(temp$coefficients[,4]) <= 0.001, ], caption='Sig (
kable(temp$coefficients[ abs(temp$coefficients[,3]) >= 4, ], caption='Sig Coef

confint(MEPlotNoT, level = 0.90)
confint(MEPlotNoT, level = 0.95)
confint(MEPlotNoT, level = 0.99)
```

Table: Sig Coefficients

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	42.1152493	1.1063232	38.067763	0e+00
E3	0.8914593	0.1741044	5.120258	3e-07
E1:E4	0.2079469	0.0130877	15.888711	0e+00
E3:E4	0.1585580	0.0183806	8.626372	0e+00
G2:G18	3.6573878	0.3088809	11.840772	0e+00

Table: Sig Coefficients

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	42.1152493	1.1063232	38.067763	0e+00
E3	0.8914593	0.1741044	5.120258	3e-07
E1:E4	0.2079469	0.0130877	15.888711	0e+00
E3:E4	0.1585580	0.0183806	8.626372	0e+00
G2:G18	3.6573878	0.3088809	11.840772	0e+00

A matrix: 5 × 2 of type dbl

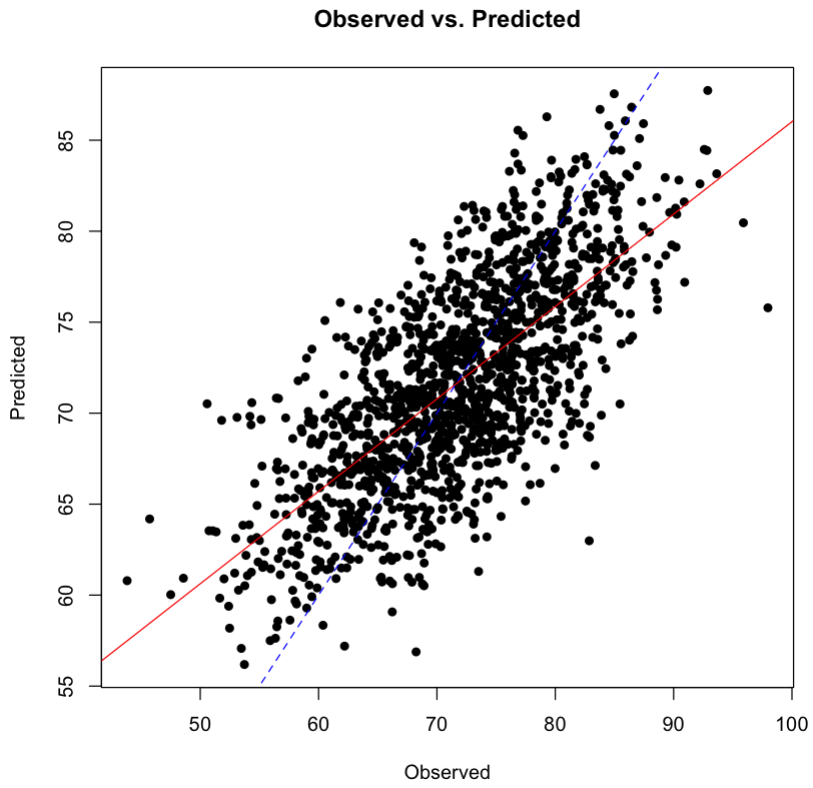
	5 %	95 %
(Intercept)	40.2943446	43.9361540
E3	0.6048998	1.1780188
E1:E4	0.1864058	0.2294881
E3:E4	0.1283052	0.1888108
G2:G18	3.1489988	4.1657769

A matrix: 5 × 2 of type dbl

	2.5 %	97.5 %
(Intercept)	39.9450817	44.2854169
E3	0.5499355	1.2329831
E1:E4	0.1822740	0.2336198
E3:E4	0.1225025	0.1946135
G2:G18	3.0514860	4.2632897

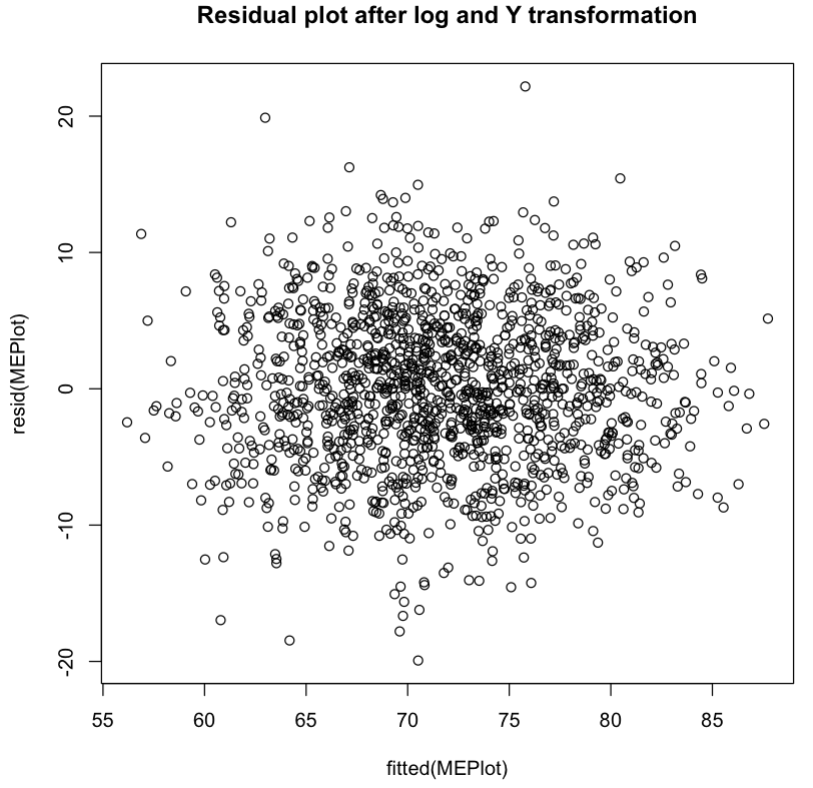
A matrix: 5 × 2 of type dbl

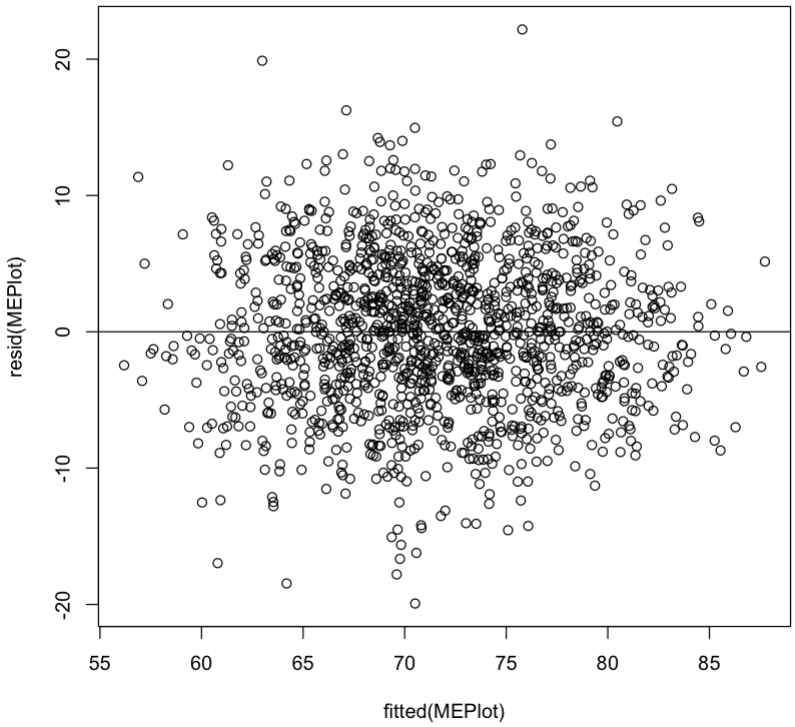
	0.5 %	99.5 %
(Intercept)	39.2617885	44.9687101
E3	0.4424043	1.3405143
E1:E4	0.1741907	0.2417031
E3:E4	0.1111502	0.2059658
G2:G18	2.8607134	4.4540623



