

Warranty Payments Case Study

MAS II Fall 2019

Table of Contents	Page
A. Table of Contents	1
B. Setup and Exploratory Data Output	1
C. Results of Fitting Models	8
C.1 Fixed effects include payment lag. Random effects include store. Variance is assumed to be constant.	8
C.2 Fixed effects include payment lag and month sold. Random effects include store. Variance is assumed to be constant.	11
C.3 Fixed effects include payment lag and month sold. Random effects include store. Variance is assumed to be a function of payment lag using the exponential variance function option.	15
C.4 Fixed effects include payment lag and month sold. Random effects include store. Variance is assumed to be a function of payment lag using the power transform of payment time option.	19

B. Setup Study and Create Exploratory Data Output

In this case study, seven different models are displayed to model the costs for warranty payments for one product. All of the models displayed are Linear Mixed Effects models and use the nlme R package.

The product is sold in 100 different stores, but only three stores are included in the sample. Those three stores are under the variable Store and are denoted by: Store_1, Store_2 and Store_3.

The product was only sold for one year and the individual payments are recorded by month of sale, Sale_MOnth, and the month of payment after the sale. The warranty expires at the end of 12 months. The month of payment is denoted by number of months since the sale. For example, payments made on the warranty for products sold in January that were paid in January would show a payment lag, Pay_Lag, of 1. Payments for products sold in the month of February that were made in March would show a payment lag of 2.

The individual payments were transformed by using the natural logarithm transformation before the start of the analysis and are recorded in the column Log_Payment. The variable standard_log_payment is the result of normalizing the Log_Payment variable.

Set up Analysis by Loading Packages and Reading in Data Set

```
## -- Attaching packages -----
----- tidyverse 1.2.1 --

## v ggplot2 3.2.0      v purrr   0.3.2
## v tibble  2.1.3      v dplyr   0.8.1
## v tidyr   0.8.3      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.4.0

## -- Conflicts -----
----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

##
## Attaching package: 'nlme'

## The following object is masked from 'package:dplyr':
##
##   collapse

## New names:
## * Sigma_Store -> Sigma_Store...10
## * Sigma_Store -> Sigma_Store...12
```

Exploratory Data Information

Summary information on the by claim data set as well as a series of graphs are displayed in this section to inform interpretation of the analysis that follows.

```
summary(Repair)

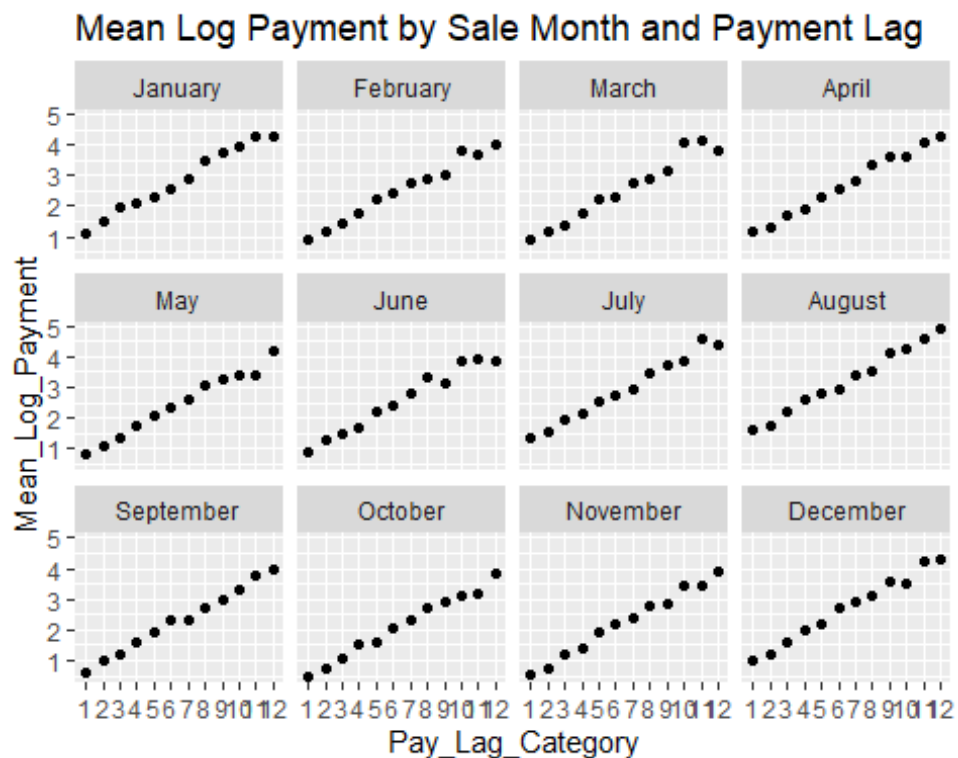
##  sale_month_lookup      Sale_Month      Store      Pay_Lag
##  Min.   : 1.0           December: 3601  Length:43200  Min.   : 1.0
##  1st Qu.: 4.0           February: 3600  Class :character  1st Qu.: 4.0
##  Median : 7.0           March    : 3600  Mode  :character  Median : 7.0
##  Mean   : 6.5           April   : 3600                      Mean   : 6.5
##  3rd Qu.:10.0          May     : 3600                      3rd Qu.:10.0
##  Max.   :12.0          June    : 3600                      Max.   :12.0
##                    (Other) :21599

##  Log_Payment      Payment      Pay_Lag_Category
##  Min.   :-7.6053   Min.   : 0      12   : 3601
##  1st Qu.: 0.9118   1st Qu.: 2      2    : 3600
##  Median : 2.2680   Median : 10     3    : 3600
##  Mean   : 2.5933   Mean   : 2199   4    : 3600
##  3rd Qu.: 4.0151   3rd Qu.: 55    5    : 3600
##  Max.   :16.8512   Max.   :20815028 6    : 3600
##                    (Other):21599
```

```
## standard_log_payment
## Min.   :-4.1539
## 1st Qu.: -0.6849
## Median :-0.1325
## Mean   : 0.0000
## 3rd Qu.: 0.5791
## Max.   : 5.8072
##

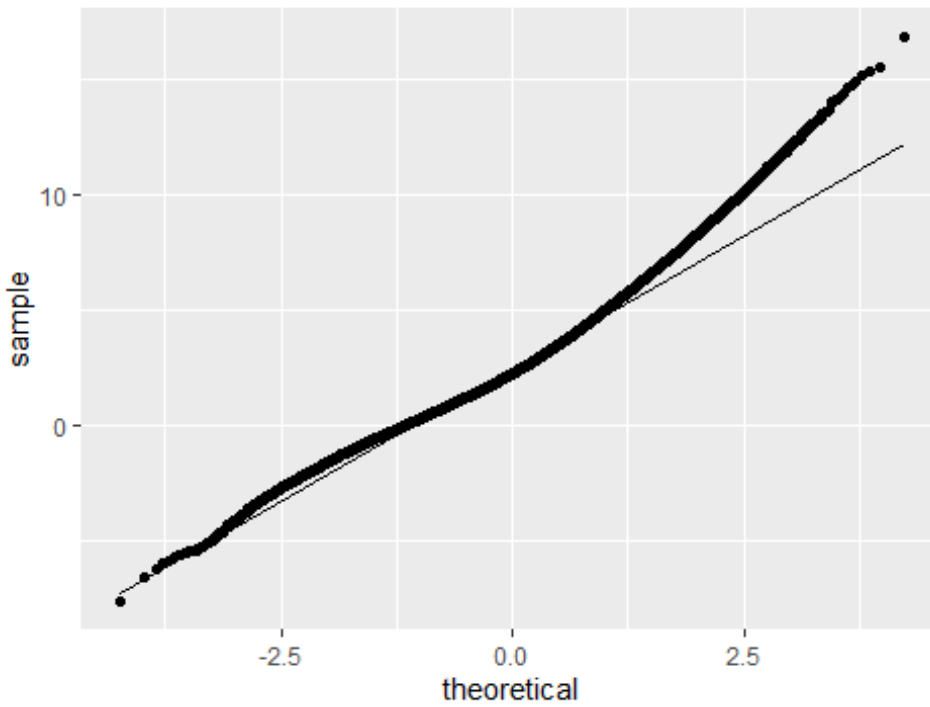
Repair_Group <- Repair %>% group_by(Sale_Month, Pay_Lag_Category) %>%
  mutate(Mean_Log_Payment = mean(Log_Payment), Std_Dev_Log_Payment = sd(Log_Payment))

ggplot (data= Repair_Group, aes(x=Pay_Lag_Category, y= Mean_Log_Payment)) +geom_point() +
  facet_wrap(vars(Sale_Month)) +labs(title ="Mean Log Payment by Sale Month and Payment Lag")
```



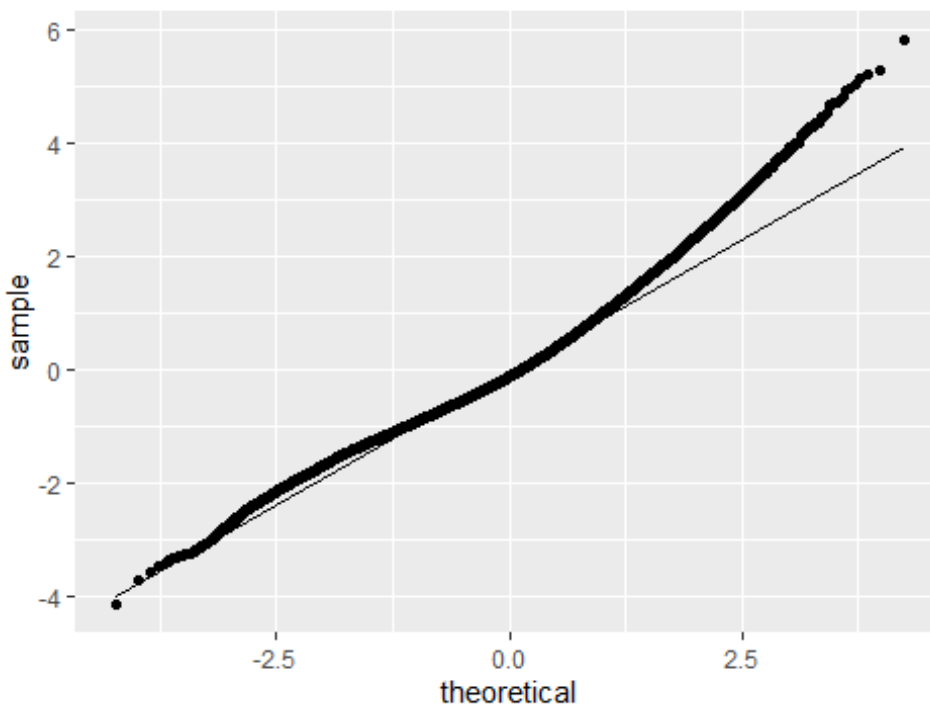
```
ggplot(data=Repair, aes(sample=Log_Payment)) + geom_qq() +geom_qq_line()+
  labs(title="QQ Plot for Log of Payments")
```

QQ Plot for Log of Payments

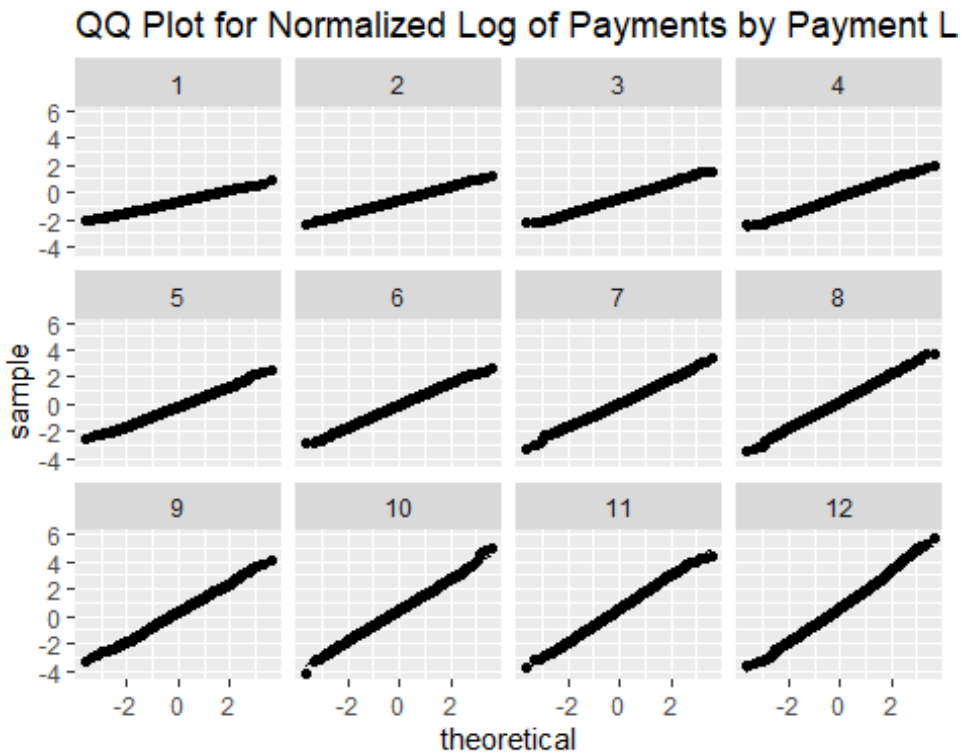


```
ggplot(data=Repair,aes(sample=standard_log_payment)) + geom_qq() +geom_qq_line()+  
  labs(title="QQ Plot for Normalized Log of Payments")
```

QQ Plot for Normalized Log of Payments

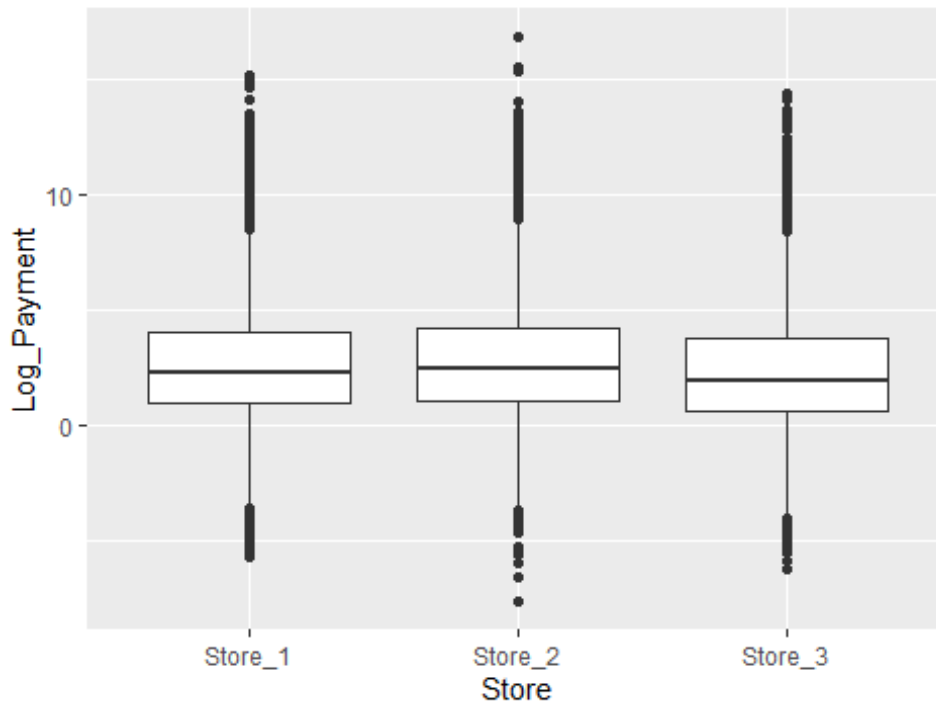


```
ggplot(data=Repair,aes(sample=standard_log_payment)) + geom_qq() +geom_qq_line()+
labs(title="QQ Plot for Normalized Log of Payments by Payment Lag") +
facet_wrap(vars(Pay_Lag_Category))
```