Bonferroni's inequality applies here, multiplying each of the coefficients p-value by 24, we see that the value is still less than the recommended 0.001.

In [22]:
```
plot(resid(MEPlot) ~ fitted(MEPlot), main = 'Residual plot after log and Y tra
plot(fitted(MEPlot), resid(MEPlot), abline(0,0))
```

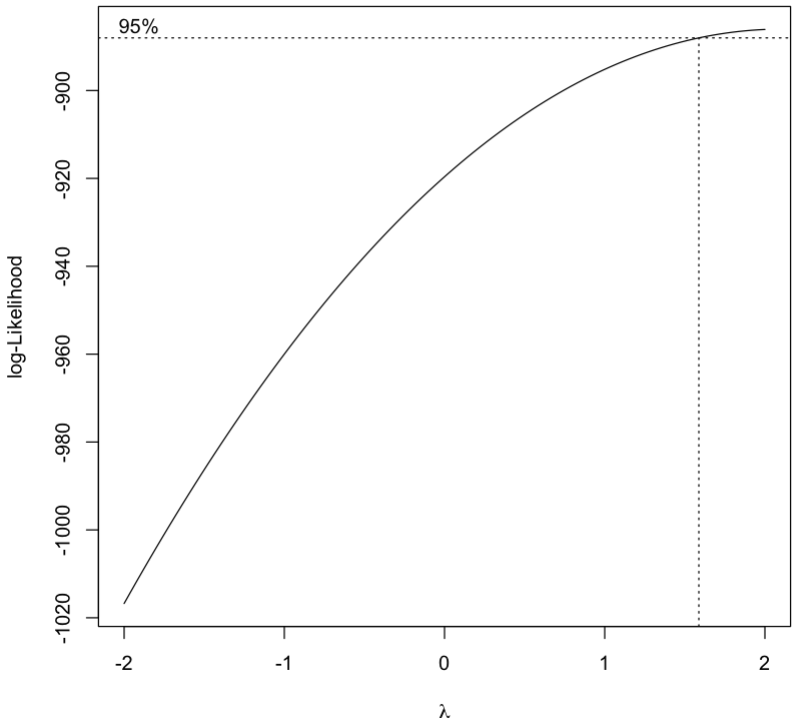




End test for best model accounting for interactions

In [23]:
```
summary(ME_testNoI)
library(MASS)
transformation <- boxcox(ME_testNoI)
```

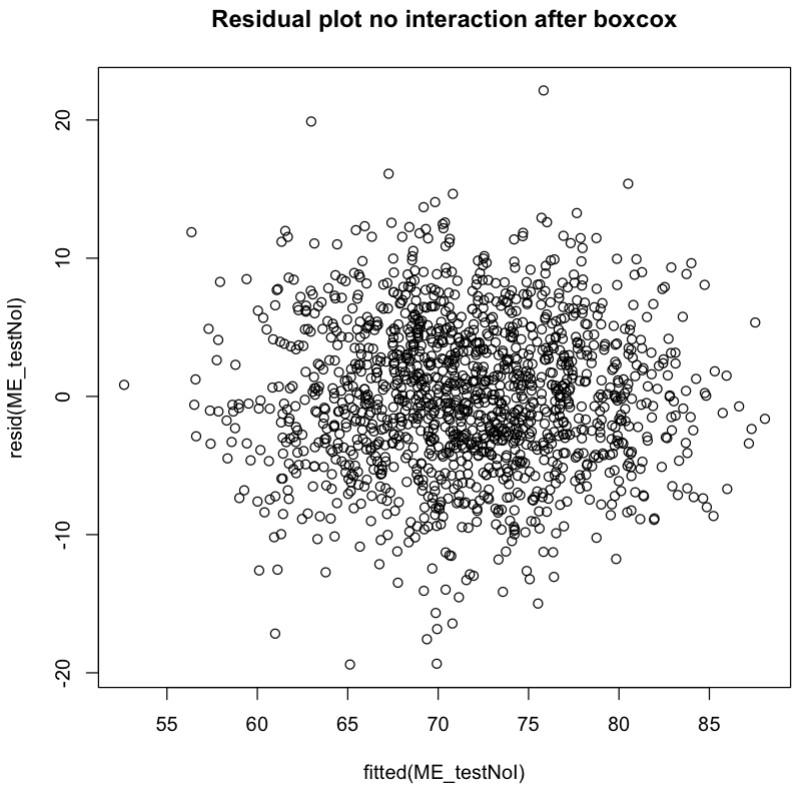
```
Call:
lm(formula = Y ~ E1 + E2 + E3 + E4 + G1 + G2 + G3 + G4 + G5 +
    G6 + G7 + G8 + G9 + G10 + G11 + G12 + G13 + G14 + G15 + G16 +
    G17 + G18 + G19 + G20, data = data)

Residuals:
     Min       1Q   Median       3Q      Max
-2.67786 -0.45085 -0.00822  0.48742  2.61488

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.813262   0.251267  31.096   <2e-16 ***
E1           0.196336   0.012862  15.265   <2e-16 ***
E2          -0.009078   0.013080  -0.694   0.4878
E3           0.265576   0.012907  20.576   <2e-16 ***
E4           0.352108   0.012726  27.669   <2e-16 ***
G1          -0.045627   0.046661  -0.978   0.3283
G2           0.399761   0.046523   8.593   <2e-16 ***
G3           0.065475   0.047564   1.377   0.1689
G4          -0.014444   0.045790  -0.315   0.7525
G5          -0.015411   0.046707  -0.330   0.7415
G6          -0.069026   0.044195  -1.562   0.1185
G7           0.014784   0.044744   0.330   0.7411
G8          -0.096178   0.046750  -2.057   0.0398 *
G9           0.051281   0.048592   1.055   0.2914
G10         -0.043765   0.046262  -0.946   0.3443
G11         -0.046808   0.047078  -0.994   0.3203
G12          0.053096   0.045923   1.156   0.2478
G13          0.044876   0.045543   0.985   0.3246
G14          0.046861   0.047816   0.980   0.3272
G15         -0.019236   0.046835  -0.411   0.6813
G16          0.038413   0.046420   0.827   0.4081
G17         -0.092722   0.045464  -2.039   0.0416 *
G18          0.452659   0.046249   9.787   <2e-16 ***
G19          0.012653   0.048676   0.260   0.7949
G20         -0.015588   0.046849  -0.333   0.7394
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7043 on 1428 degrees of freedom
Multiple R-squared:  0.526,     Adjusted R-squared:  0.518
F-statistic: 66.02 on 24 and 1428 DF,  p-value: < 2.2e-16
```



In [24]:
```
ME_testNoI <- lm(Y^1.6 ~ E1+E2+E3+E4+G1+G2+G3+G4+G5+G6+G7+G8+G9+G10+G11+G12+G1
summary(ME_testNoI)
```

```
Call:
lm(formula = Y^1.6 ~ E1 + E2 + E3 + E4 + G1 + G2 + G3 + G4 +
    G5 + G6 + G7 + G8 + G9 + G10 + G11 + G12 + G13 + G14 + G15 +
    G16 + G17 + G18 + G19 + G20, data = data)

Residuals:
     Min       1Q   Median       3Q      Max
-19.3996  -3.5511  -0.1492   3.8278  22.1371

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 19.39643    1.97978   9.797   <2e-16 ***
E1           1.55853    0.10134  15.379   <2e-16 ***
E2          -0.06375    0.10306  -0.619   0.5363
E3           2.09957    0.10169  20.646   <2e-16 ***
E4           2.78192    0.10027  27.745   <2e-16 ***
G1          -0.35305    0.36765  -0.960   0.3371
G2           3.12686    0.36657   8.530   <2e-16 ***
G3           0.52584    0.37476   1.403   0.1608
G4          -0.12568    0.36079  -0.348   0.7276
G5          -0.11093    0.36801  -0.301   0.7631
G6          -0.56544    0.34822  -1.624   0.1046
G7           0.11920    0.35254   0.338   0.7353
G8          -0.76724    0.36835  -2.083   0.0374 *
G9           0.40882    0.38287   1.068   0.2858
G10         -0.33092    0.36451  -0.908   0.3641
G11         -0.34401    0.37094  -0.927   0.3539
G12          0.42758    0.36184   1.182   0.2375
G13          0.35096    0.35885   0.978   0.3282
G14          0.36590    0.37676   0.971   0.3316
G15         -0.15104    0.36902  -0.409   0.6824
G16          0.28532    0.36575   0.780   0.4355
G17         -0.72138    0.35822  -2.014   0.0442 *
G18          3.54989    0.36440   9.742   <2e-16 ***
G19          0.07436    0.38353   0.194   0.8463
G20         -0.12174    0.36913  -0.330   0.7416
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.549 on 1428 degrees of freedom
Multiple R-squared:  0.5272,    Adjusted R-squared:  0.5193
F-statistic: 66.36 on 24 and 1428 DF,  p-value: < 2.2e-16
```

In [25]:
```
plot(resid(ME_testNoI) ~ fitted(ME_testNoI), main = 'Residual plot no interact
```



In [26]:
```
install.packages('leaps', dependencies=TRUE, repos='http://cran.rstudio.com/')
library(leaps)
MF <- regsubsets(model.matrix(ME_testNoI)[,-1], I((data$Y)^1.6),
                 nbest = 1, nvmax = 5,
                 method = 'forward', intercept = TRUE)
temp <- summary(MF)
```