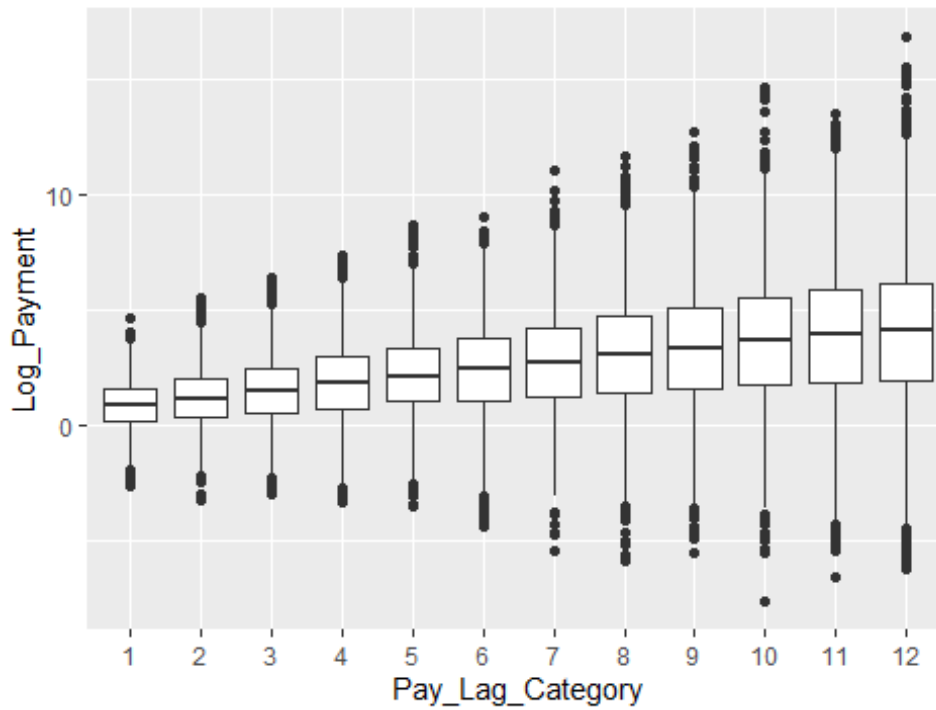
```
ggplot(data = Repair, aes(x=Store,y=Log_Payment))+ geom_boxplot()+
labs(title="Box Plot of Log Payments by Store")
```

Box Plot of Log Payments by Store

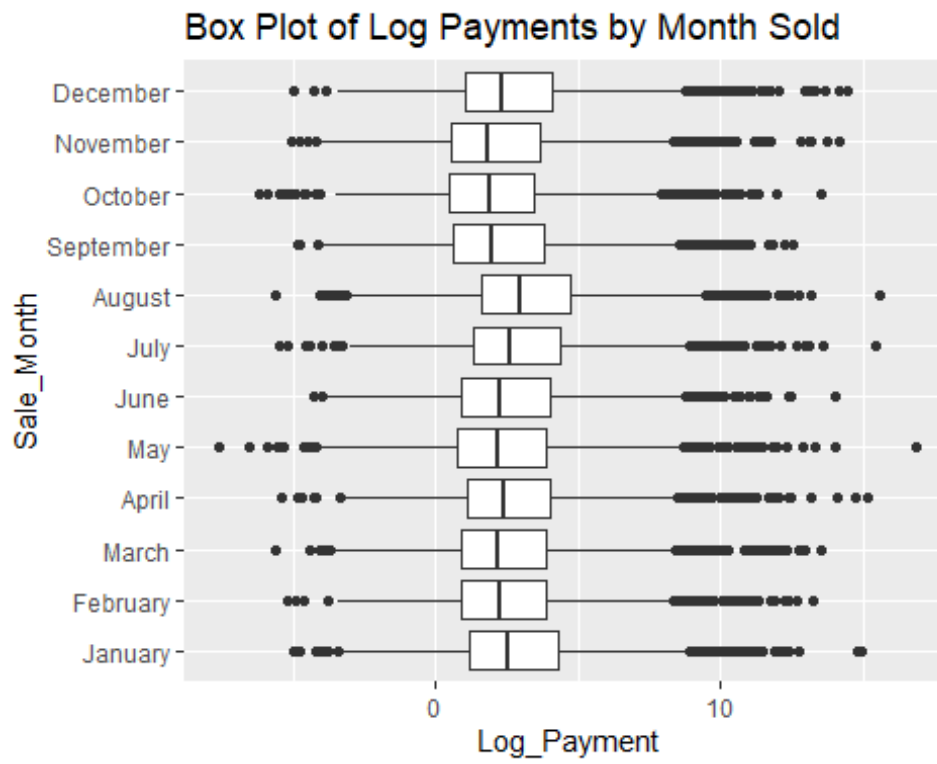


```
ggplot(data = Repair, aes(x=Pay_Lag_Category,y=Log_Payment))+ geom_boxplot()+  
labs(title="Box Plot of Log Payments by Payment Lag")
```

Box Plot of Log Payments by Payment Lag



```
ggplot(data = Repair, aes(x=Sale_Month,y=Log_Payment))+ geom_boxplot()+  
labs(title="Box Plot of Log Payments by Month Sold")+ coord_flip()
```