```
The downloaded binary packages are in
        /var/folders/rh/bp1_zwxj76z6nrtwmrl525600000gn/T//Rtmpgy5Fms/downloade
d_packages
```

In [27]:
```
install.packages('knitr', dependencies=TRUE, repos='http://cran.rstudio.com/')
library(knitr)
```

```
The downloaded binary packages are in
        /var/folders/rh/bp1_zwxj76z6nrtwmrl525600000gn/T//Rtmpgy5Fms/downloade
d_packages
```

In [28]:
```
var <- colnames(model.matrix(ME_testNoI))
M_test <- apply(temp$which, 1,
                function(x) paste0(var[x], collapse='+'))
kable(data.frame(cbind(model = M_test, adjR2 = temp$adjr2, BIC = temp$bic)),
     caption = 'test model summary')
```

Table: test model summary

model	adjR2	BIC
(Intercept)+E4	0.244386580951173	-393.605745152781
(Intercept)+E3+E4	0.38432295360598	-684.911154212549
(Intercept)+E1+E3+E4	0.461373765890979	-872.899247373236
(Intercept)+E1+E3+E4+G18	0.494300860591575	-958.276180590858
(Intercept)+E1+E3+E4+G2+G18	0.518803625687692	-1024.16395651035

In [29]:
```
AMT <- lm(Y^1.6 ~ E1+E3+E4+G2+G18, data = data)
temp <- summary(AMT)
kable(temp$coefficients[ abs(temp$coefficients[,4]) <= 0.001, ], caption='Sig (

#Not needed.
```

Table: Sig Coefficients

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	18.058343	1.3971536	12.925095	0
E1	1.561913	0.1005219	15.538035	0
E3	2.104107	0.1009883	20.835154	0
E4	2.781438	0.0999332	27.832964	0
G2	3.145662	0.3638783	8.644819	0
G18	3.549359	0.3614085	9.820907	0

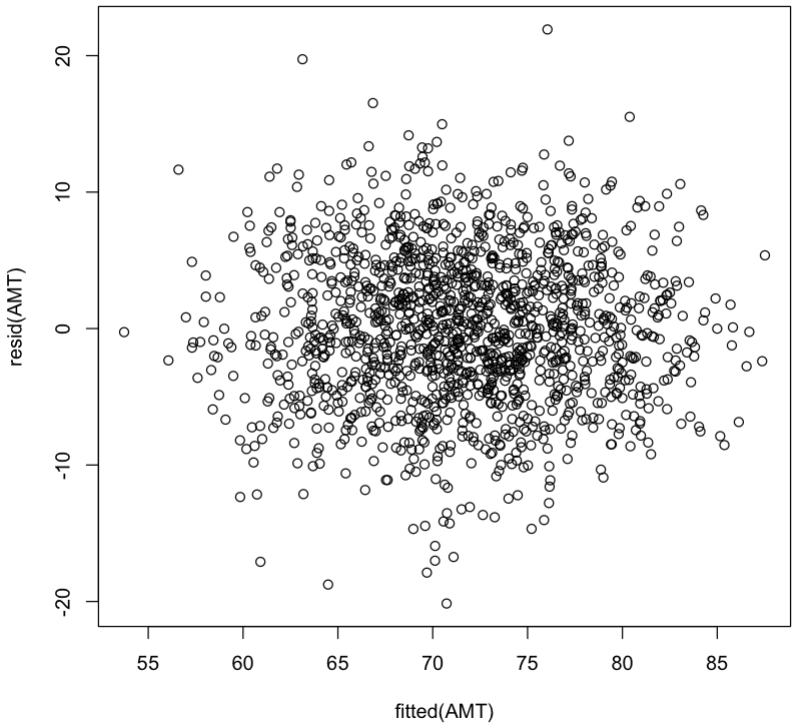
In [30]:
```
AMT_withI <- lm( Y^1.6 ~ (E1+E3+E4+G2+G18)^2, data=data)
temp <- summary(AMT_withI)
kable(temp$coefficients[ abs(temp$coefficients[,3]) >= 2, ])