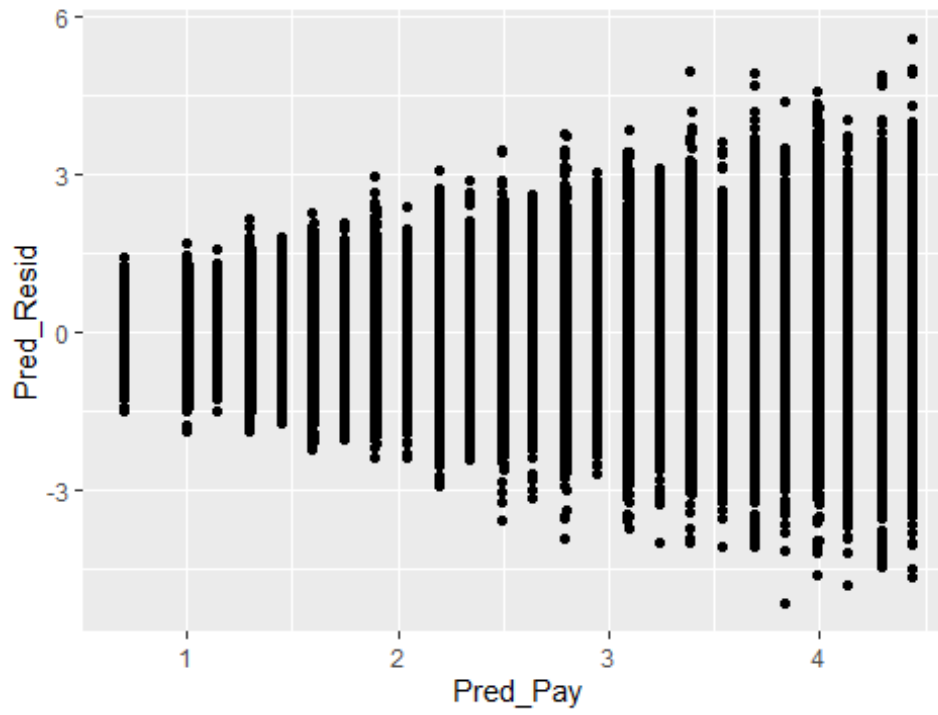
C. Results of Fitting Models

Create Model 1

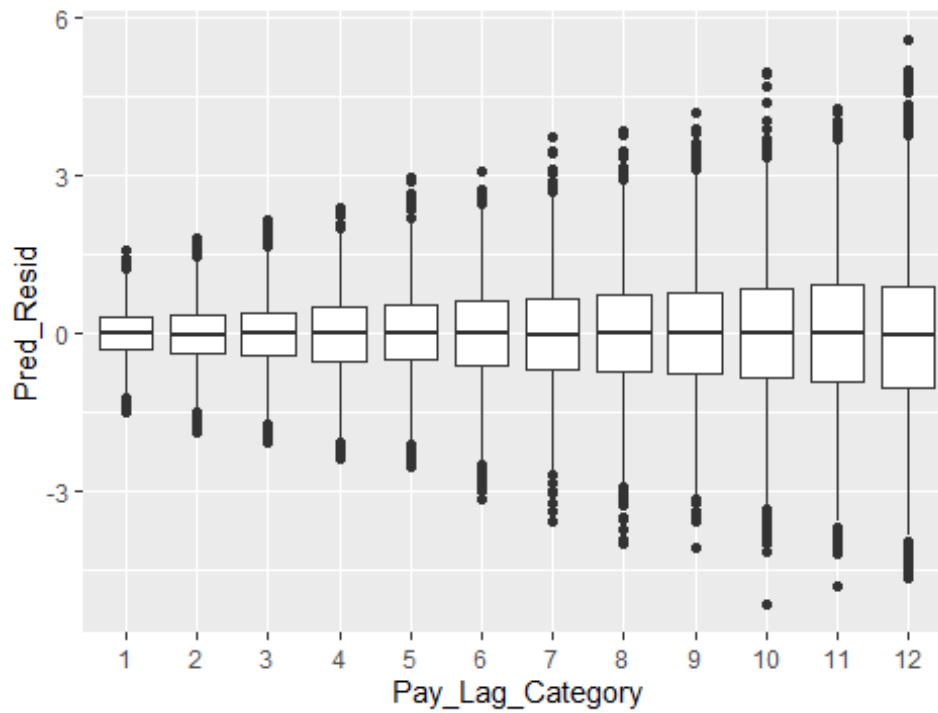
Fixed effects include payment lag. Random effects include store. Variance is assumed to be constant.

```
## Linear mixed-effects model fit by REML
## Data: Repair
##      AIC      BIC   logLik
## 191545.6 191580.3 -95768.8
##
## Random effects:
## Formula: ~1 | Store
##      (Intercept) Residual
## StdDev:  0.2302979 2.220299
##
## Fixed effects: Log_Payment ~ Pay_Lag
##              Value Std.Error   DF  t-value p-value
## (Intercept) 0.6518274 0.13489914 43196  4.83196    0
## Pay_Lag      0.2986713 0.00309451 43196 96.51643    0
## Correlation:
##      (Intr)
## Pay_Lag -0.149
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -5.152857995 -0.587180201 -0.002330143  0.581518009  5.593045490
##
## Number of Observations: 43200
## Number of Groups: 3
##
##              numDF denDF  F-value p-value
## (Intercept)      1 43196  377.956 <.0001
## Pay_Lag          1 43196 9315.422 <.0001
```

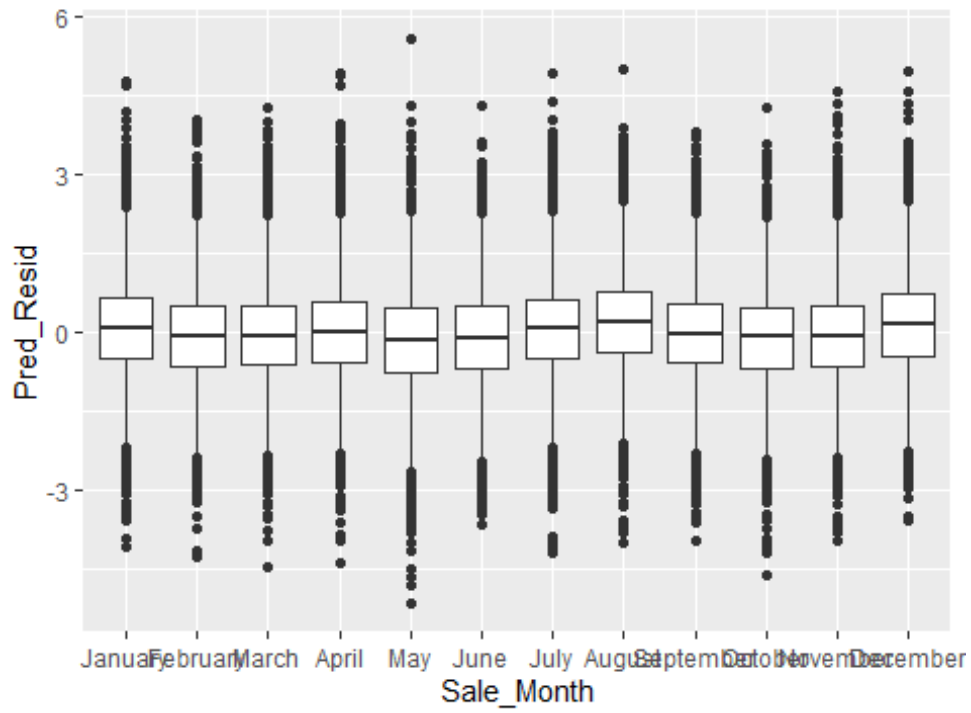

Model I Normalized Residuals vs. Predicted



Model I Box Plot Normalized Residuals vs. Payment Lag



Model I Box Plot Normalized Residuals vs. Sale Month



Create Model 2

Fixed effects include payment lag and month sold. Random effects include store. Variance is assumed to be constant.