temp <- summary(MEPlot)
kable(temp$coefficients[ abs(temp$coefficients[,3]) >= 4, ])
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	16.069150	7.9482218	2.021729	0.0433894
E3	1.675016	0.8028940	2.086223	0.0371344
E4	3.308664	0.8141968	4.063716	0.0000509
G2	7.227341	3.4783300	2.077819	0.0379032
G18	6.838992	3.3412310	2.046848	0.0408546

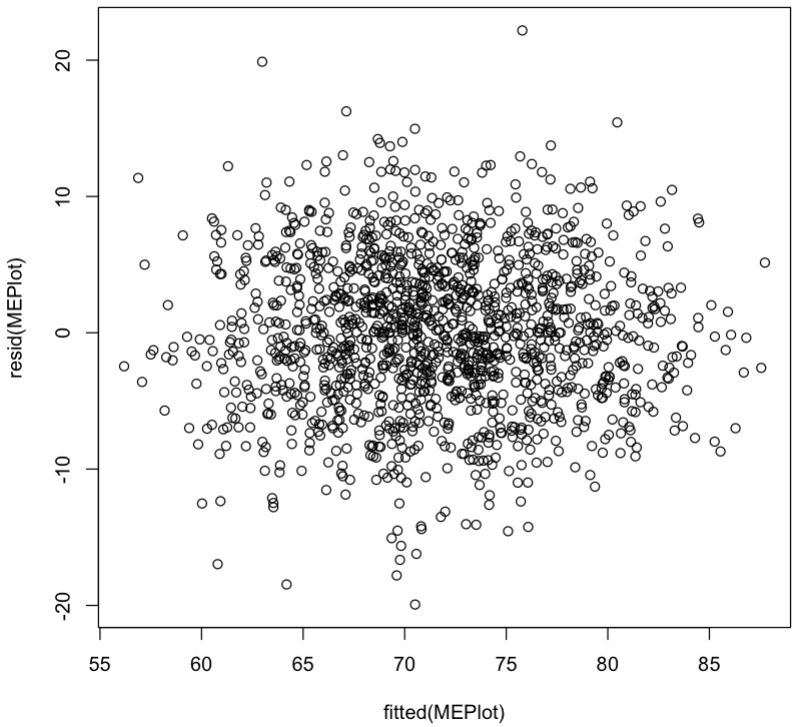
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	15.362947	2.1705785	7.077812	0e+00
sqrt(E3)	4.751377	0.9594144	4.952372	8e-07
sqrt_E1_E4	3.127764	0.2002088	15.622509	0e+00
sqrt_E3_E4	2.406085	0.2817205	8.540682	0e+00
G2:G18	3.650070	0.3078515	11.856594	0e+00

In [31]:
```
plot(resid(AMT) ~ fitted(AMT), main = 'Residual plot no interaction after boxc(
plot(resid(MEPlot) ~ fitted(MEPlot), main = 'Residual plot of interaction and
```

Residual plot no interaction after boxcox and variable selection



Residual plot of interaction and transformations



Our previous model including the interactions seems to be correct, since a model without interactions selected only variables that were in the model with interactions,

meaning that these variables are statistically significant.