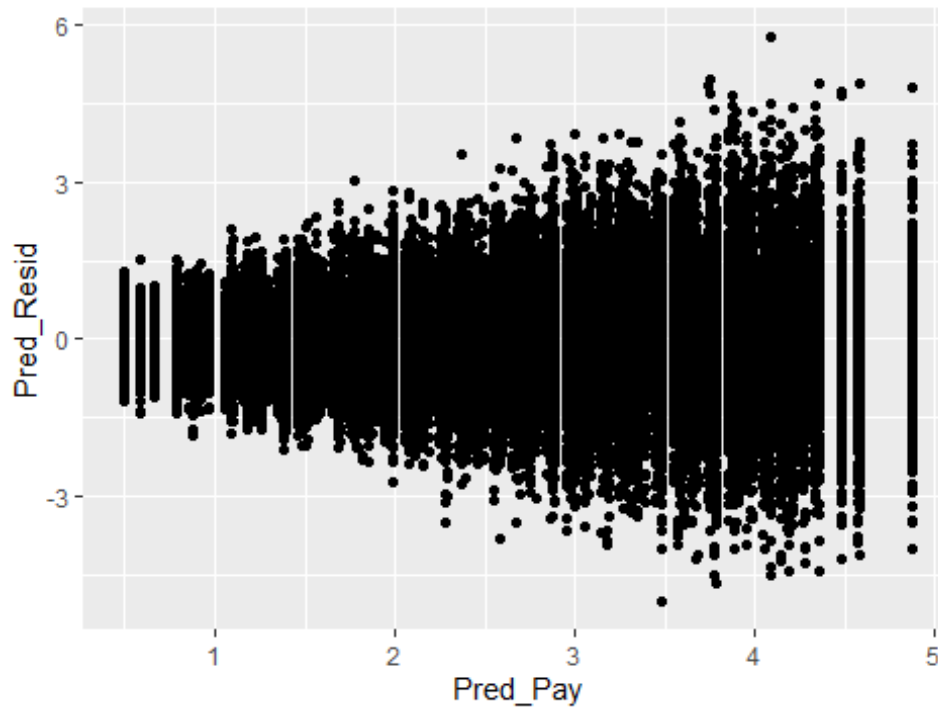
```
## Linear mixed-effects model fit by REML
## Data: Repair
##      AIC      BIC    logLik
## 191136.2 191266.3 -95553.12
##
## Random effects:
## Formula: ~1 | Store
##      (Intercept) Residual
## StdDev:  0.04907785 2.208503
##
## Fixed effects: Log_Payment ~ Pay_Lag + Sale_Month
##
##              Value Std.Error   DF  t-value p-value
## (Intercept)   0.8945594 0.06453183 43185 13.86230  0.0000
## Pay_Lag        0.2986655 0.00307807 43185 97.03005  0.0000
## Sale_MonthFebruary -0.3342177 0.05205854 43185 -6.42004  0.0000
## Sale_MonthMarch   -0.2885262 0.05205854 43185 -5.54234  0.0000
## Sale_MonthApril  -0.1265996 0.05205854 43185 -2.43187  0.0150
## Sale_MonthMay    -0.3960827 0.08676038 43185 -4.56525  0.0000
## Sale_MonthJune   -0.2724004 0.08676038 43185 -3.13969  0.0017
## Sale_MonthJuly    0.1019019 0.08676038 43185  1.17452  0.2402
## Sale_MonthAugust  0.3891668 0.08676038 43185  4.48554  0.0000
## Sale_MonthSeptember -0.5303462 0.08676038 43185 -6.11277  0.0000
## Sale_MonthOctober -0.7028963 0.08676038 43185 -8.10158  0.0000
## Sale_MonthNovember -0.6114114 0.08676038 43185 -7.04713  0.0000
## Sale_MonthDecember -0.1409673 0.08675821 43185 -1.62483  0.1042
## Correlation:
##              (Intr) Pay_Lg Sl_MnF Sl_MnthMr Sl_MnthAp Sl_MnthMy
## Pay_Lag      -0.310
## Sale_MonthFebruary -0.403  0.000
## Sale_MonthMarch   -0.403  0.000  0.500
## Sale_MonthApril  -0.403  0.000  0.500  0.500
## Sale_MonthMay    -0.672  0.000  0.300  0.300  0.300
## Sale_MonthJune   -0.672  0.000  0.300  0.300  0.300  0.820
## Sale_MonthJuly   -0.672  0.000  0.300  0.300  0.300  0.820
## Sale_MonthAugust -0.672  0.000  0.300  0.300  0.300  0.820
## Sale_MonthSeptember -0.672  0.000  0.300  0.300  0.300  0.500
## Sale_MonthOctober -0.672  0.000  0.300  0.300  0.300  0.500
## Sale_MonthNovember -0.672  0.000  0.300  0.300  0.300  0.500
## Sale_MonthDecember -0.672  0.000  0.300  0.300  0.300  0.500
```

```

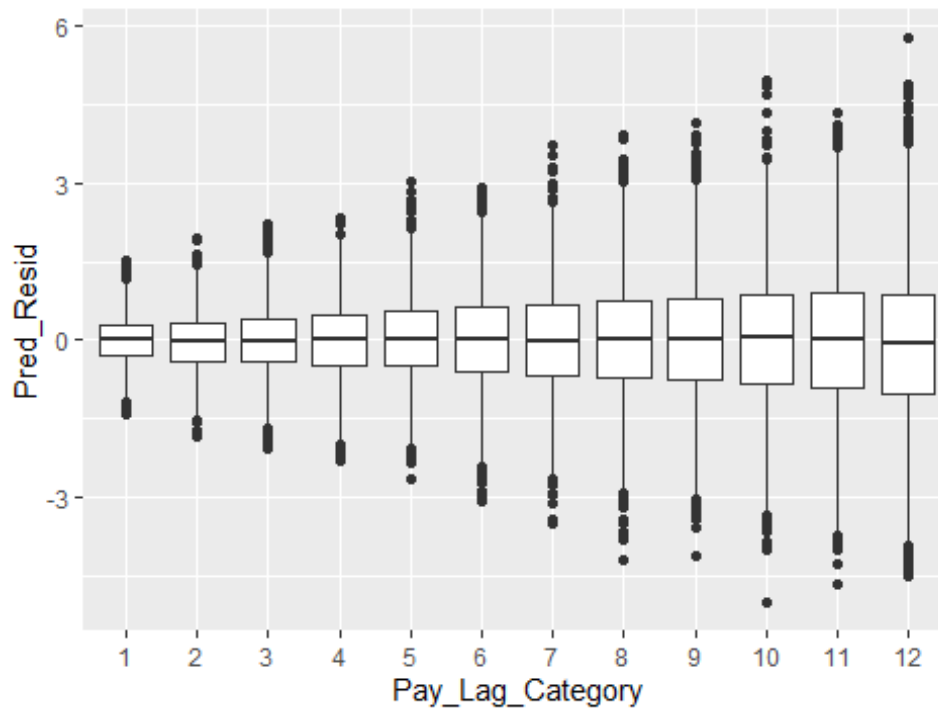
##           Sl_MnthJn Sl_MnthJl Sl_MnthAg Sl_MnS Sl_MnO Sl_MnN
## Pay_Lag
## Sale_MonthFebruary
## Sale_MonthMarch
## Sale_MonthApril
## Sale_MonthMay
## Sale_MonthJune
## Sale_MonthJuly      0.820
## Sale_MonthAugust    0.820      0.820
## Sale_MonthSeptember 0.500      0.500      0.500
## Sale_MonthOctober   0.500      0.500      0.500      0.820
## Sale_MonthNovember  0.500      0.500      0.500      0.820  0.820
## Sale_MonthDecember  0.500      0.500      0.500      0.820  0.820  0.820
##
## Standardized Within-Group Residuals:
##           Min           Q1           Med           Q3           Max
## -5.021685091 -0.584161394 -0.002552077  0.579781948  5.781617861
##
## Number of Observations: 43200
## Number of Groups: 3
##
##           numDF denDF  F-value p-value
## (Intercept)      1 43185 7343.471 <.0001
## Pay_Lag           1 43185 9415.180 <.0001
## Sale_Month       11 43185  46.410 <.0001

```

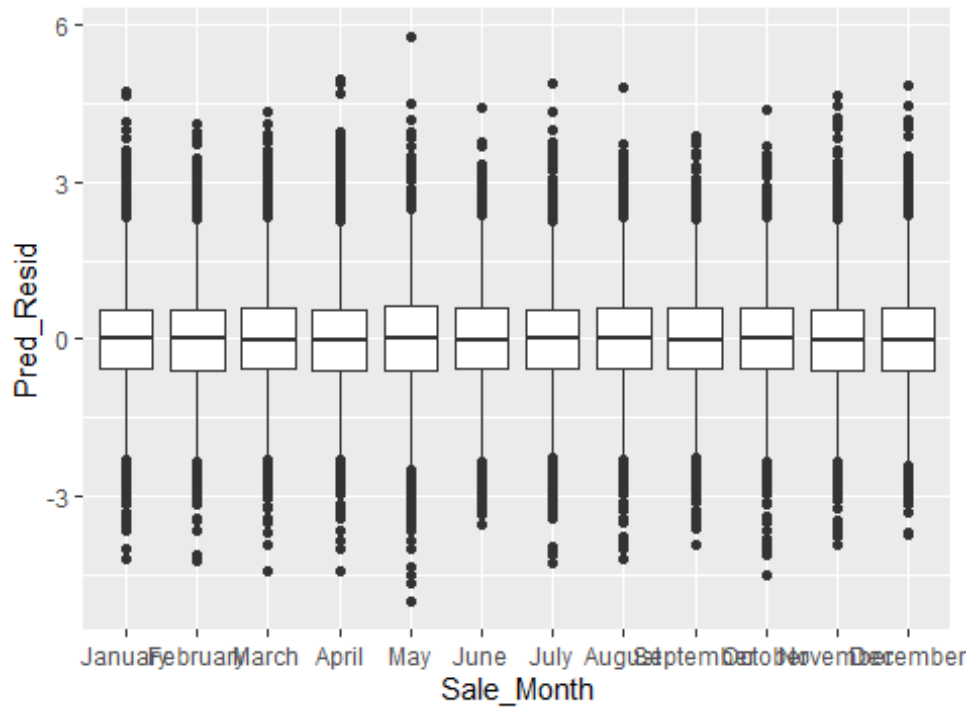
Model II Normalized Residuals vs. Predicted



Model II Box Plot Normalized Residuals vs. Payment La



Model II Box Plot Normalized Residuals vs. Sale Month



Create Model 3

Fixed effects include payment lag and month sold. Random effects include store. Variance is assumed to be a function of payment lag using the exponential variance function option.

```
wgt_function_1 <- varExp(form = ~Pay_Lag/2)

Model_III <- lme (Log_Payment ~ Pay_Lag + Sale_Month, data =Repair,
                 random = ~1|Store,
                 weights=wgt_function_1)
summary(Model_III)

## Linear mixed-effects model fit by REML
## Data: Repair
##      AIC      BIC    logLik
## 181627.8 181766.6 -90797.89
##
## Random effects:
## Formula: ~1 | Store
##      (Intercept) Residual
## StdDev:  0.02249708 1.012199
##
## Variance function:
## Structure: Exponential of variance covariate
## Formula: ~Pay_Lag/2
## Parameter estimates:
##      expon
## 0.2061607
## Fixed effects: Log_Payment ~ Pay_Lag + Sale_Month
##
##              Value Std.Error   DF  t-value p-value
## (Intercept)    0.8726579 0.03879148 43185  22.49612  0.0000
## Pay_Lag        0.3025872 0.00281533 43185 107.47826  0.0000
## Sale_MonthFebruary -0.3083311 0.04133114 43185  -7.46002  0.0000
## Sale_MonthMarch   -0.3059846 0.04133114 43185  -7.40325  0.0000
## Sale_MonthApril   -0.1204501 0.04133114 43185  -2.91427  0.0036
## Sale_MonthMay     -0.3872547 0.05215842 43185  -7.42459  0.0000
## Sale_MonthJune    -0.2721021 0.05215842 43185  -5.21684  0.0000
## Sale_MonthJuly     0.1135696 0.05215842 43185   2.17740  0.0295
## Sale_MonthAugust   0.3874351 0.05215842 43185   7.42805  0.0000
## Sale_MonthSeptember -0.5276125 0.05215842 43185 -10.11558  0.0000
## Sale_MonthOctober -0.6868433 0.05215842 43185 -13.16841  0.0000
## Sale_MonthNovember -0.6349874 0.05215842 43185 -12.17421  0.0000
## Sale_MonthDecember -0.1619186 0.05215784 43185  -3.10440  0.0019
```

```

## Correlation:
##              (Intr) Pay_Lg SI_MnF SI_MnthMr SI_MnthAp SI_MnthMy
## Pay_Lag        -0.310
## Sale_MonthFebruary -0.533  0.000
## Sale_MonthMarch   -0.533  0.000  0.500
## Sale_MonthApril   -0.533  0.000  0.500  0.500
## Sale_MonthMay     -0.672  0.000  0.396  0.396  0.396
## Sale_MonthJune    -0.672  0.000  0.396  0.396  0.396  0.686
## Sale_MonthJuly    -0.672  0.000  0.396  0.396  0.396  0.686
## Sale_MonthAugust  -0.672  0.000  0.396  0.396  0.396  0.686
## Sale_MonthSeptember -0.672  0.000  0.396  0.396  0.396  0.500
## Sale_MonthOctober -0.672  0.000  0.396  0.396  0.396  0.500
## Sale_MonthNovember -0.672  0.000  0.396  0.396  0.396  0.500
## Sale_MonthDecember -0.672  0.000  0.396  0.396  0.396  0.500
##              SI_MnthJn SI_MnthJl SI_MnthAg SI_MnS SI_MnO SI_MnN
## Pay_Lag
## Sale_MonthFebruary
## Sale_MonthMarch
## Sale_MonthApril
## Sale_MonthMay
## Sale_MonthJune
## Sale_MonthJuly      0.686
## Sale_MonthAugust    0.686  0.686
## Sale_MonthSeptember 0.500  0.500  0.500
## Sale_MonthOctober   0.500  0.500  0.500  0.686
## Sale_MonthNovember  0.500  0.500  0.500  0.686  0.686
## Sale_MonthDecember  0.500  0.500  0.500  0.686  0.686  0.686
##
## Standardized Within-Group Residuals:
##           Min           Q1           Med           Q3           Max
## -4.029197902 -0.669171884 -0.002832774  0.662012553  3.994009865
##
## Number of Observations: 43200
## Number of Groups: 3

anova(Model_III)

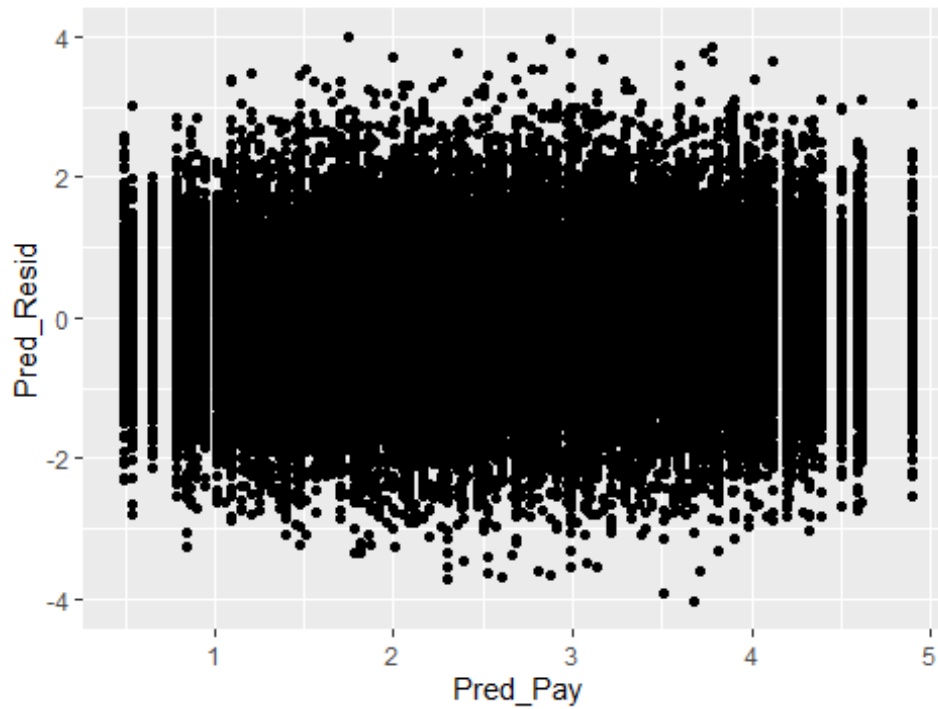
##           numDF denDF  F-value p-value
## (Intercept)     1 43185 15381.830 <.0001
## Pay_Lag          1 43185 11551.994 <.0001
## Sale_Month      11 43185   79.906 <.0001

Model_III_Results <- Repair %>% mutate(Pred_Pay =fitted.values(Model_III), Pr
ed_Resid =residuals(Model_III,type= "normalized"))

ggplot( data=Model_III_Results, aes(x=Pred_Pay, y=Pred_Resid)) +geom_point()
+
  labs( title ="Model III Normalized Residuals vs. Predicted")

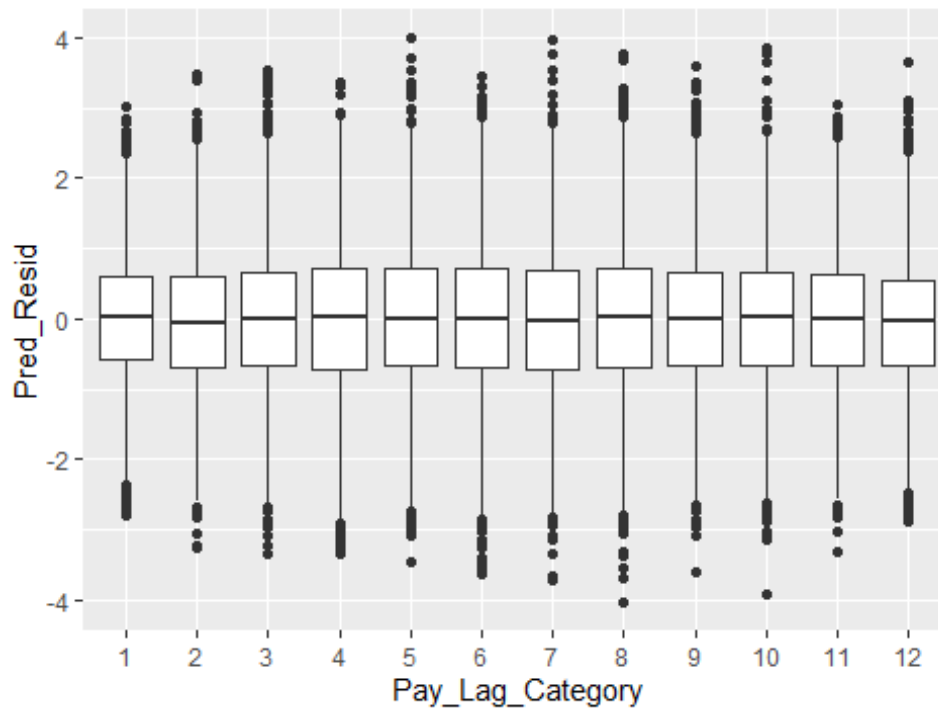
```