In [32]:
```
testM <- lm(Y^1.6 ~ (E1 + E3 + E4 + G2:G18), data = data)
summary(testM)
```

```
Call:
lm(formula = Y^1.6 ~ (E1 + E3 + E4 + G2:G18), data = data)

Residuals:
     Min       1Q   Median       3Q      Max 
-20.1238  -3.7734  -0.0747   3.9101  22.4470 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  21.0326     1.3624   15.44   <2e-16 ***
E1            1.5709     0.1016   15.47   <2e-16 ***
E3            2.0898     0.1021   20.48   <2e-16 ***
E4            2.7904     0.1010   27.63   <2e-16 ***
G2:G18        3.6545     0.3071   11.90   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.608 on 1448 degrees of freedom
Multiple R-squared:  0.5104,    Adjusted R-squared:  0.509 
F-statistic: 377.3 on 4 and 1448 DF,  p-value: < 2.2e-16
```

Final Model

In [33]:
```
MEPlotNoT <- lm(Y^1.6 ~ ((E3) + E1:E4 + E3:E4 + G2:G18), data = data)
data$Y1.6 = (data$Y)^1.6
data$predicted <- predict(MEPlotNoT)

plot(data$Y1.6, data$predicted, main = "Observed vs. Predicted", xlab = "Obser

abline(lm(data$predicted ~ data$Y1.6), col = "red")

abline(0, 1, col = "blue", lty = 2)

plot(fitted(MEPlot), resid(MEPlot), abline(0,0), main = 'Residual Plot Final M
temp <- summary(MEPlotNoT)
kable(temp$coefficients[ abs(temp$coefficients[,4]) <= 0.001, ], caption='Sig (
kable(temp$coefficients[ abs(temp$coefficients[,3]) >= 4, ], caption='Sig Coef

confint(MEPlotNoT, level = 0.90)
confint(MEPlotNoT, level = 0.95)
confint(MEPlotNoT, level = 0.99)

summary(MEPlotNoT)
```

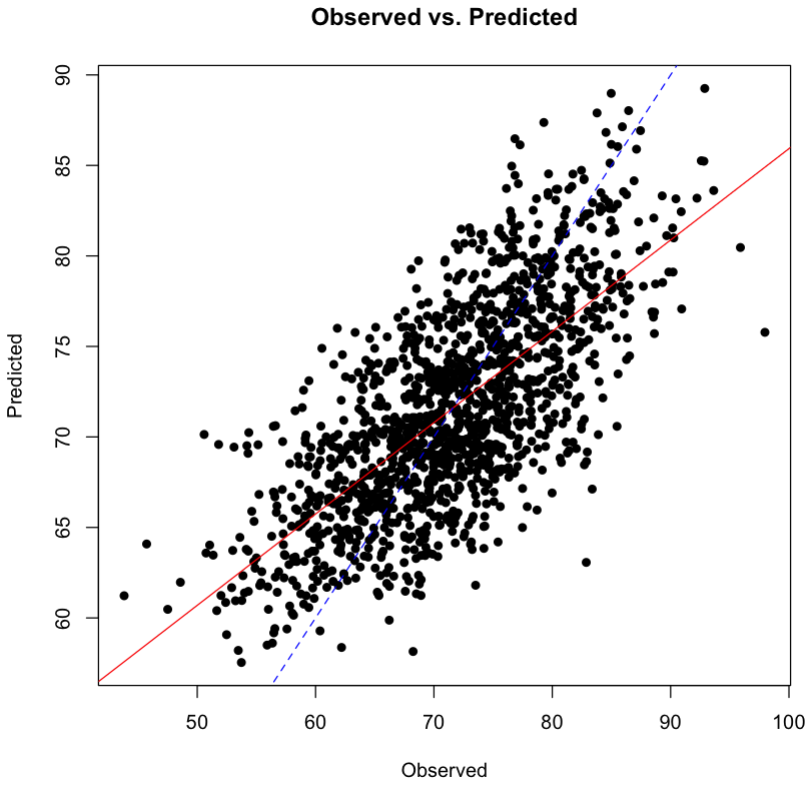


Table: Sig Coefficients Final Model

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	42.1152493	1.1063232	38.067763	0e+00
E3	0.8914593	0.1741044	5.120258	3e-07
E1:E4	0.2079469	0.0130877	15.888711	0e+00
E3:E4	0.1585580	0.0183806	8.626372	0e+00
G2:G18	3.6573878	0.3088809	11.840772	0e+00