Model III Normalized Residuals vs. Predicted

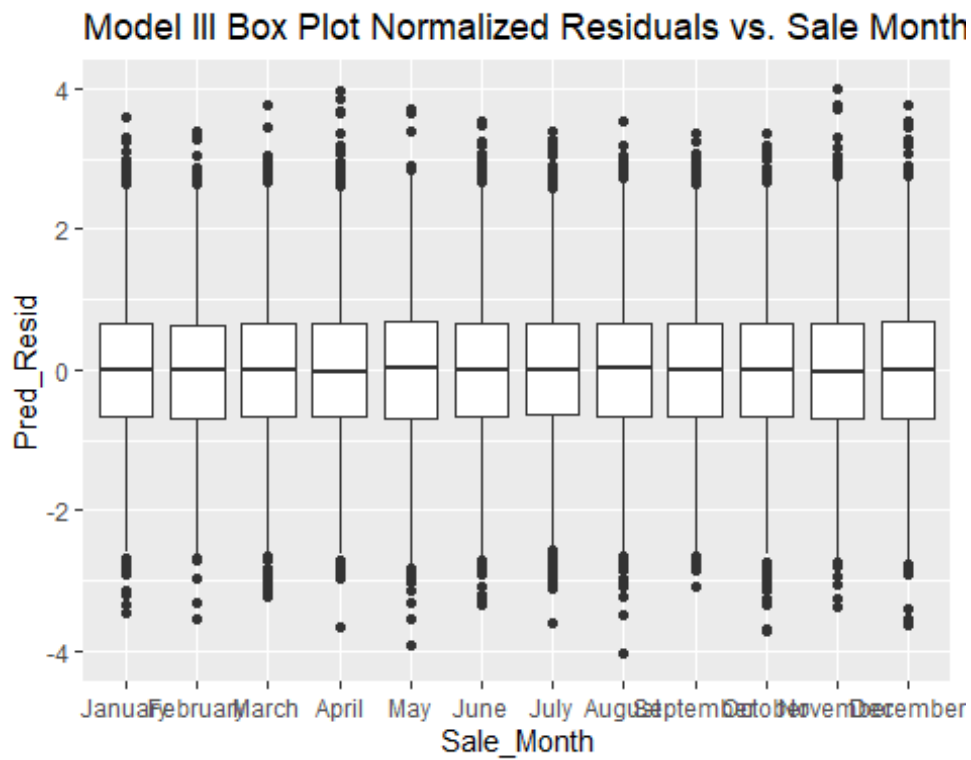


```
ggplot( data=Model_III_Results, aes(x=Pay_Lag_Category, y=Pred_Resid)) +geom_boxplot() +  
labs( title ="Model III Box Plot Normalized Residuals vs. Payment Lag")
```

Model III Box Plot Normalized Residuals vs. Payment Lag



```
ggplot( data=Model_III_Results, aes(x=Sale_Month, y=Pred_Resid)) +geom_boxplo  
t() +  
  labs( title ="Model III Box Plot Normalized Residuals vs. Sale Month")
```



Create Model 4

Fixed effects include payment lag and month sold. Random effects include store. Variance is assumed to be a function of payment lag using the power transform of payment time option.

```
wgt_function_3 <- varPower(form = ~Pay_Lag)

Model_IV <- lme (Log_Payment ~ Pay_Lag + Sale_Month, data =Repair,
                random = ~1|Store,
                weights=wgt_function_3)
summary(Model_IV)

## Linear mixed-effects model fit by REML
## Data: Repair
##      AIC      BIC    logLik
## 181693.1 181831.8 -90830.53
##
## Random effects:
## Formula: ~1 | Store
##      (Intercept) Residual
## StdDev:  0.01945993 0.8774883
##
## Variance function:
## Structure: Power of variance covariate
## Formula: ~Pay_Lag
## Parameter estimates:
##      power
## 0.4884623
## Fixed effects: Log_Payment ~ Pay_Lag + Sale_Month
##              Value Std.Error   DF  t-value p-value
## (Intercept)  0.8701807 0.03589243 43185  24.24413  0.0000
## Pay_Lag      0.3013128 0.00255453 43185 117.95252  0.0000
## Sale_MonthFebruary -0.2967367 0.04020954 43185  -7.37976  0.0000
## Sale_MonthMarch   -0.2965901 0.04020954 43185  -7.37611  0.0000
## Sale_MonthApril   -0.1041056 0.04020954 43185  -2.58908  0.0096
## Sale_MonthMay     -0.3808330 0.04872560 43185  -7.81587  0.0000
## Sale_MonthJune    -0.2711228 0.04872560 43185  -5.56428  0.0000
## Sale_MonthJuly     0.1276205 0.04872560 43185   2.61917  0.0088
## Sale_MonthAugust   0.3969035 0.04872560 43185   8.14569  0.0000
## Sale_MonthSeptember -0.5260494 0.04872560 43185 -10.79616  0.0000
## Sale_MonthOctober -0.6828486 0.04872560 43185 -14.01416  0.0000
## Sale_MonthNovember -0.6354767 0.04872560 43185 -13.04195  0.0000
## Sale_MonthDecember -0.1577172 0.04872484 43185  -3.23690  0.0012
```

```

## Correlation:
##              (Intr) Pay_Lg SI_MnF SI_MnthMr SI_MnthAp SI_MnthMy
## Pay_Lag        -0.280
## Sale_MonthFebruary -0.560  0.000
## Sale_MonthMarch   -0.560  0.000  0.500
## Sale_MonthApril   -0.560  0.000  0.500  0.500
## Sale_MonthMay     -0.679  0.000  0.413  0.413  0.413
## Sale_MonthJune    -0.679  0.000  0.413  0.413  0.413  0.660
## Sale_MonthJuly    -0.679  0.000  0.413  0.413  0.413  0.660
## Sale_MonthAugust  -0.679  0.000  0.413  0.413  0.413  0.660
## Sale_MonthSeptember -0.679  0.000  0.413  0.413  0.413  0.500
## Sale_MonthOctober -0.679  0.000  0.413  0.413  0.413  0.500
## Sale_MonthNovember -0.679  0.000  0.413  0.413  0.413  0.500
## Sale_MonthDecember -0.679  0.000  0.413  0.413  0.413  0.500
##              SI_MnthJn SI_MnthJl SI_MnthAg SI_MnS SI_MnO SI_MnN
## Pay_Lag
## Sale_MonthFebruary
## Sale_MonthMarch
## Sale_MonthApril
## Sale_MonthMay
## Sale_MonthJune
## Sale_MonthJuly      0.660
## Sale_MonthAugust    0.660  0.660
## Sale_MonthSeptember 0.500  0.500  0.500
## Sale_MonthOctober   0.500  0.500  0.500  0.660
## Sale_MonthNovember  0.500  0.500  0.500  0.660  0.660
## Sale_MonthDecember  0.500  0.500  0.500  0.660  0.660  0.660
##
## Standardized Within-Group Residuals:
##           Min           Q1           Med           Q3           Max
## -4.110764166 -0.668393442 -0.001169045  0.662217363  4.315151676
##
## Number of Observations: 43200
## Number of Groups: 3

anova(Model_IV)

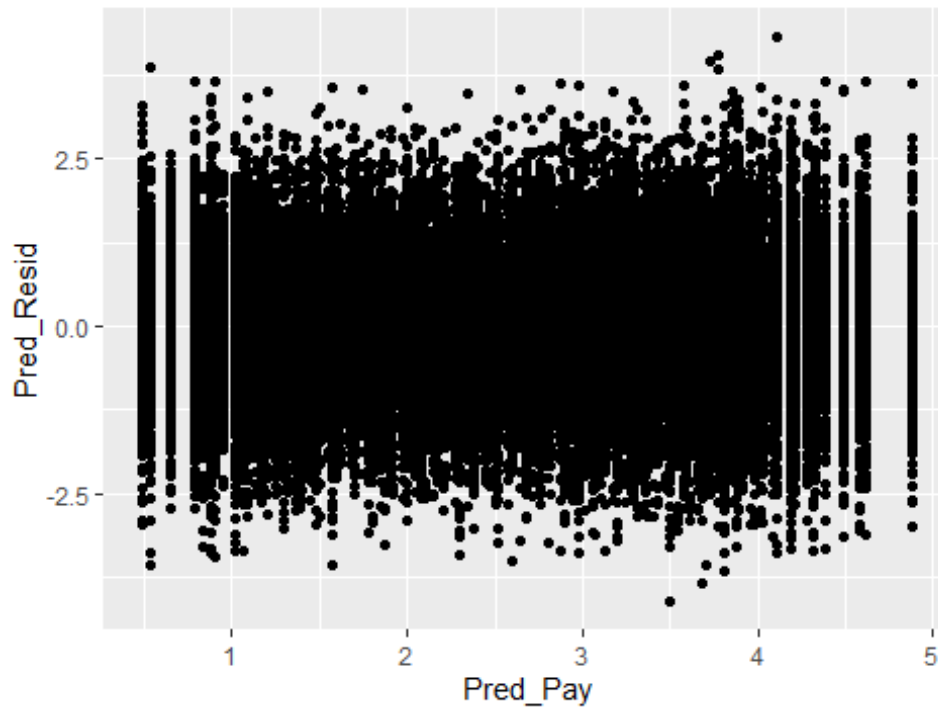
##           numDF denDF  F-value p-value
## (Intercept)     1 43185 17069.147 <.0001
## Pay_Lag          1 43185 13913.386 <.0001
## Sale_Month      11 43185   88.186 <.0001

Model_IV_Results <- Repair %>% mutate(Pred_Pay =fitted.values(Model_IV), Pred
_Resid =residuals(Model_IV,type= "normalized"))

ggplot( data=Model_IV_Results, aes(x=Pred_Pay, y=Pred_Resid)) +geom_point() +
labs( title ="Model IV Normalized Residuals vs. Predicted")

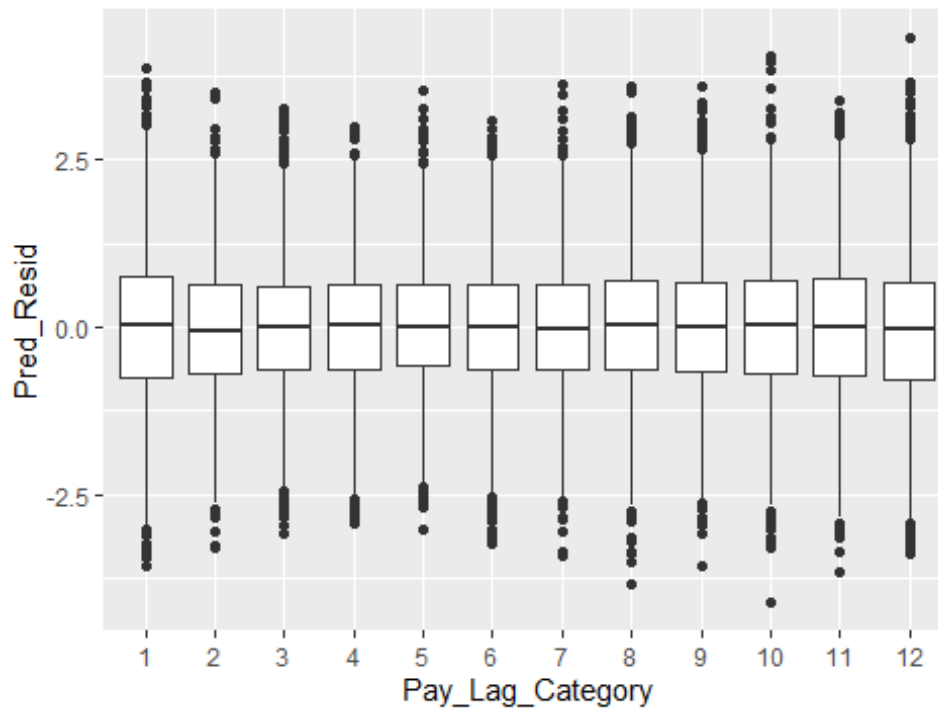
```

Model IV Normalized Residuals vs. Predicted



```
ggplot( data=Model_IV_Results, aes(x=Pay_Lag_Category, y=Pred_Resid)) +geom_boxplot() +  
labs( title ="Model IV Box Plot Normalized Residuals vs. Payment Lag")
```

Model IV Box Plot Normalized Residuals vs. Payment



```
ggplot( data=Model_IV_Results, aes(x=Sale_Month, y=Pred_Resid)) +geom_boxplot  
( ) +  
labs( title ="Model IV Box Plot Normalized Residuals vs. Sale Month")
```