Table: Sig Coefficients Final Model

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	42.1152493	1.1063232	38.067763	0e+00
E3	0.8914593	0.1741044	5.120258	3e-07
E1:E4	0.2079469	0.0130877	15.888711	0e+00
E3:E4	0.1585580	0.0183806	8.626372	0e+00
G2:G18	3.6573878	0.3088809	11.840772	0e+00

A matrix: 5 × 2 of type dbl

	5 %	95 %
(Intercept)	40.2943446	43.9361540
E3	0.6048998	1.1780188
E1:E4	0.1864058	0.2294881
E3:E4	0.1283052	0.1888108
G2:G18	3.1489988	4.1657769

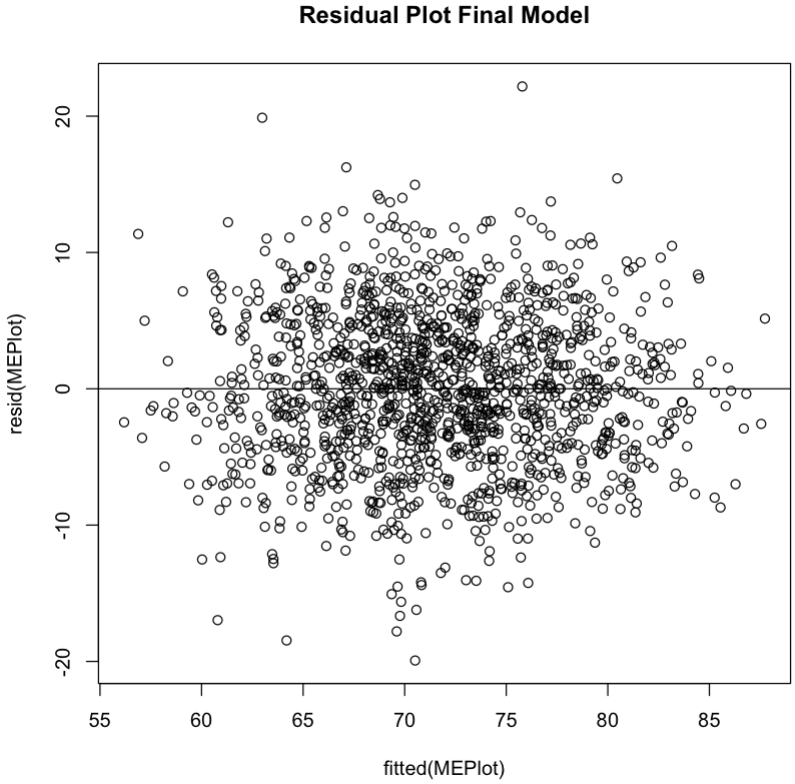
A matrix: 5 × 2 of type dbl

	2.5 %	97.5 %
(Intercept)	39.9450817	44.2854169
E3	0.5499355	1.2329831
E1:E4	0.1822740	0.2336198
E3:E4	0.1225025	0.1946135
G2:G18	3.0514860	4.2632897

A matrix: 5 × 2 of type dbl

	0.5 %	99.5 %
(Intercept)	39.2617885	44.9687101
E3	0.4424043	1.3405143
E1:E4	0.1741907	0.2417031
E3:E4	0.1111502	0.2059658
G2:G18	2.8607134	4.4540623

```
Call:
lm(formula = Y^1.6 ~ ((E3) + E1:E4 + E3:E4 + G2:G18), data = data)

Residuals:
     Min       1Q   Median       3Q      Max
-19.5538  -3.8397  -0.0525   3.9236  22.1887

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 42.11525    1.10632  38.068  < 2e-16 ***
E3           0.89146    0.17410   5.120 3.46e-07 ***
E1:E4        0.20795    0.01309  15.889  < 2e-16 ***
E3:E4        0.15856    0.01838   8.626  < 2e-16 ***
G2:G18       3.65739    0.30888  11.841  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.641 on 1448 degrees of freedom
Multiple R-squared:  0.5047,    Adjusted R-squared:  0.5033
F-statistic: 368.8 on 4 and 1448 DF,  p-value: < 2.2e-16
```


Residual Plot Final Model
resid(MEPlot)
fitted(MEPlot)
20
10
0
-10
-20
55
60
65
70
75
80
85

In [ ]: