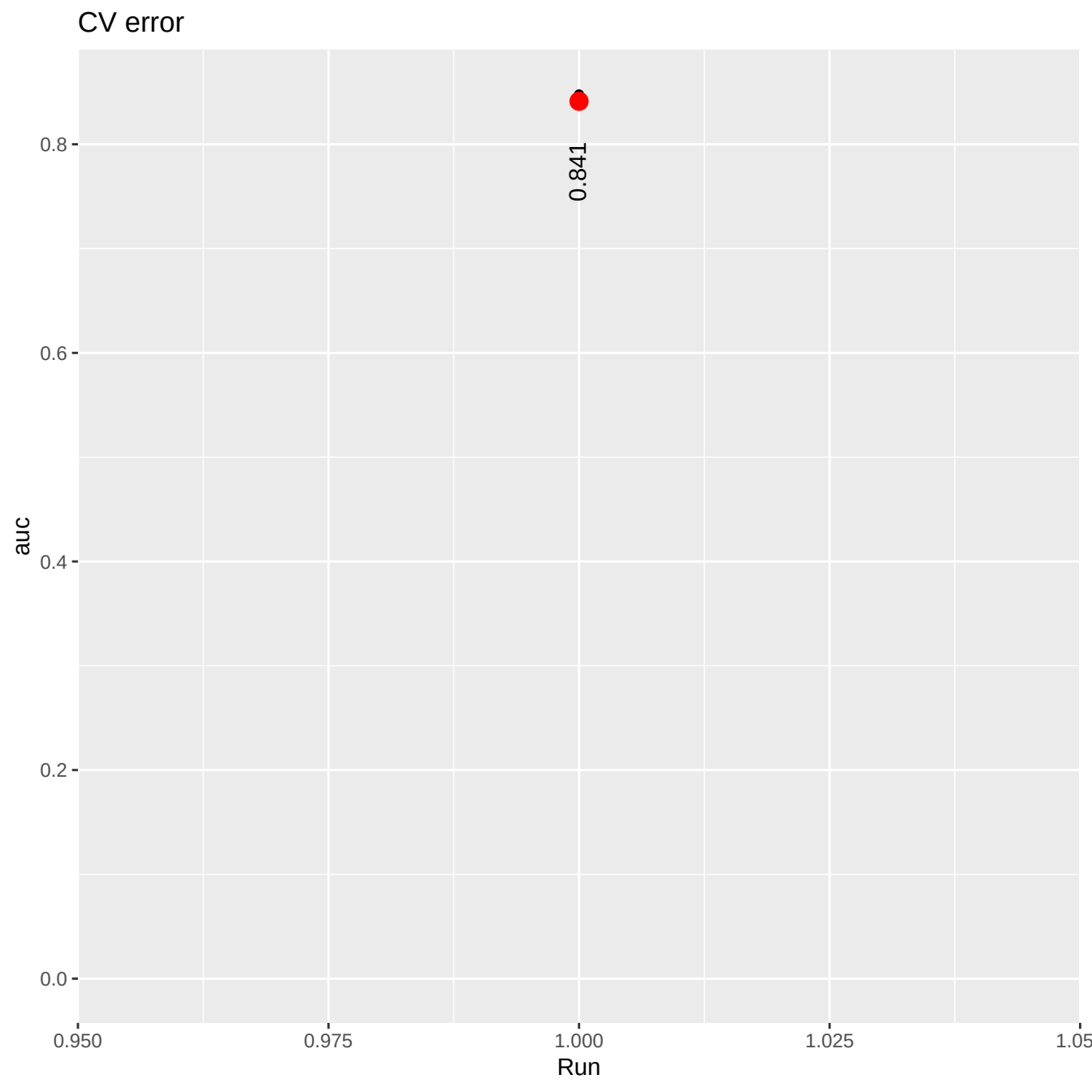
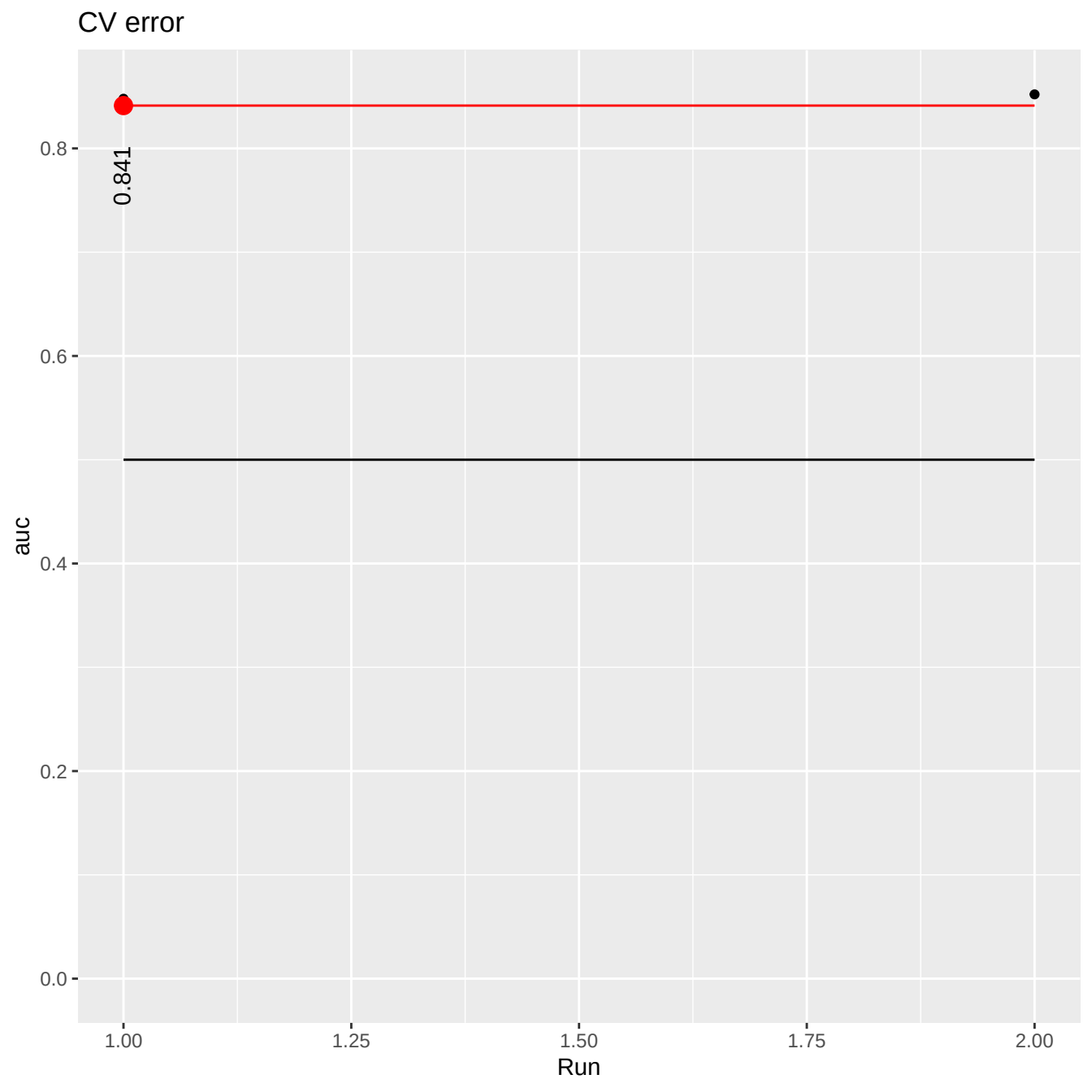
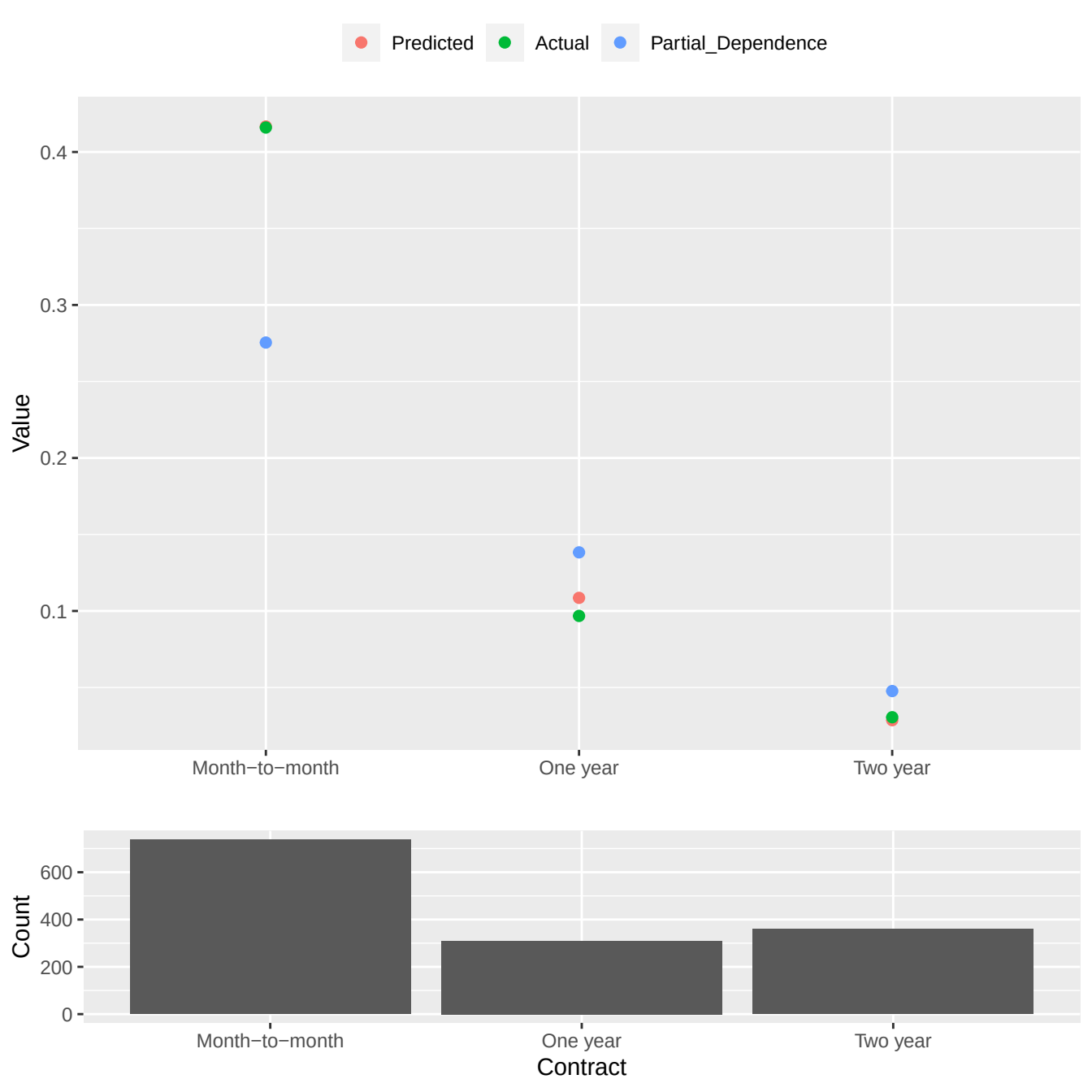
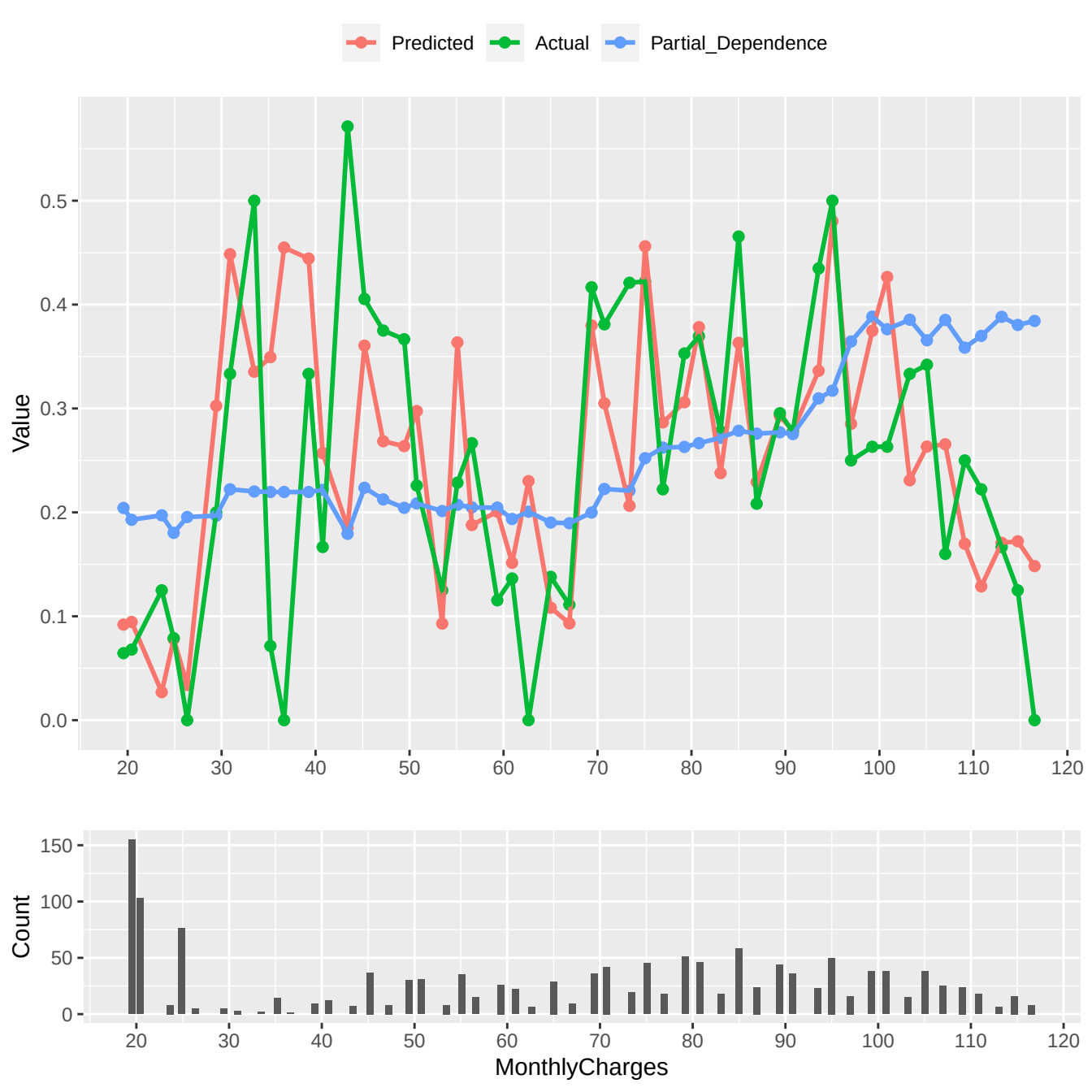

CV error
0.841
auc
0.0
0.2
0.4
0.6
0.8
0.950
0.975
1.000
1.025
1.05
Run

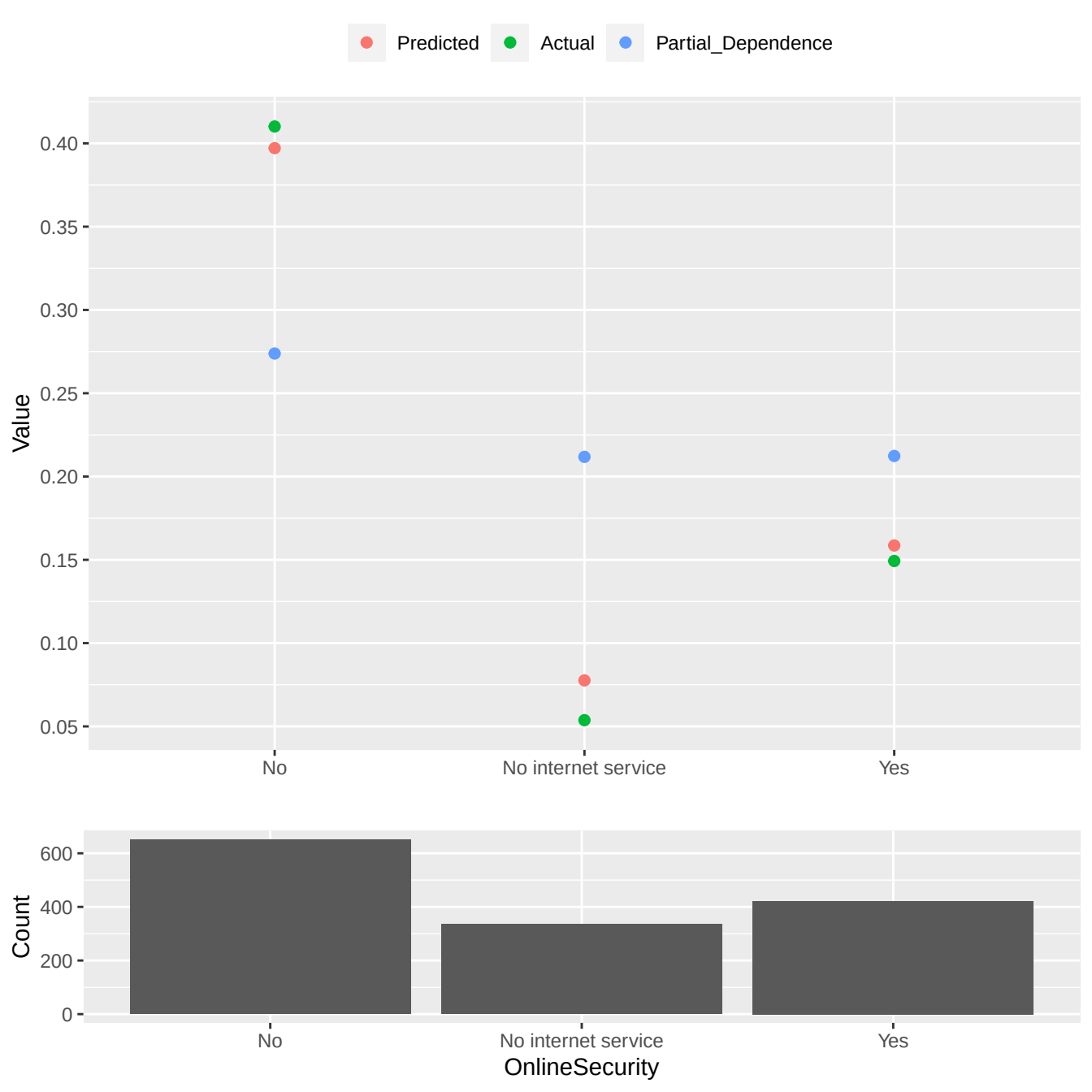

CV error
0.841
0.8
0.6
0.4
0.2
0.0
auc
1.00
1.25
1.50
1.75
2.00
Run

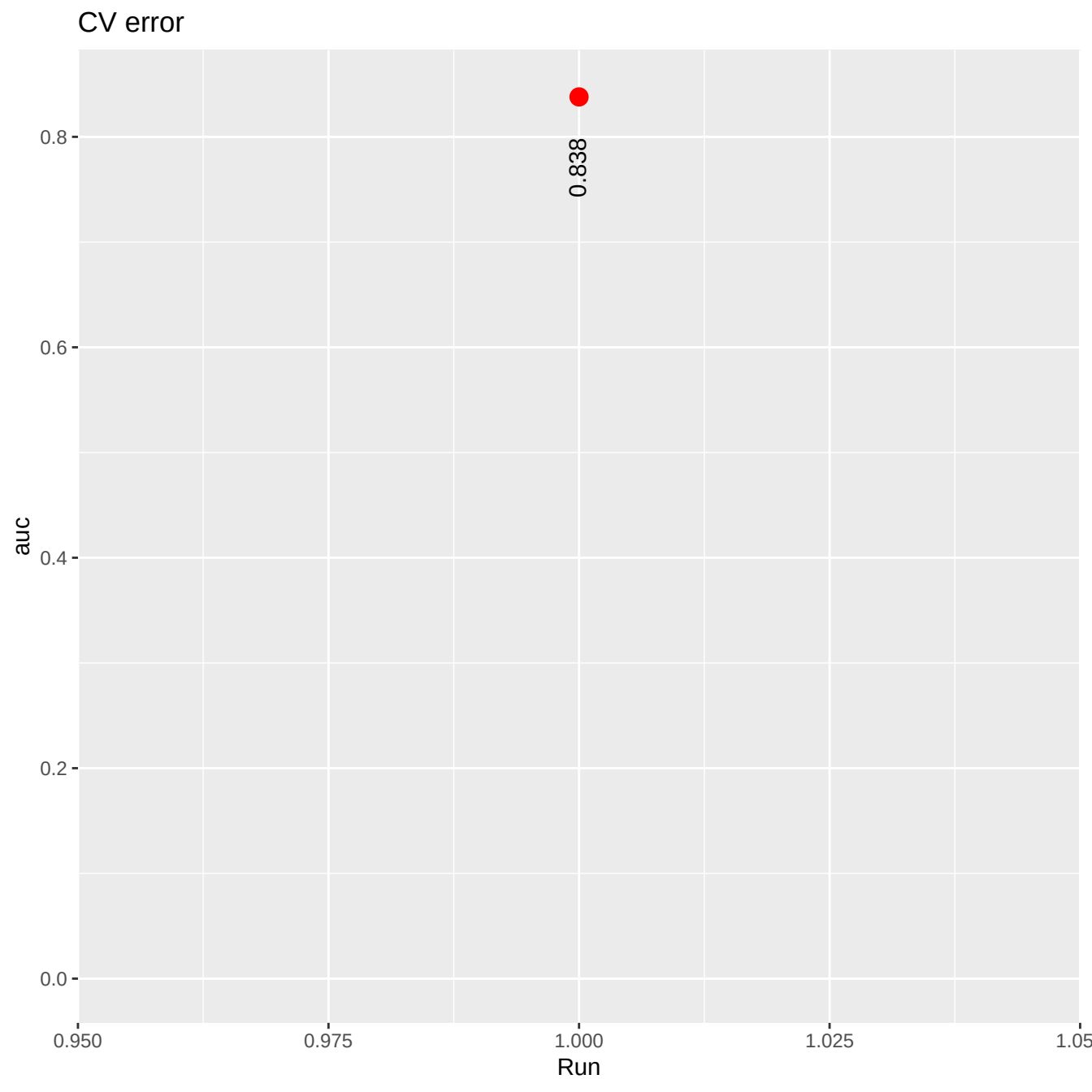

Predicted
Actual
Partial_Dependence
0.4
0.3
0.2
0.1
Value
Month−to−month
One year
Two year
600
400
200
0
Count
Month−to−month
One year
Two year
Contract

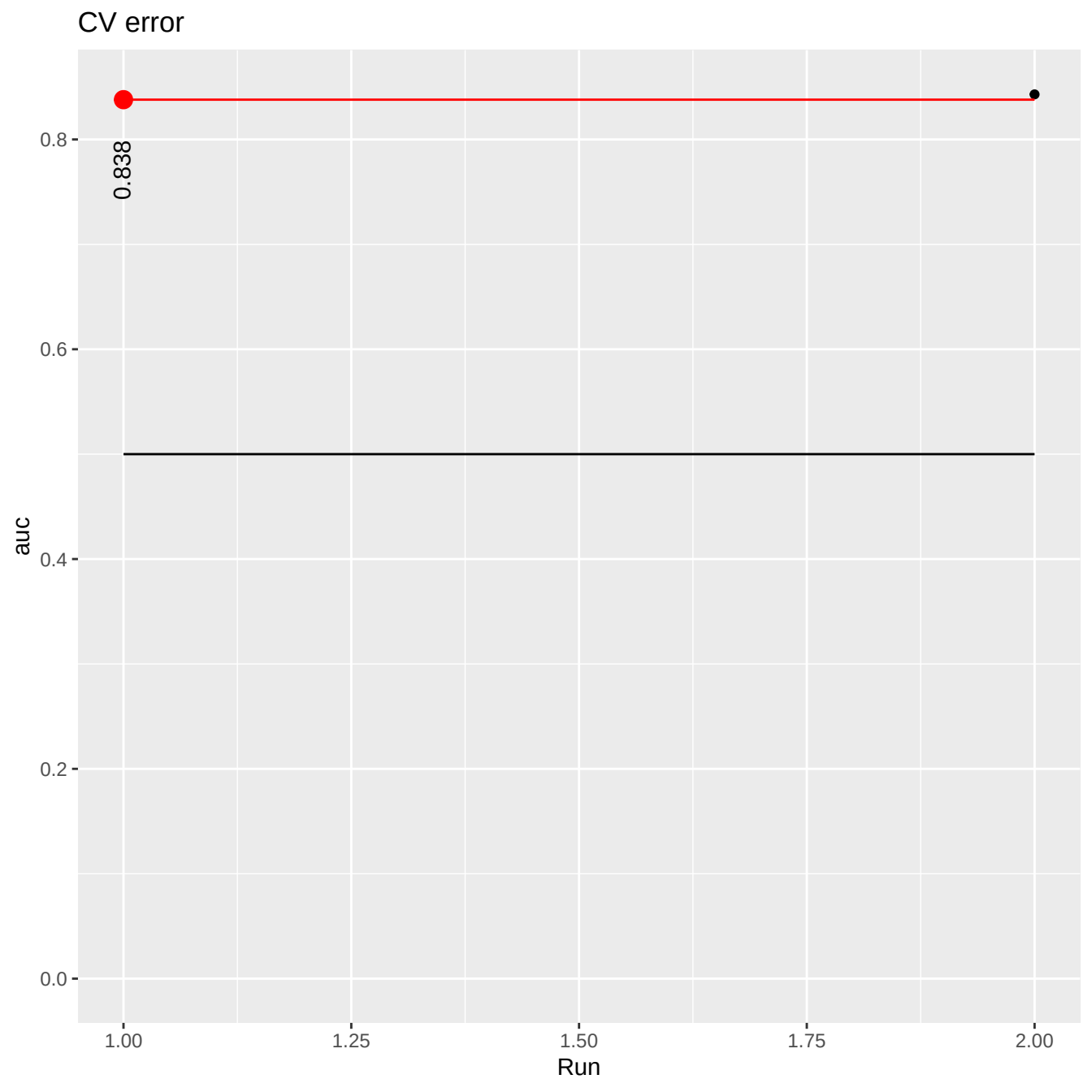

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
20
30
40
50
60
70
80
90
100
110
120
Count
0
50
100
150
MonthlyCharges

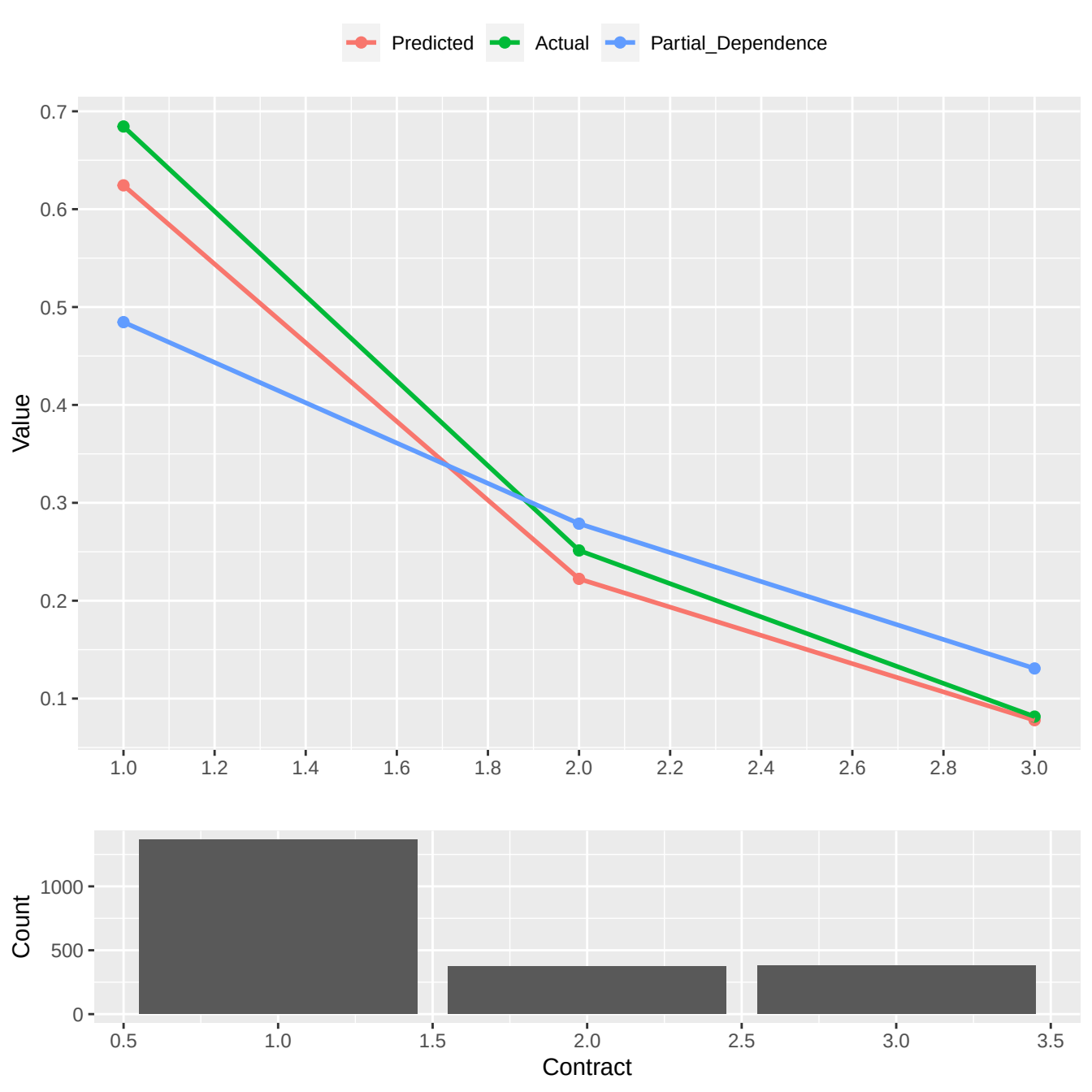

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
10
20
30
40
50
60
70
Count
0
50
100
150
tenure

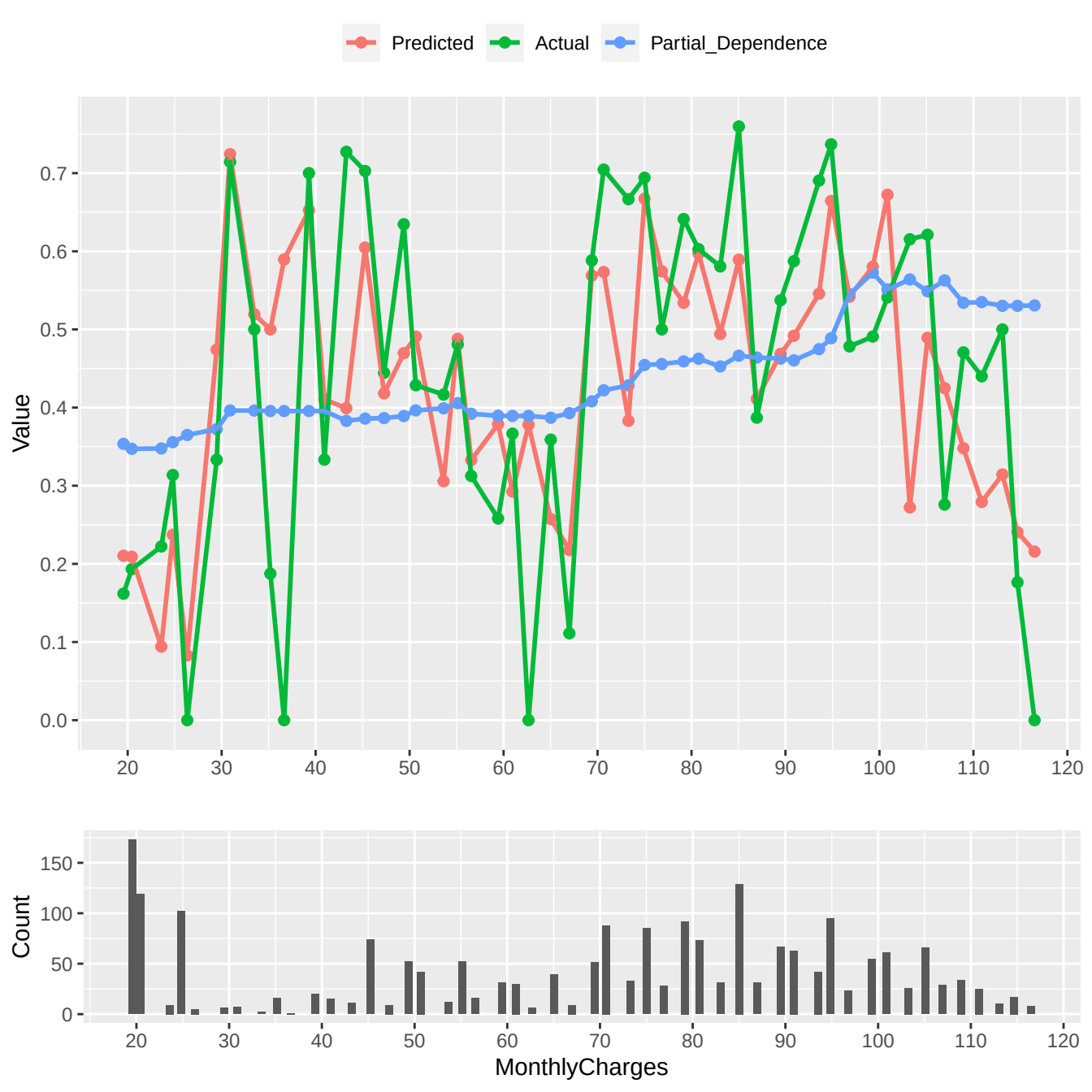

Predicted
Actual
Partial_Dependence
Value
0.5
0.4
0.3
0.2
0.1
0.0
0
1000
2000
3000
4000
5000
6000
7000
8000
Count
200
150
100
50
0
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
TotalCharges

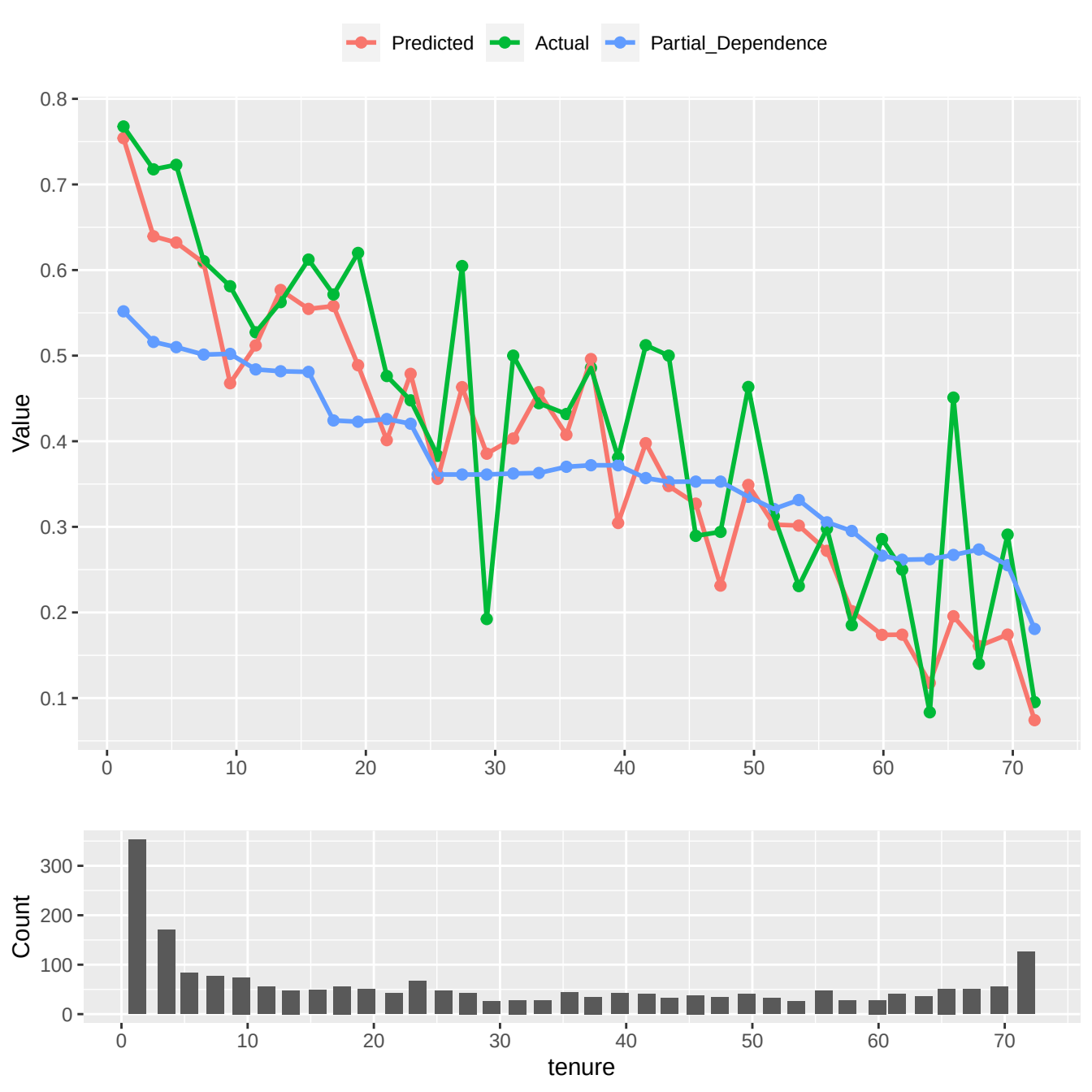

Predicted
Actual
Partial_Dependence
Value
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
No
No internet service
Yes
Count
600
400
200
0
No
No internet service
Yes
OnlineSecurity

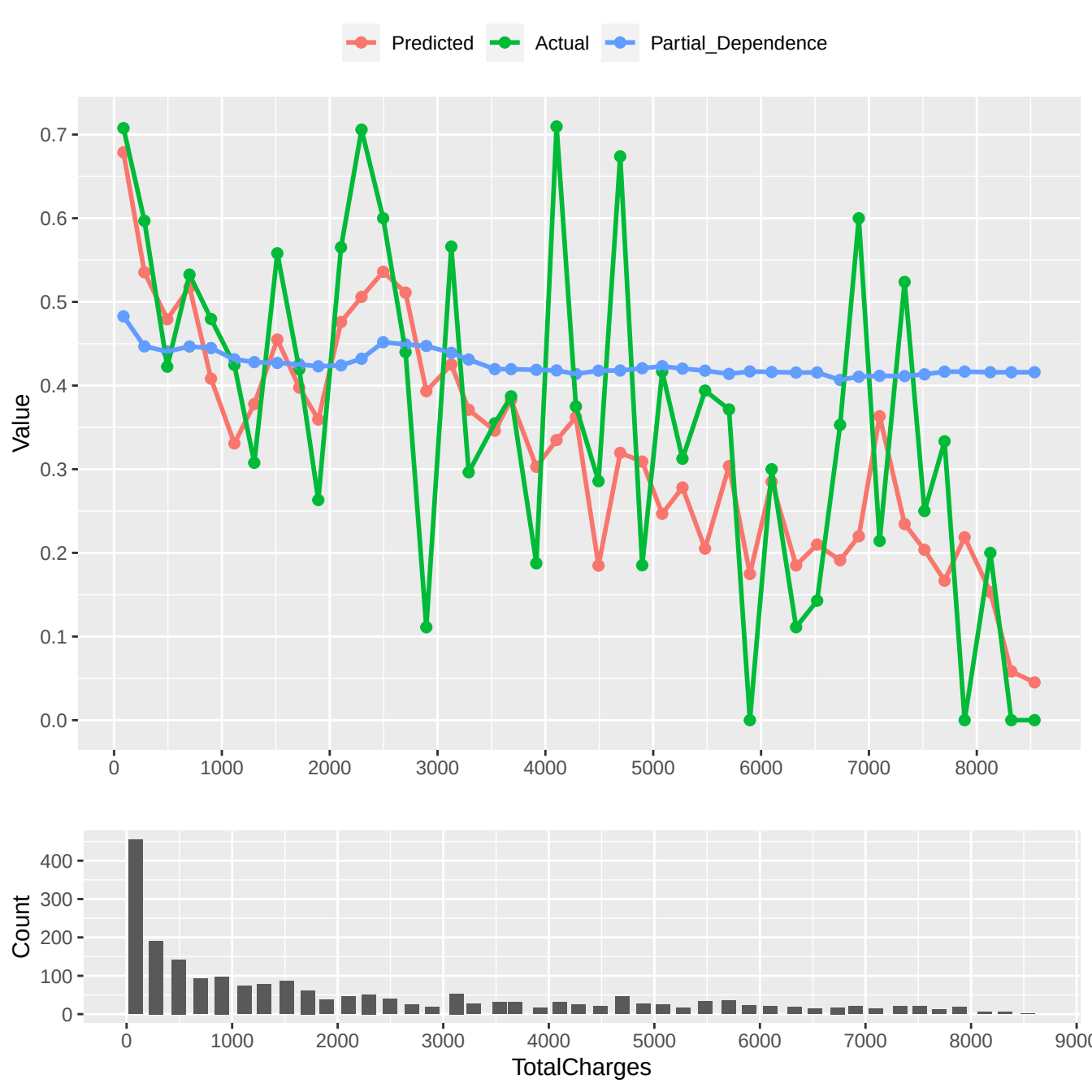

CV error
0.838
auc
0.0
0.2
0.4
0.6
0.8
0.950
0.975
1.000
1.025
1.05
Run

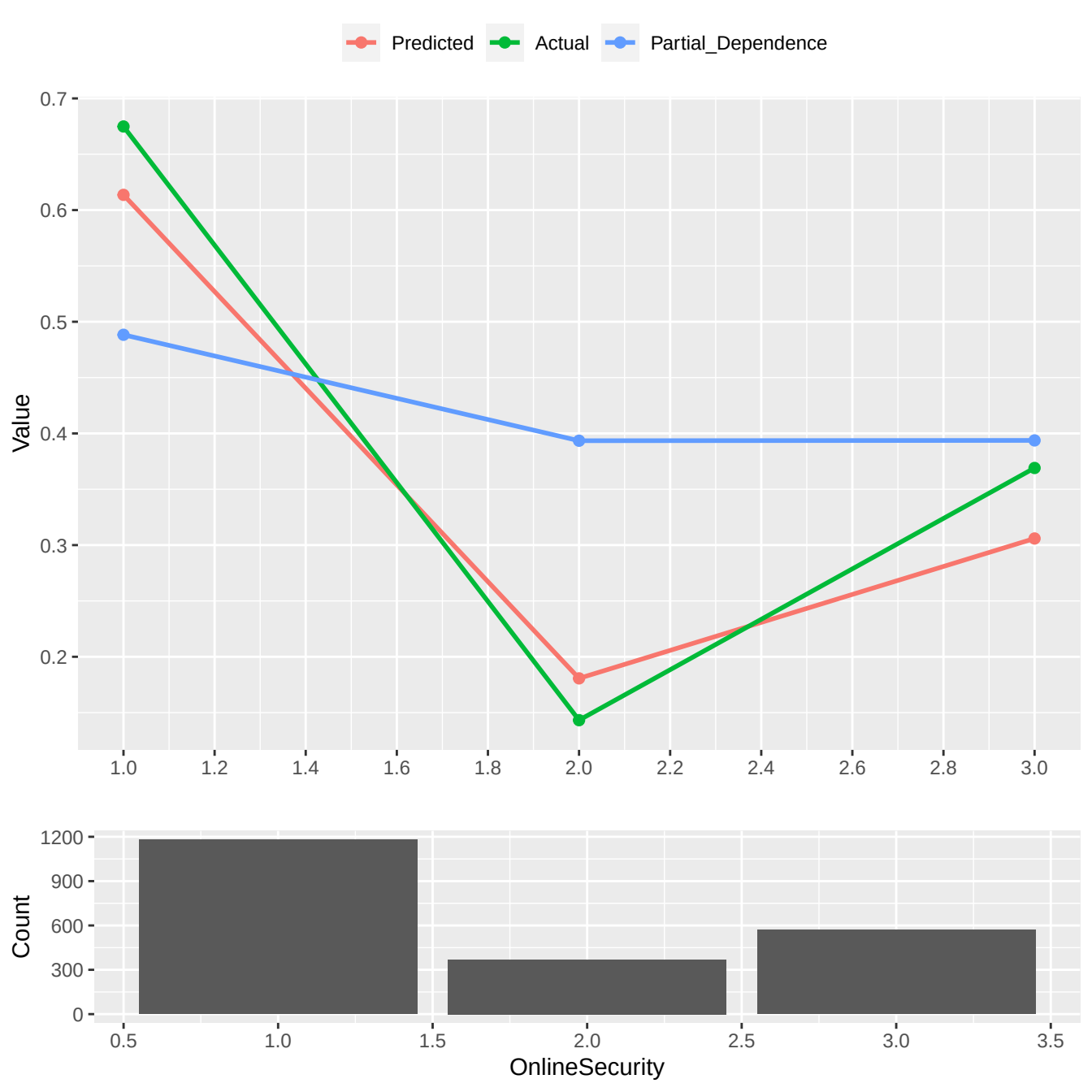

CV error
0.838
0.8
0.6
0.4
0.2
0.0
auc
1.00
1.25
1.50
1.75
2.00
Run

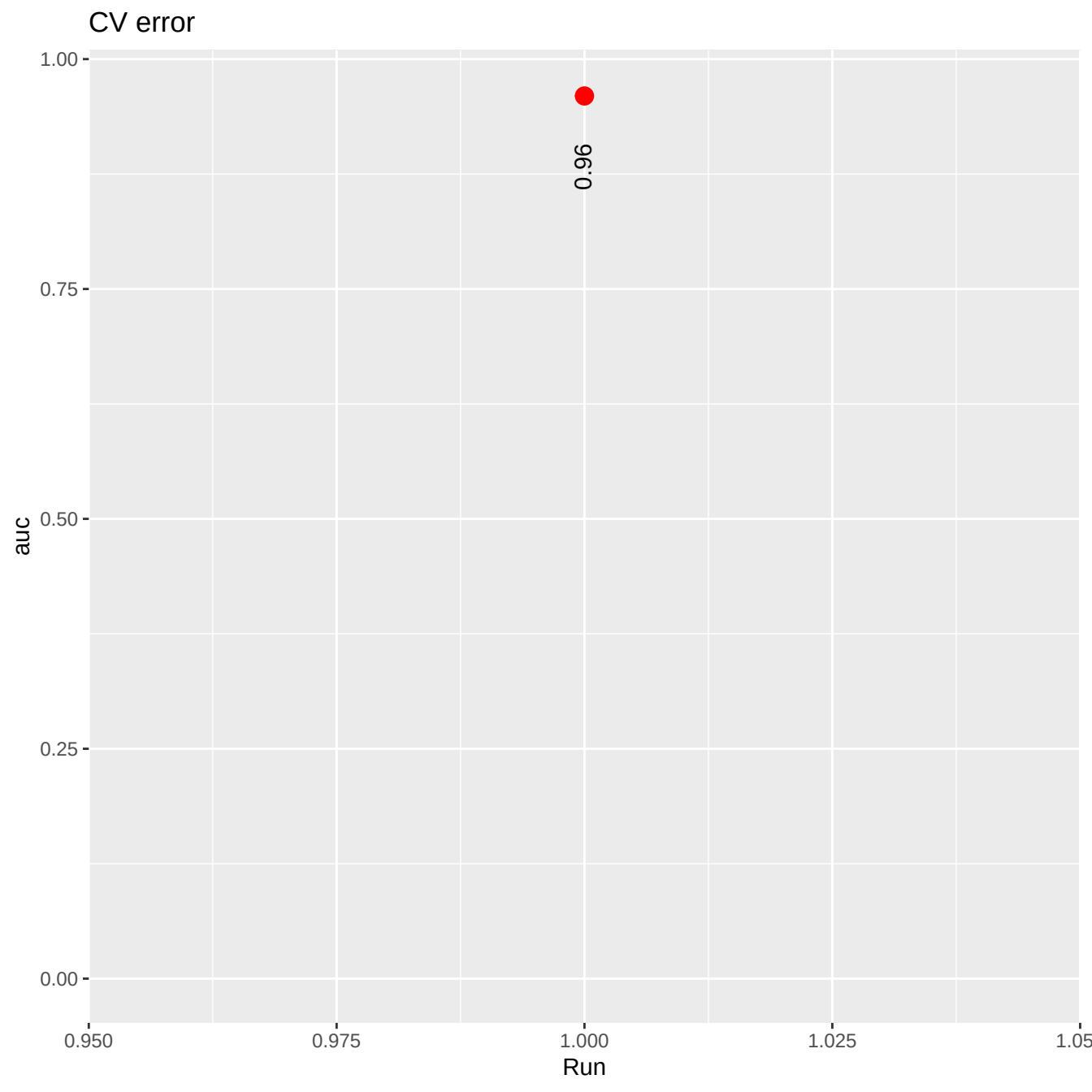

Predicted
Actual
Partial_Dependence
Value
0.7
0.6
0.5
0.4
0.3
0.2
0.1
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
1000
500
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Contract

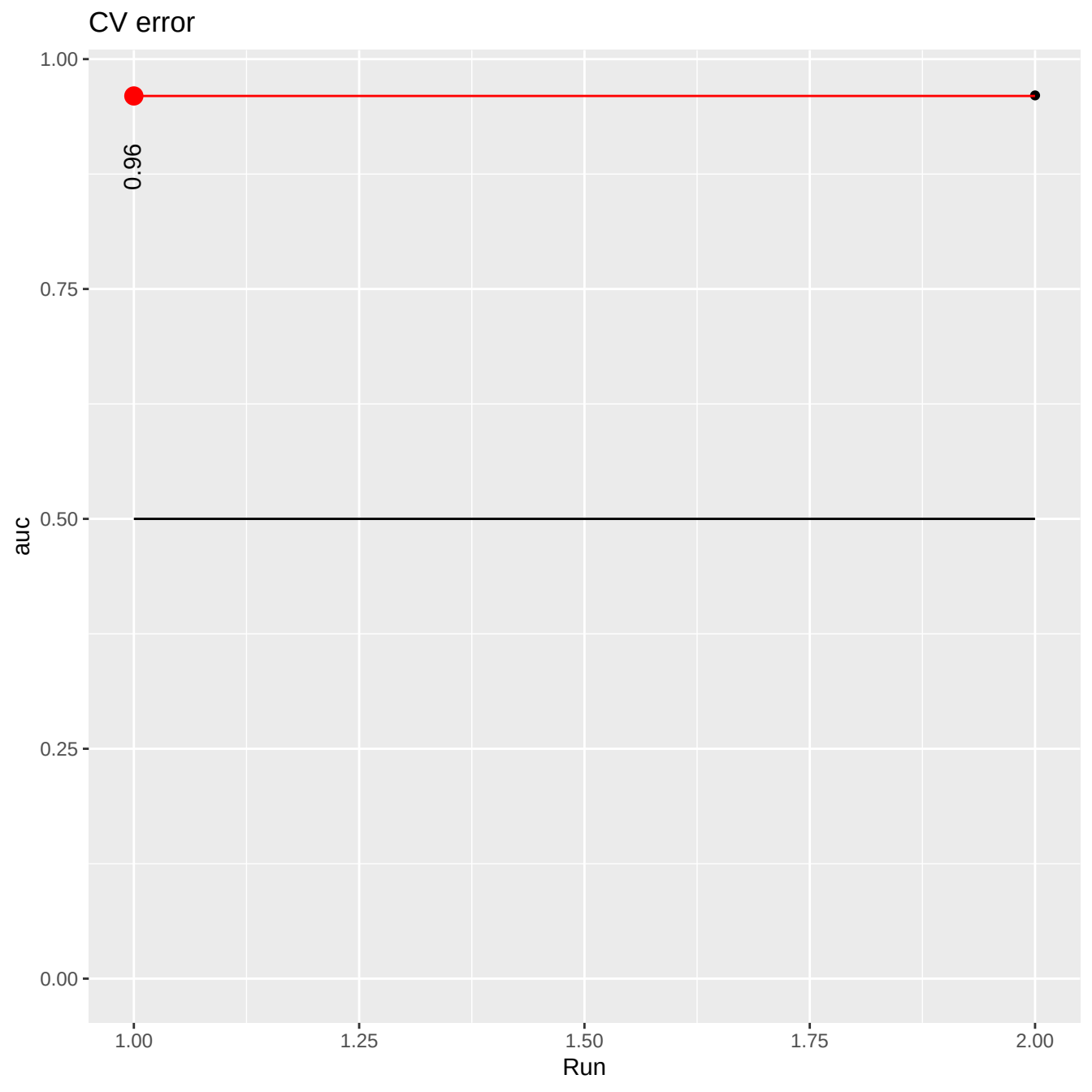

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
20
30
40
50
60
70
80
90
100
110
120
Count
0
50
100
150
MonthlyCharges

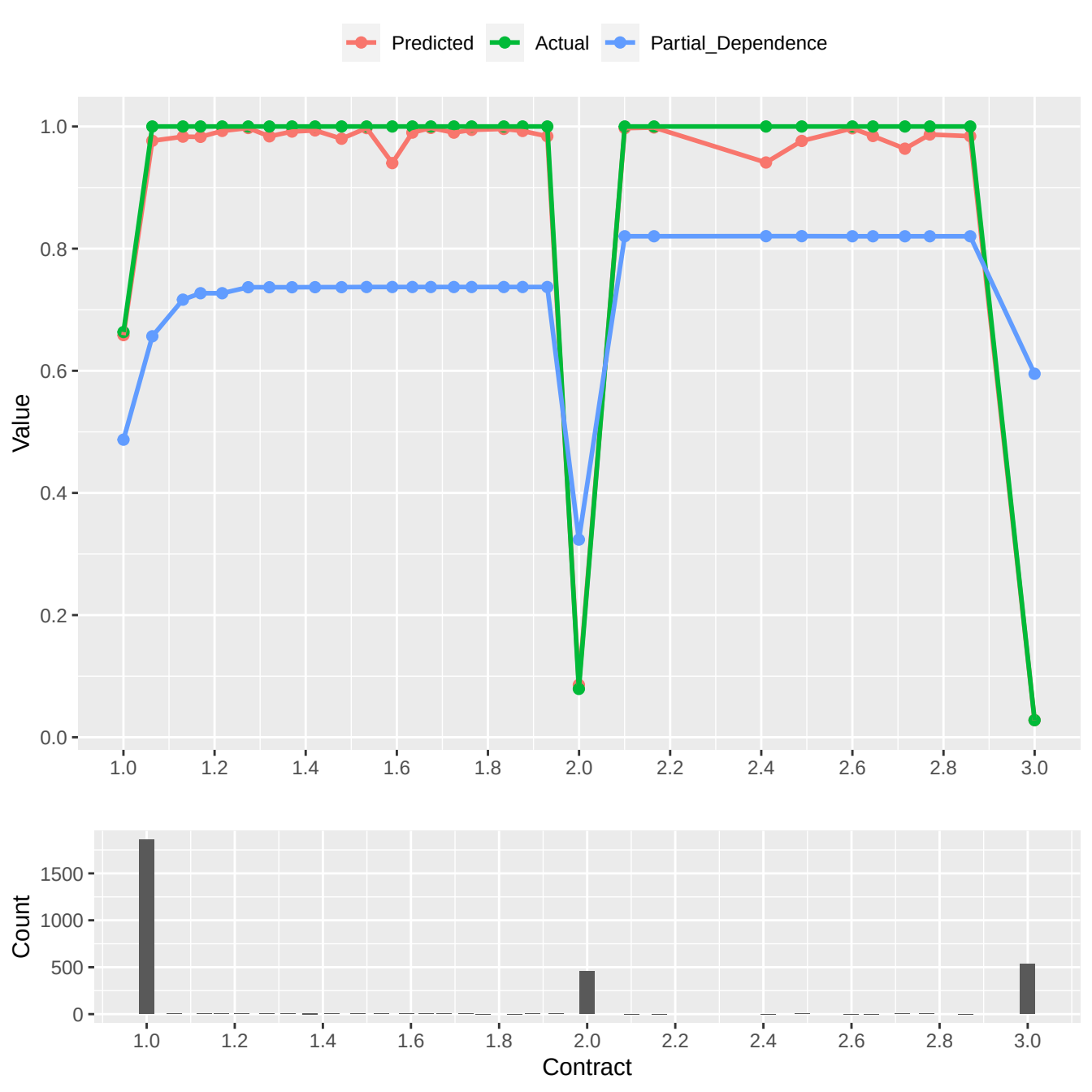

Predicted
Actual
Partial_Dependence
Value
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
10
20
30
40
50
60
70
Count
300
200
100
0
tenure

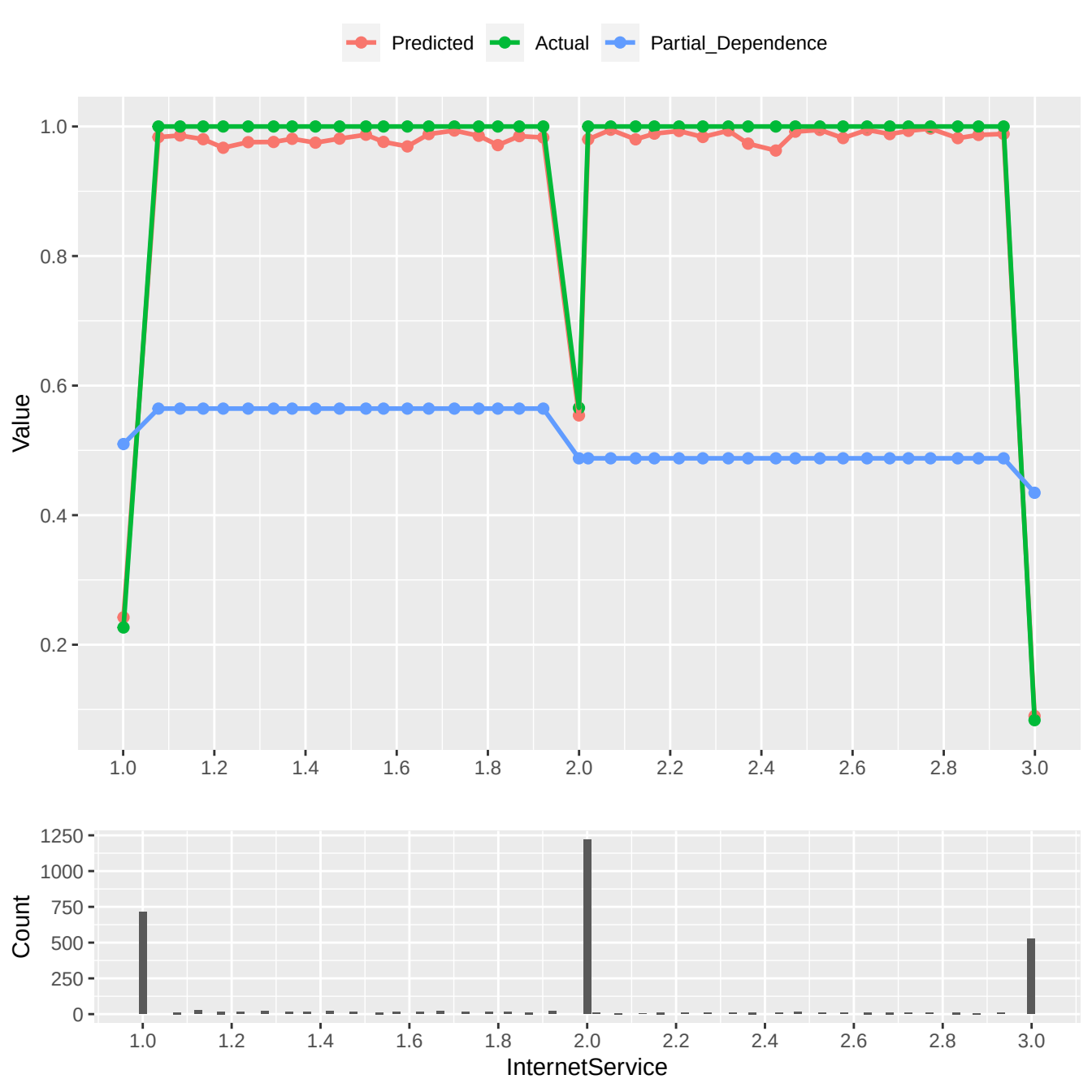

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0
1000
2000
3000
4000
5000
6000
7000
8000
Count
0
100
200
300
400
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
TotalCharges

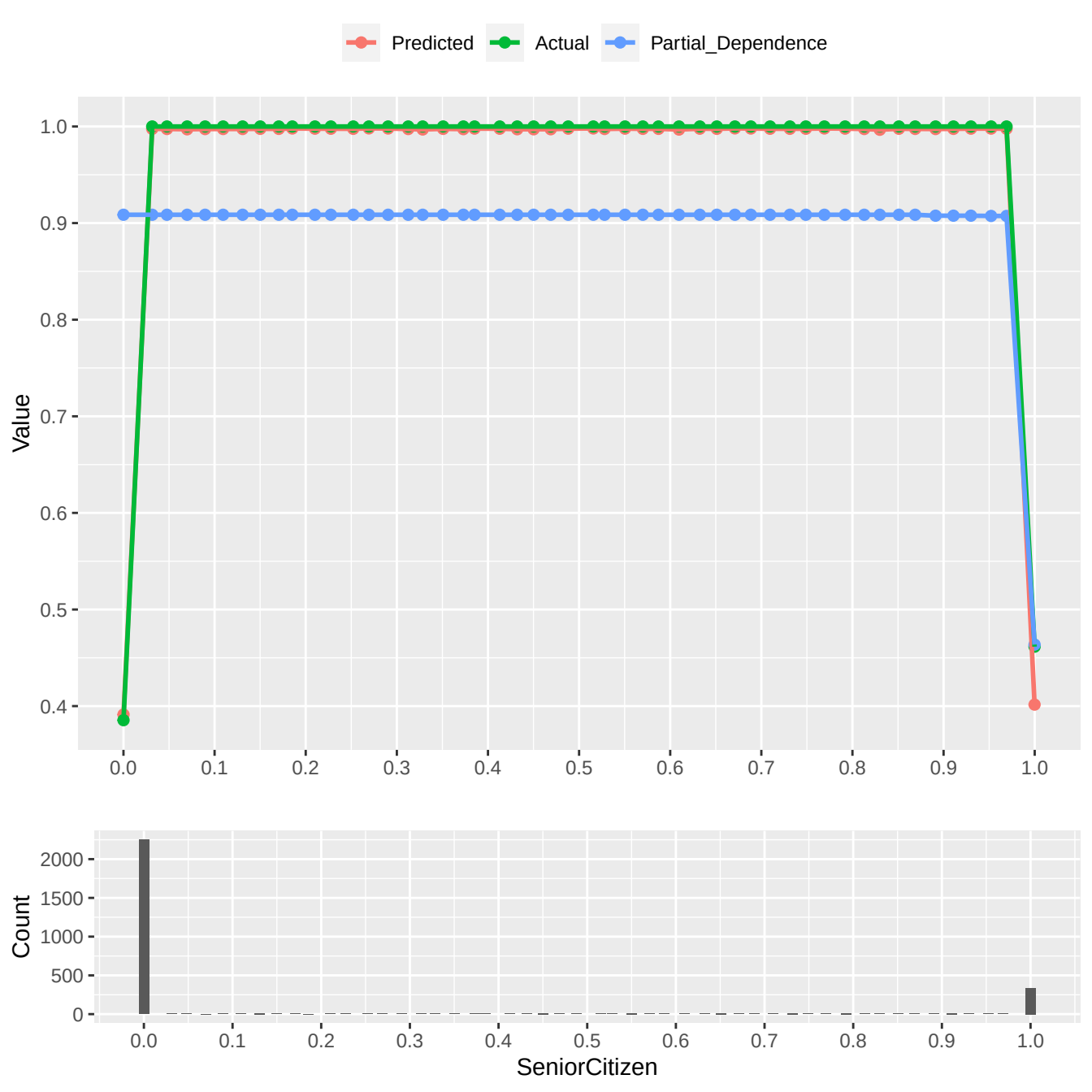

Predicted
Actual
Partial_Dependence
Value
0.7
0.6
0.5
0.4
0.3
0.2
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
1200
900
600
300
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
OnlineSecurity

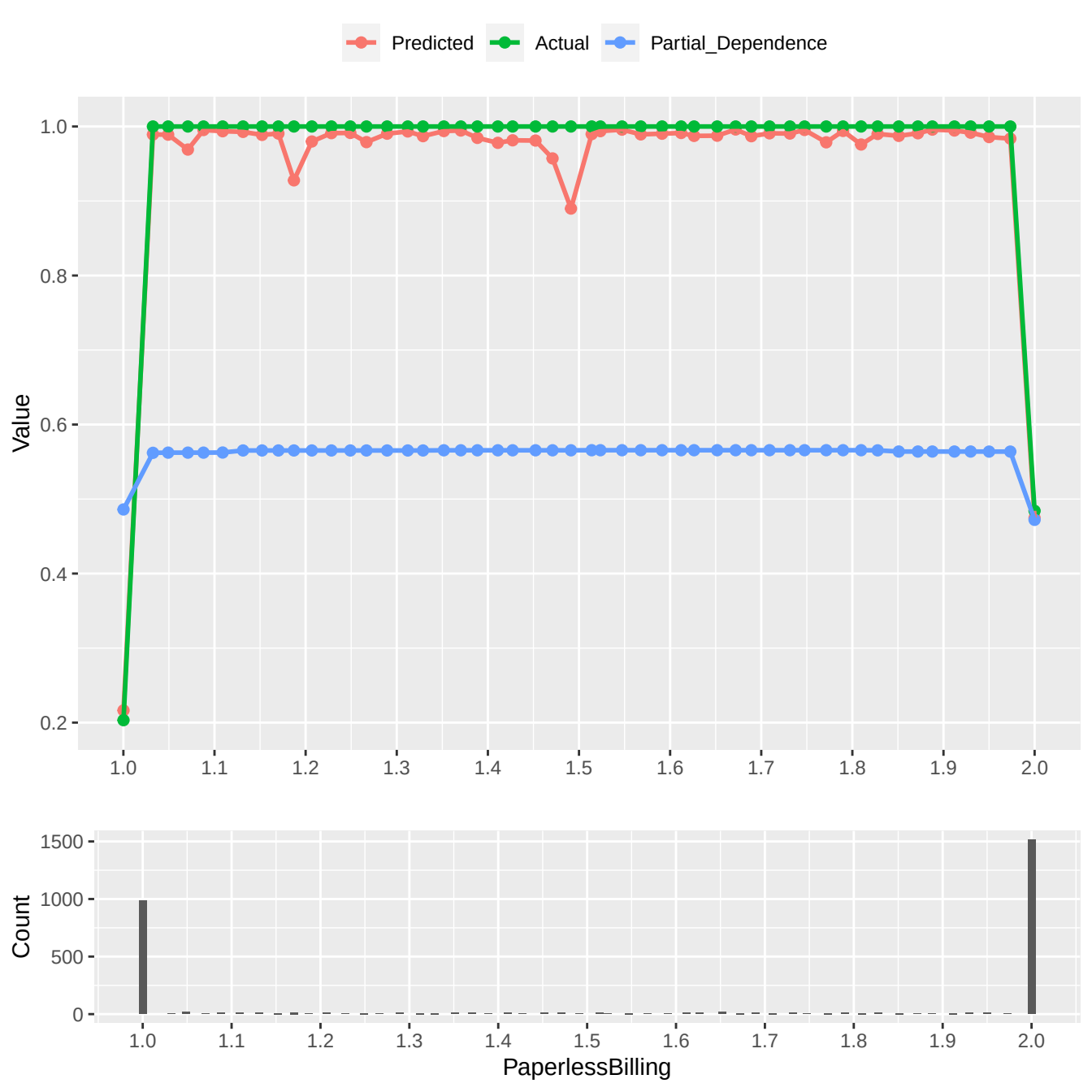

CV error
1.00
0.75
0.50
0.25
0.00
auc
0.96
0.950
0.975
1.000
1.025
1.05
Run

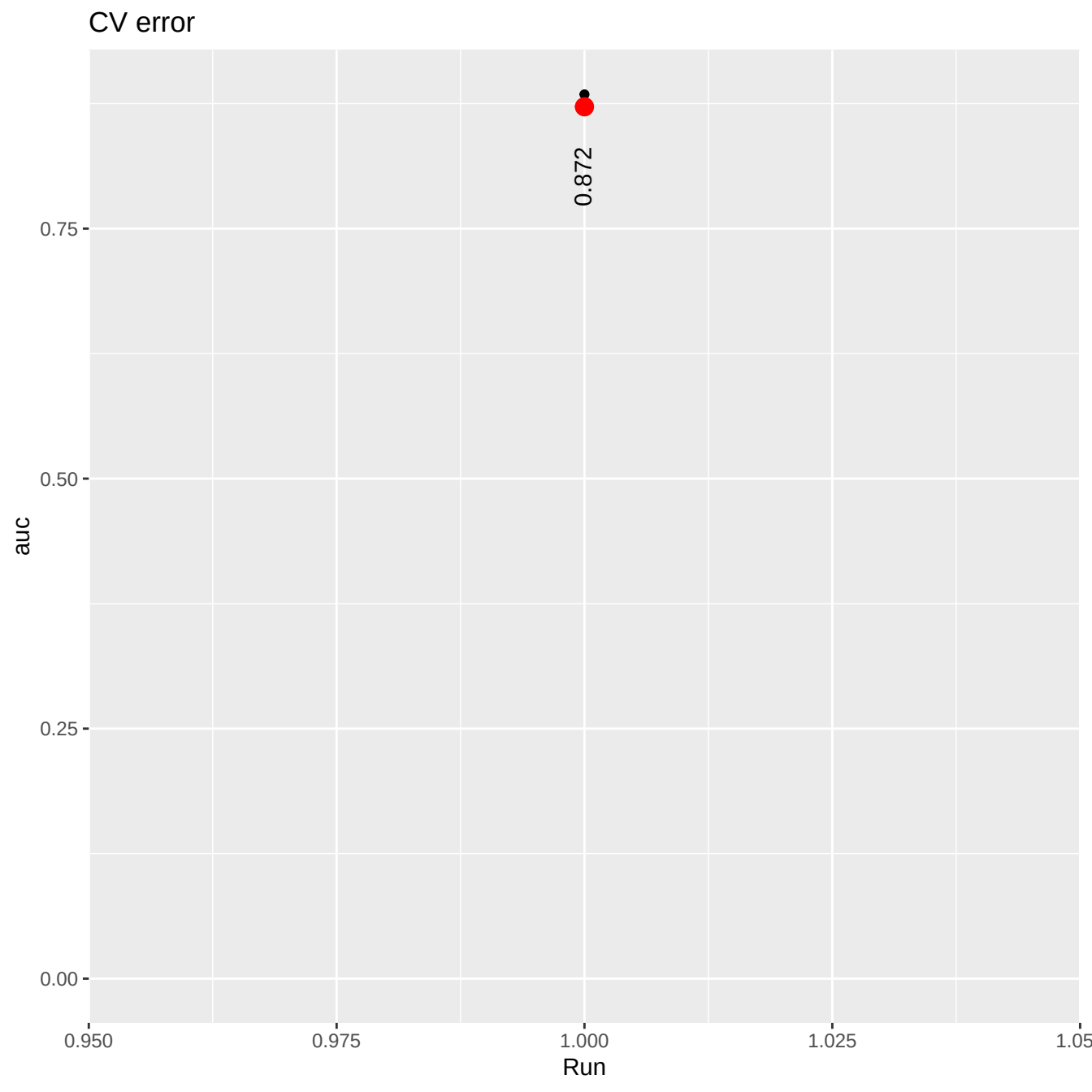

CV error
1.00
0.75
0.50
0.25
0.00
auc
0.96
1.00
1.25
1.50
1.75
2.00
Run

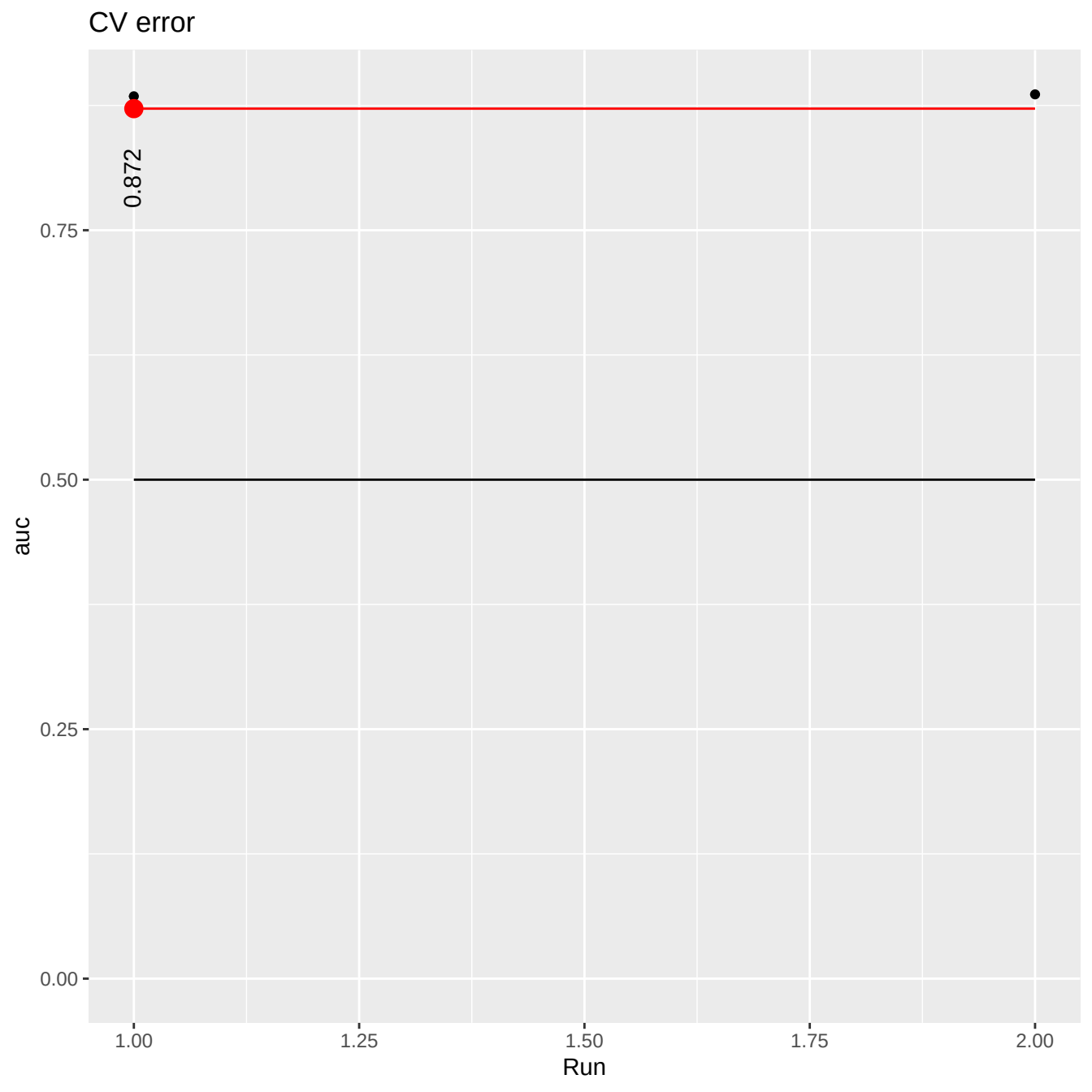

Predicted
Actual
Partial_Dependence
Value
1.0
0.8
0.6
0.4
0.2
0.0
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
1500
1000
500
0
Contract

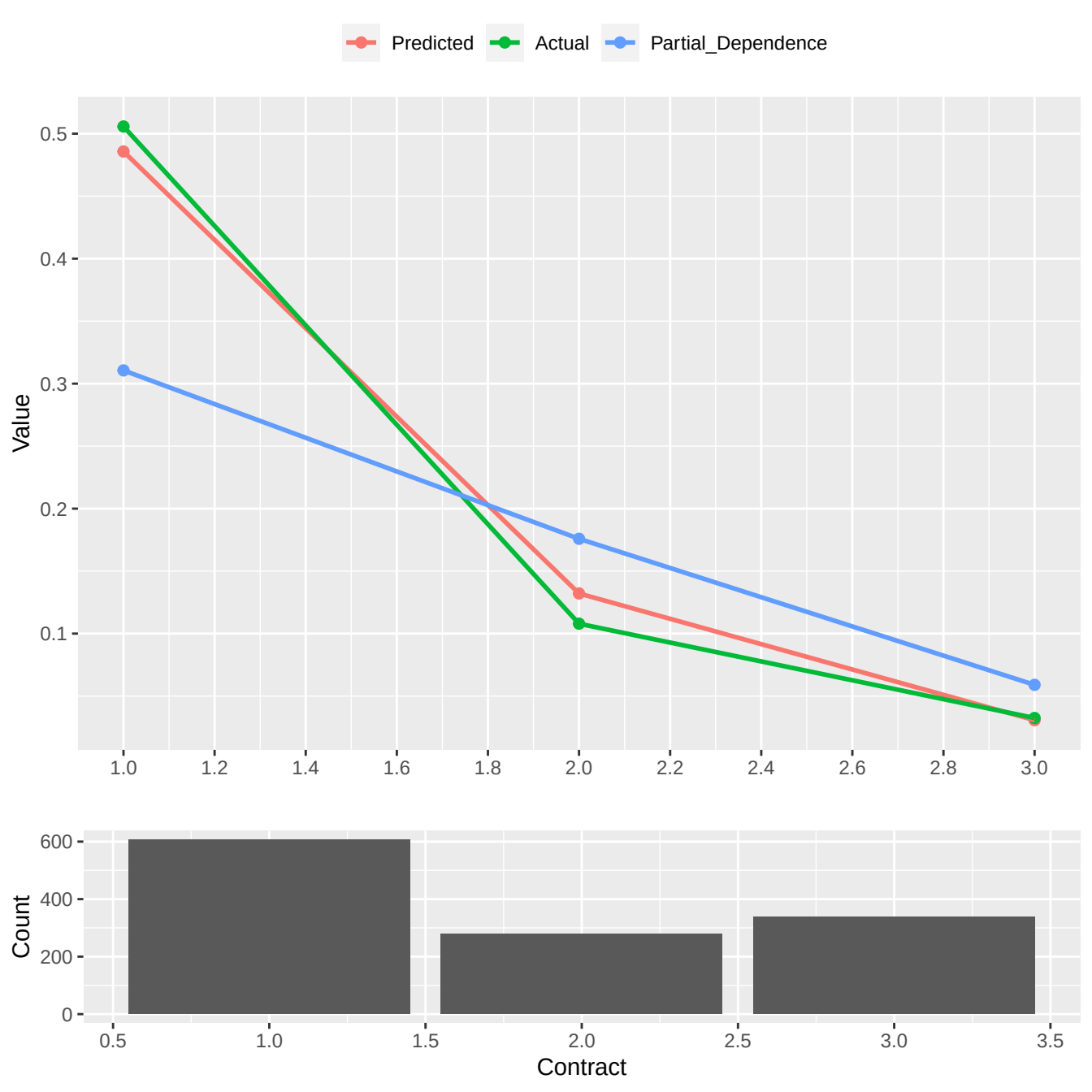

Predicted
Actual
Partial_Dependence
Value
1.0
0.8
0.6
0.4
0.2
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
1250
1000
750
500
250
0
InternetService

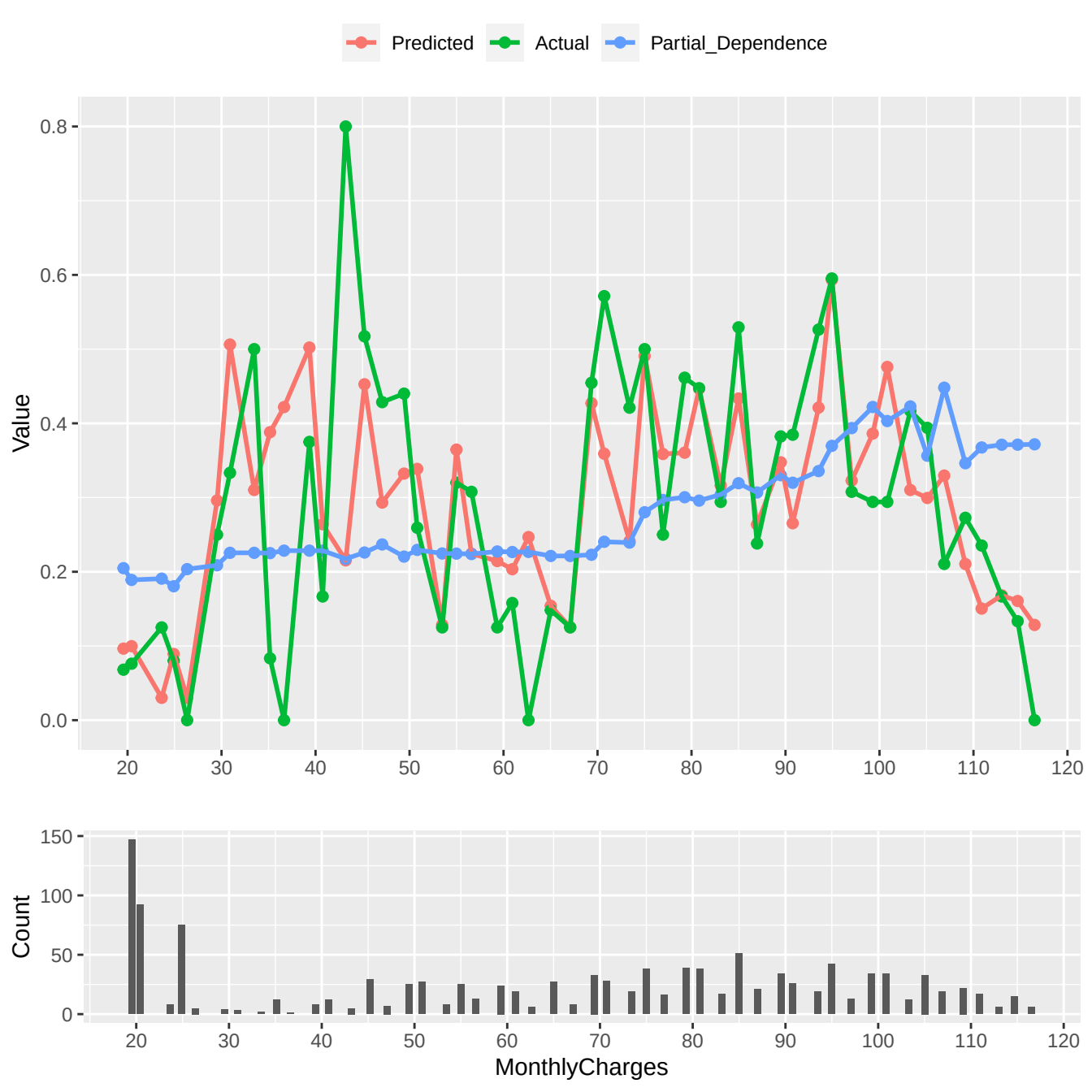

Predicted
Actual
Partial_Dependence
Value
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Count
2000
1500
1000
500
0
SeniorCitizen


Predicted
Actual
Partial_Dependence
Value
1.0
0.8
0.6
0.4
0.2
1.0
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
2.0
Count
1500
1000
500
0
PaperlessBilling

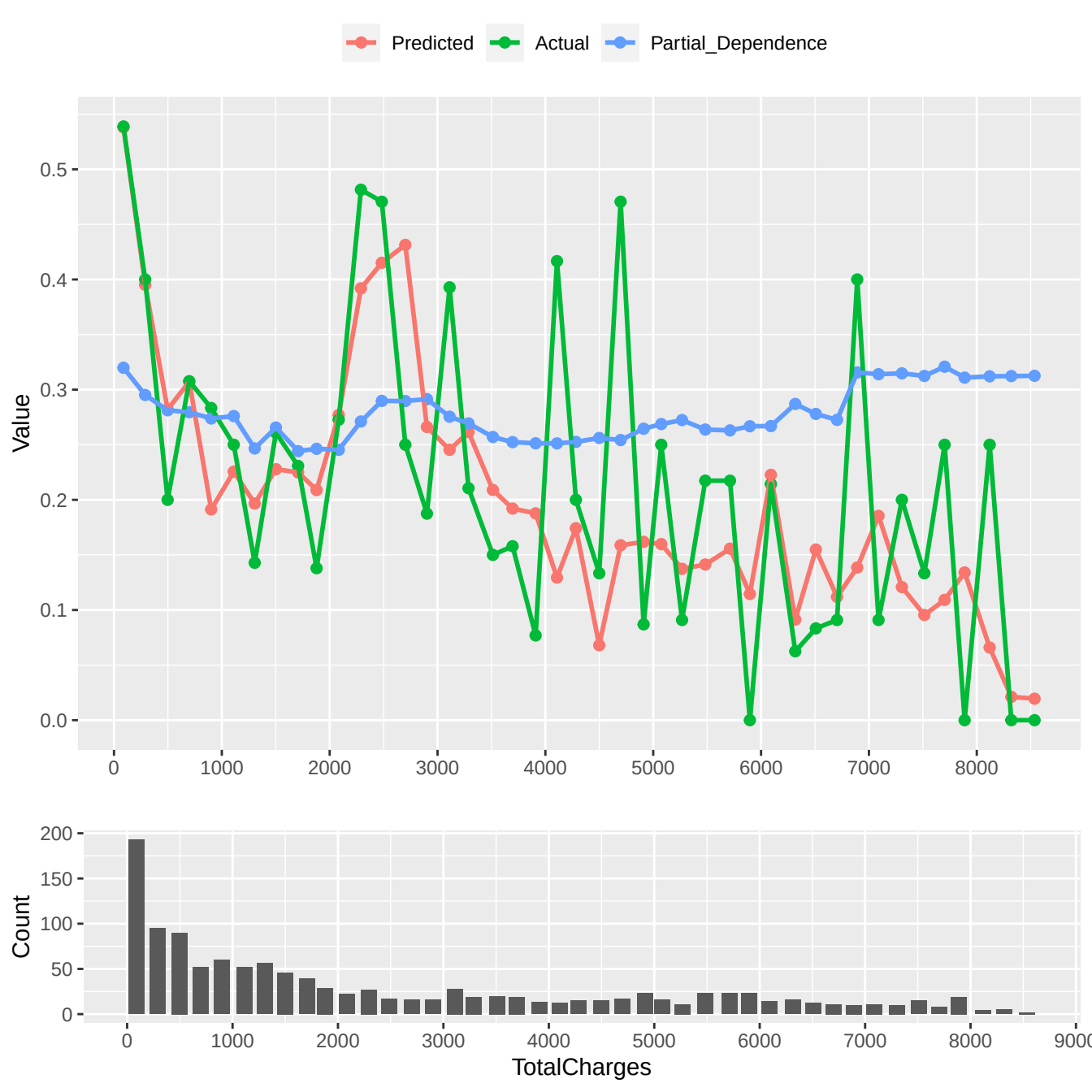

CV error
0.872
0.75
0.50
0.25
0.00
auc
0.950
0.975
1.000
1.025
1.05
Run

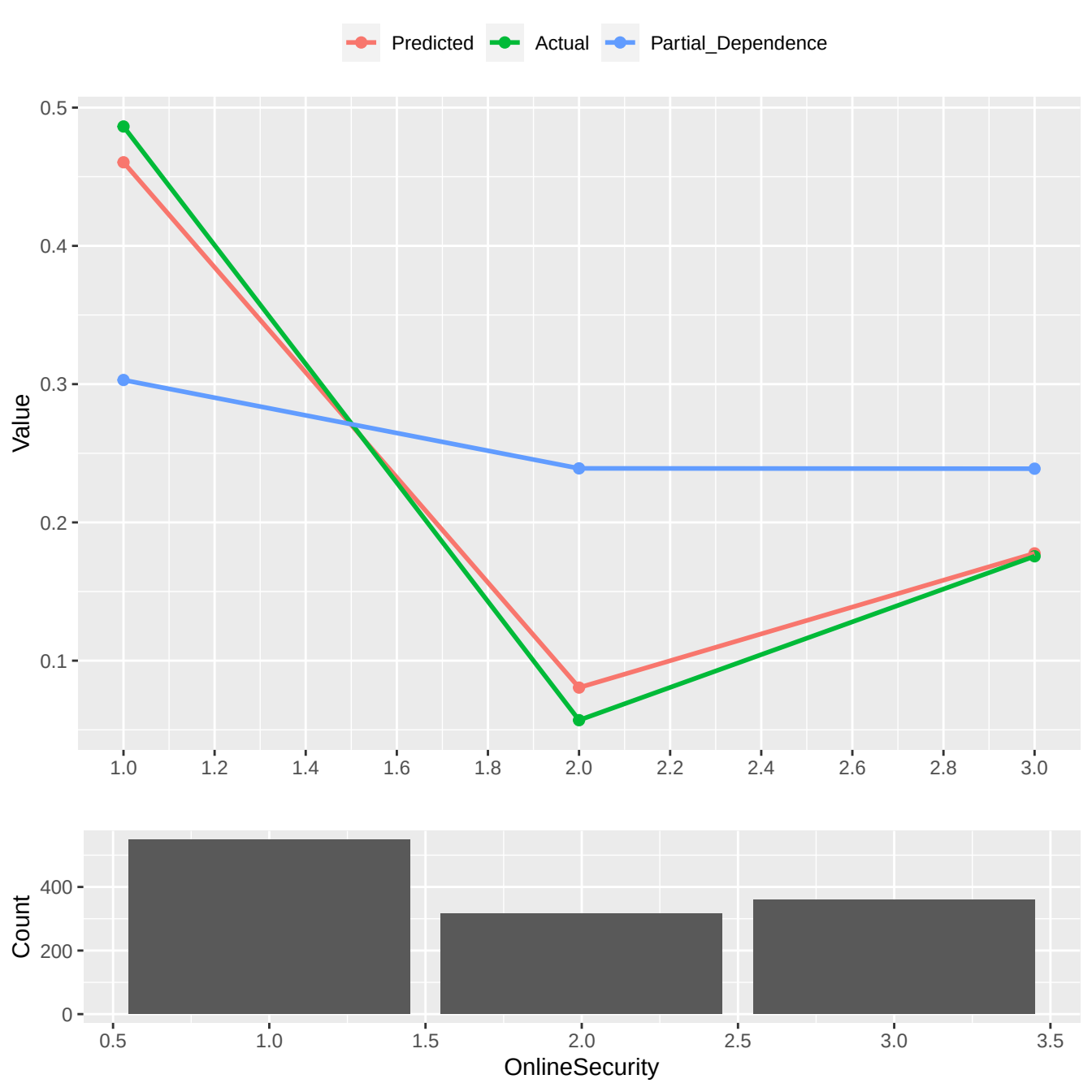

CV error
0.872
0.75
0.50
0.25
0.00
auc
1.00
1.25
1.50
1.75
2.00
Run

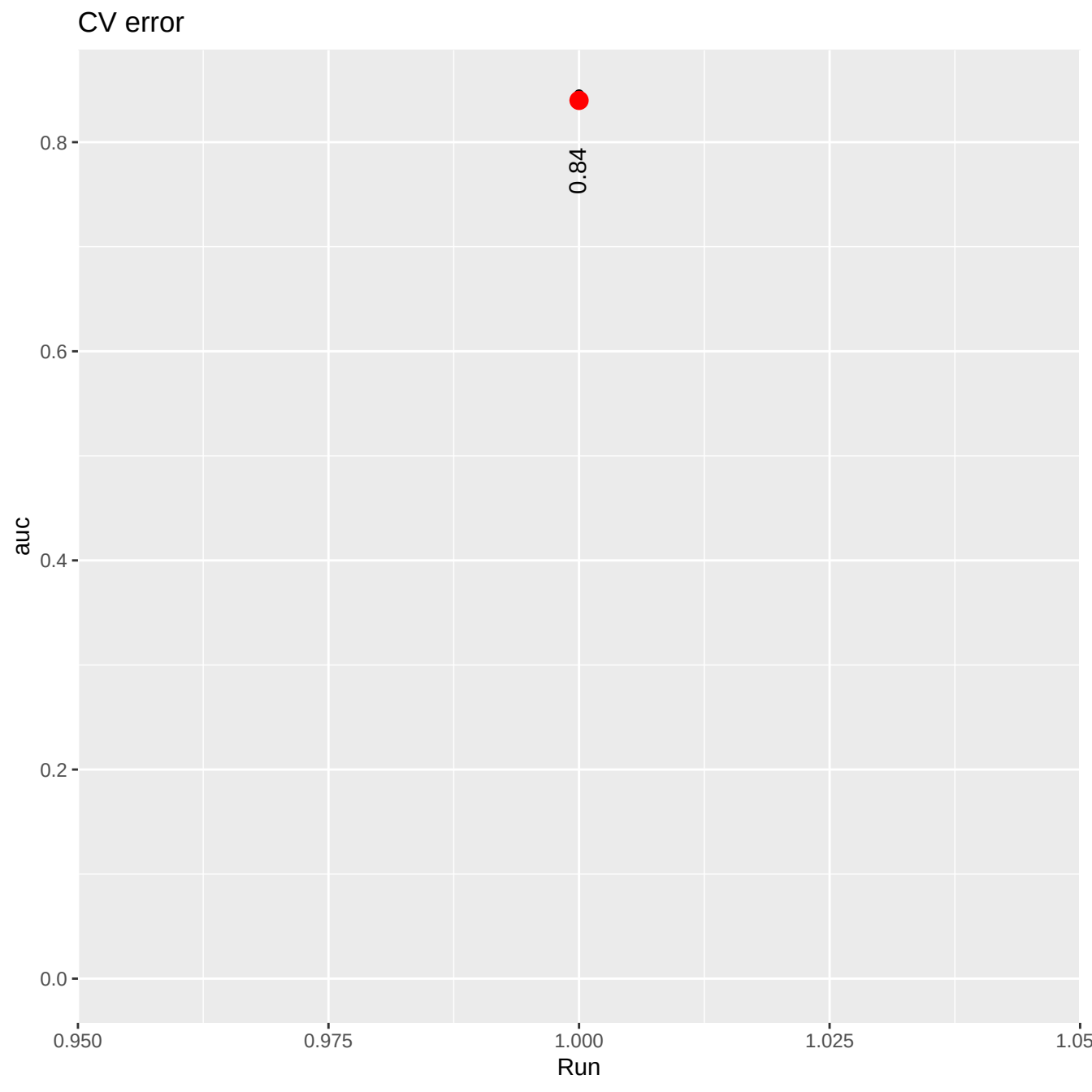

Predicted
Actual
Partial_Dependence
Value
0.5
0.4
0.3
0.2
0.1
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
600
400
200
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Contract

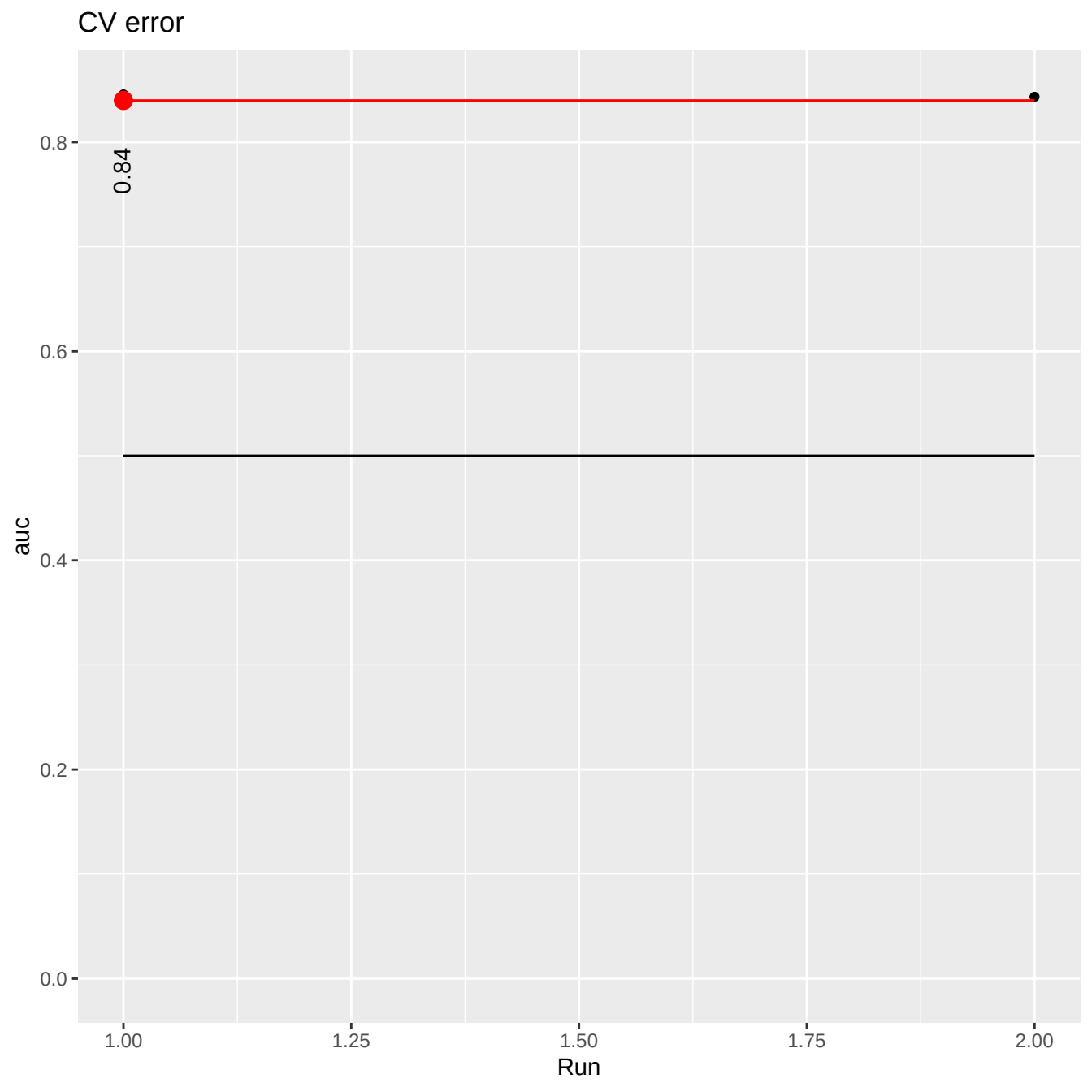

Predicted
Actual
Partial_Dependence
Value
0.0
0.2
0.4
0.6
0.8
20
30
40
50
60
70
80
90
100
110
120
Count
0
50
100
150
MonthlyCharges

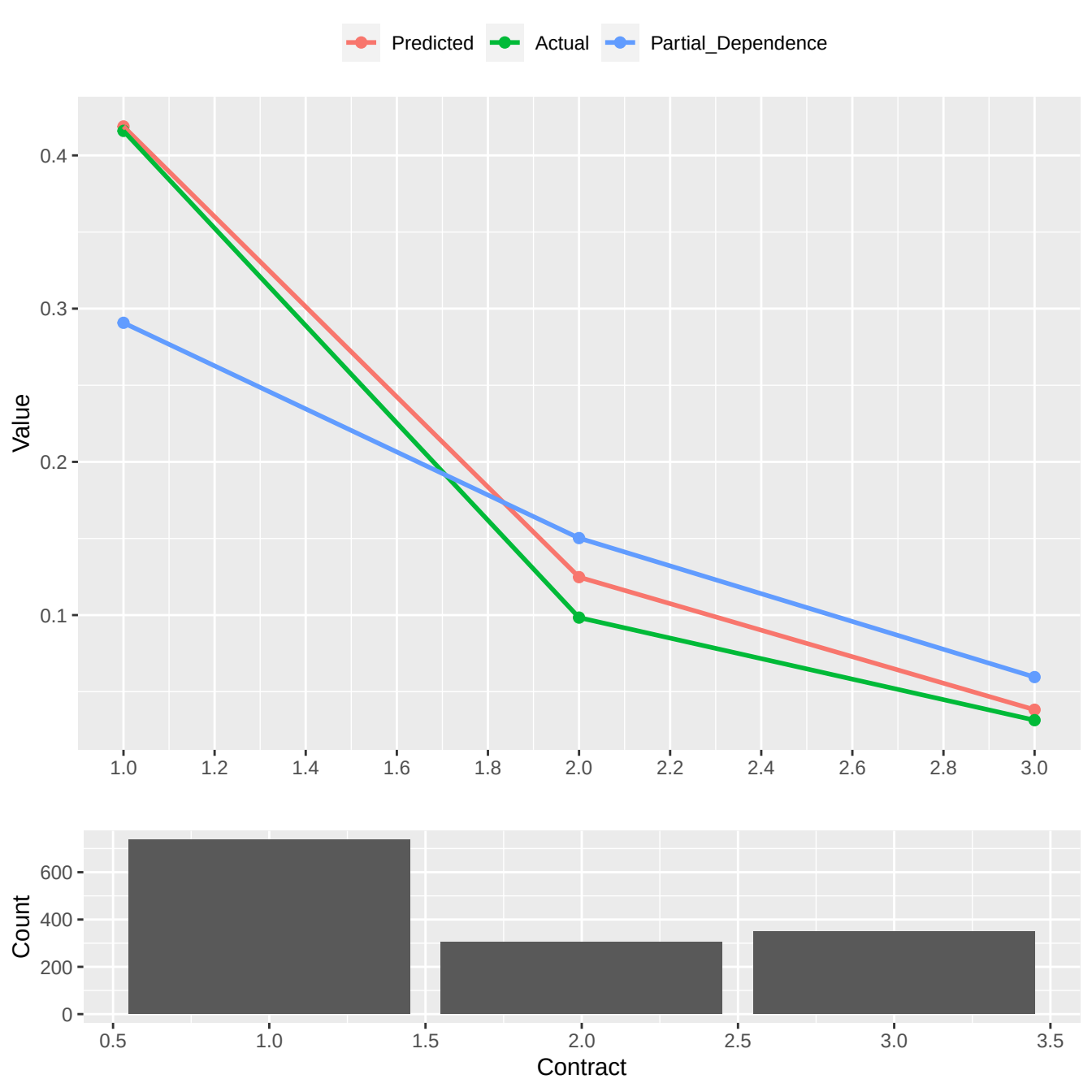

Predicted
Actual
Partial_Dependence
Value
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
10
20
30
40
50
60
70
Count
100
50
0
0
10
20
30
40
50
60
70
tenure

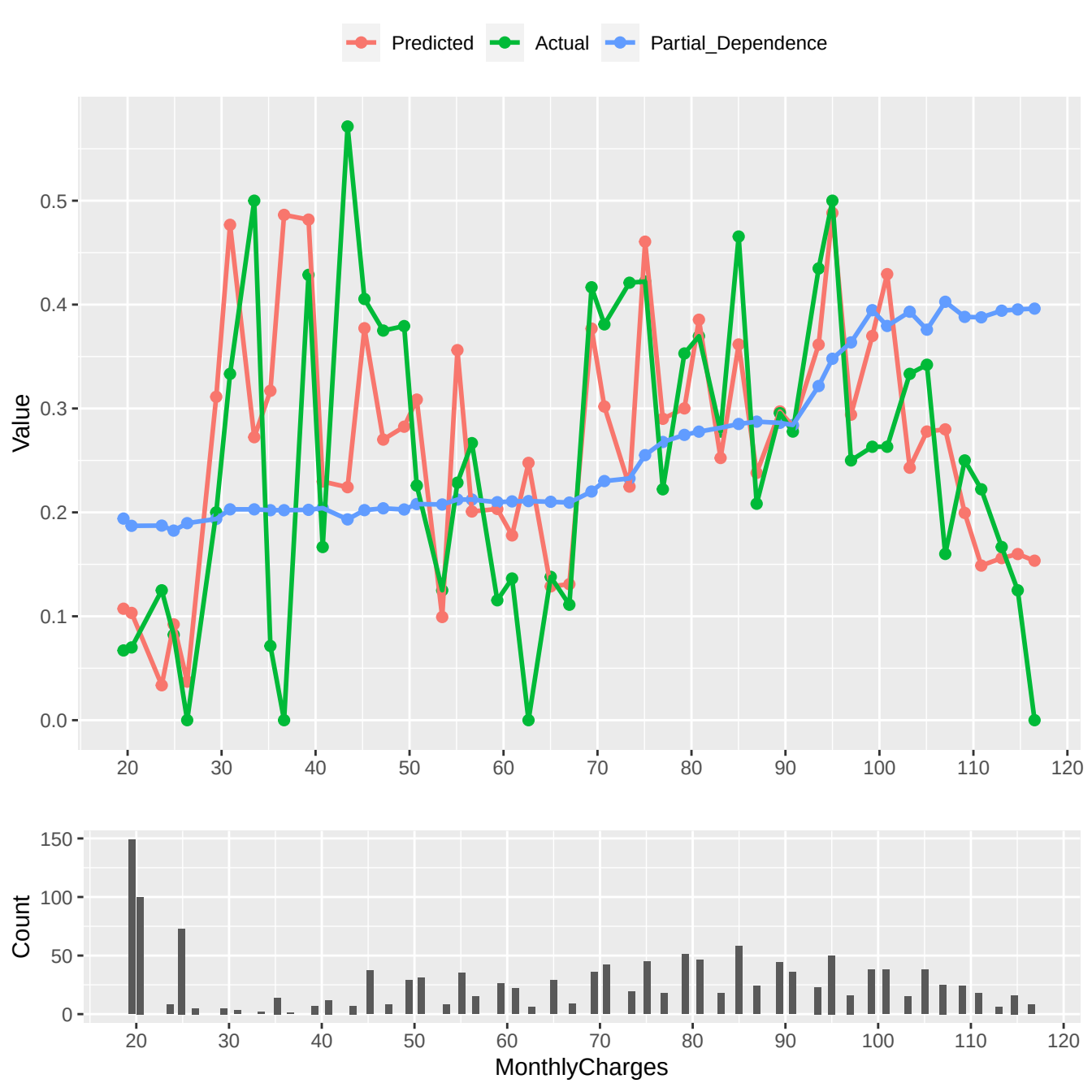

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
0
1000
2000
3000
4000
5000
6000
7000
8000
Count
0
50
100
150
200
9000
TotalCharges


Predicted
Actual
Partial_Dependence
Value
0.5
0.4
0.3
0.2
0.1
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
400
200
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
OnlineSecurity

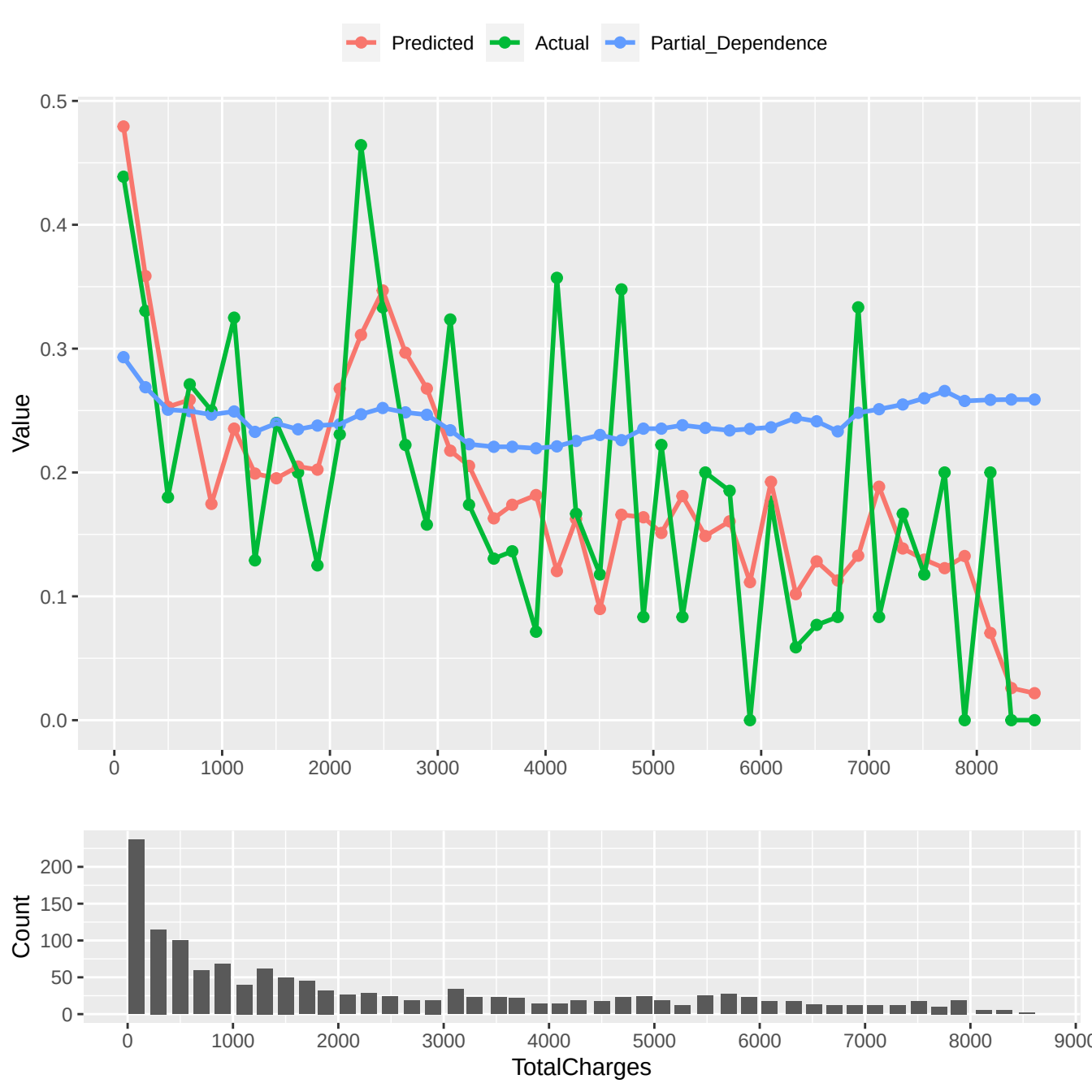

CV error
0.84
auc
0.0
0.2
0.4
0.6
0.8
0.950
0.975
1.000
1.025
1.05
Run

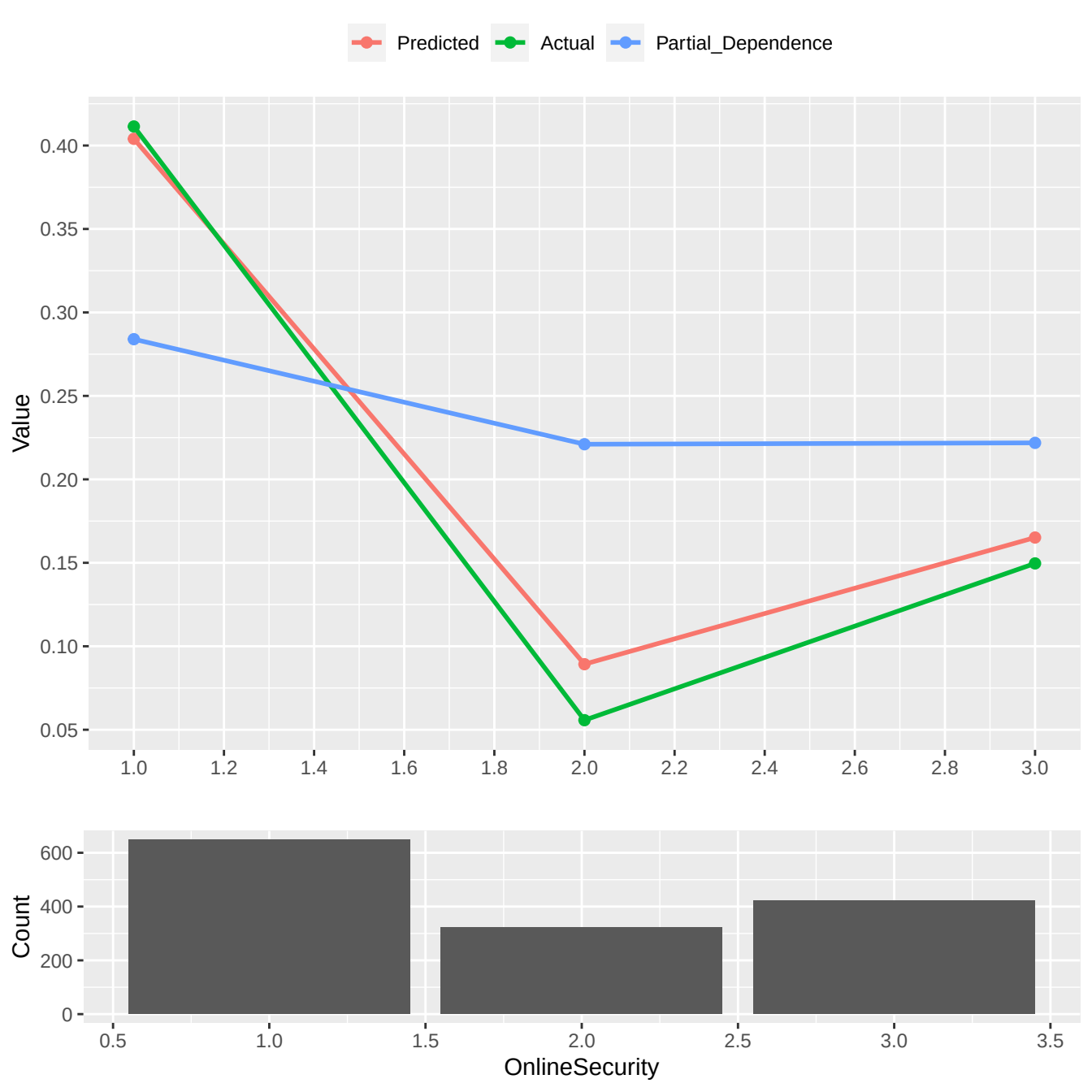

CV error
0.84
auc
0.0
0.2
0.4
0.6
0.8
1.00
1.25
1.50
1.75
2.00
Run

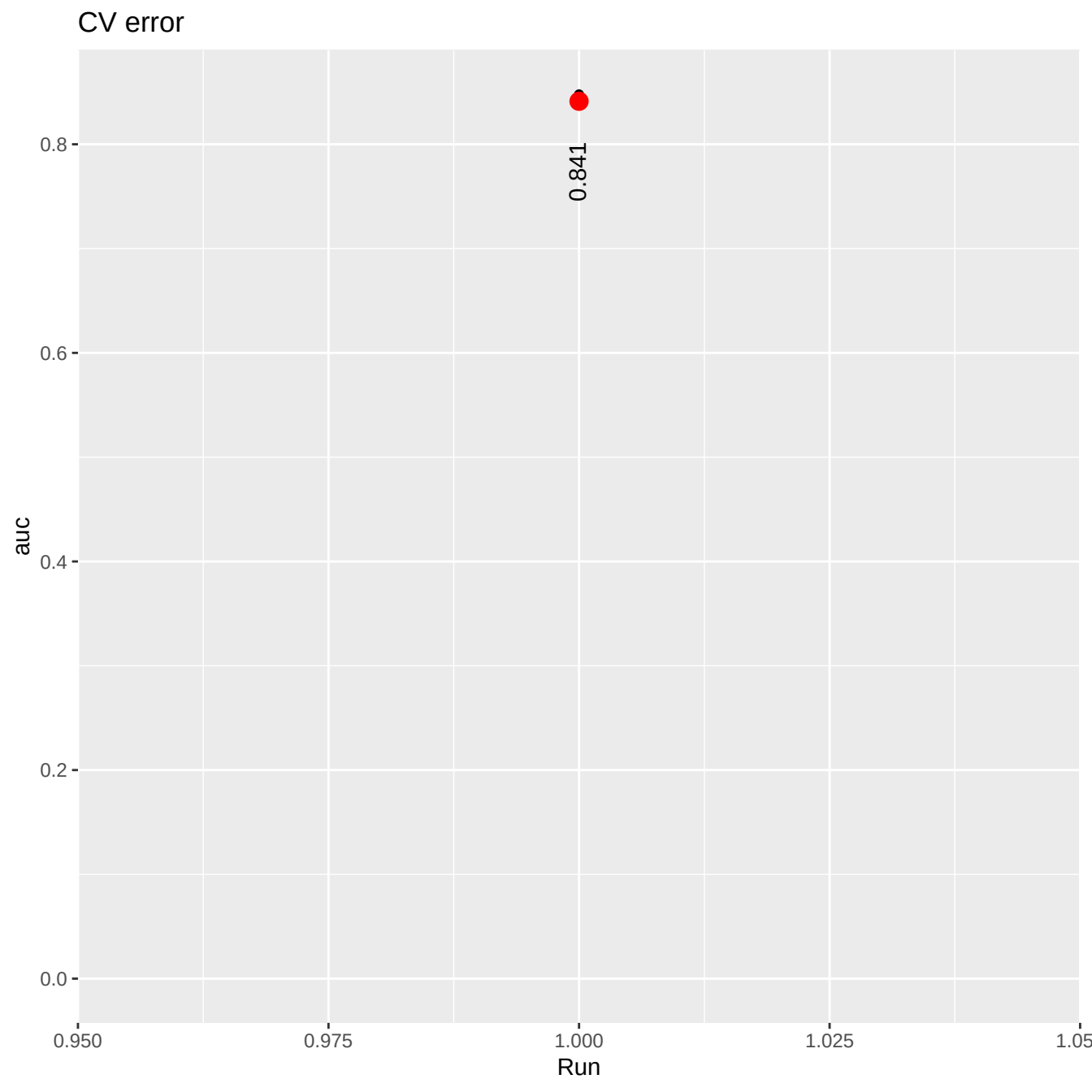

Predicted
Actual
Partial_Dependence
Value
0.4
0.3
0.2
0.1
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
600
400
200
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Contract

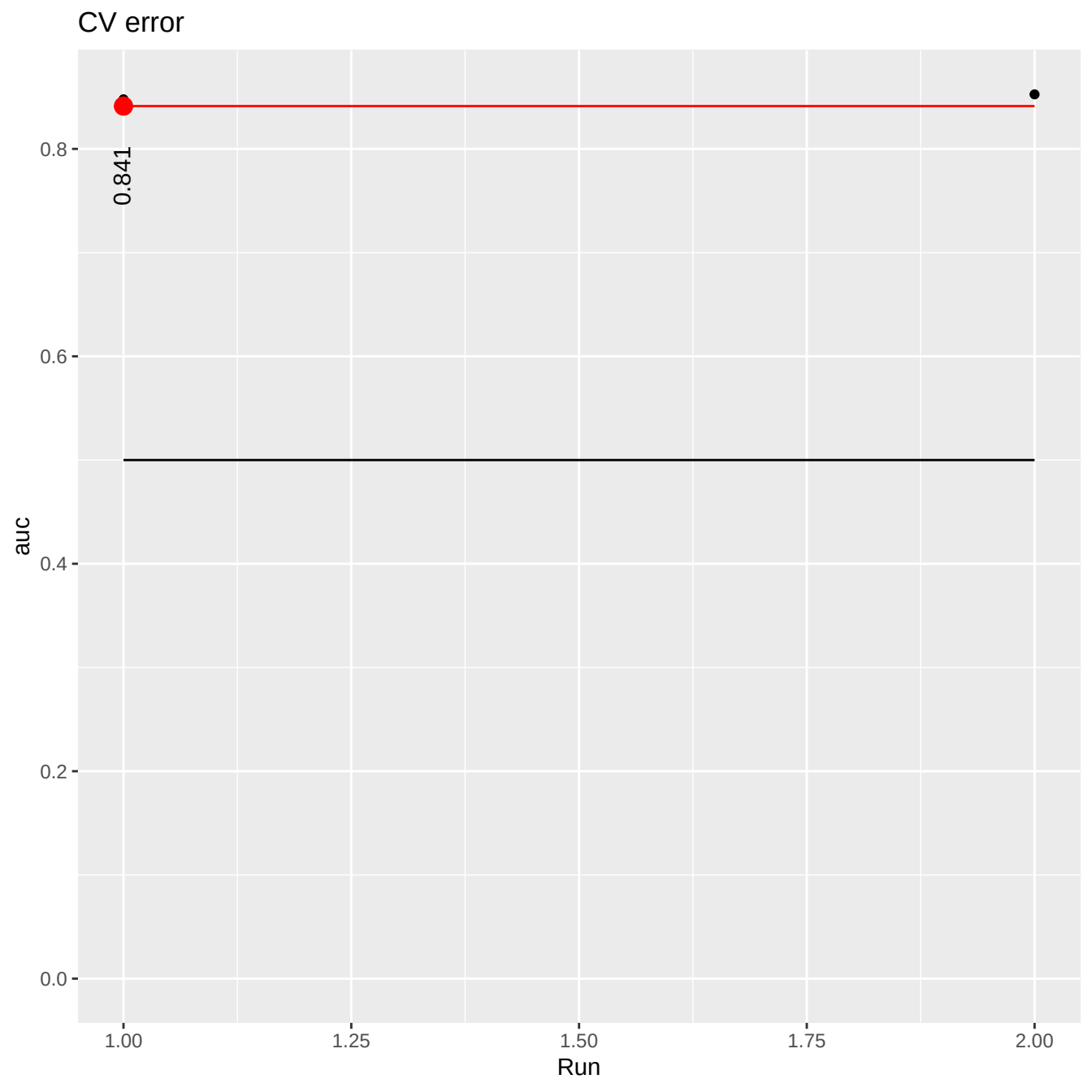

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
20
30
40
50
60
70
80
90
100
110
120
Count
0
50
100
150
MonthlyCharges

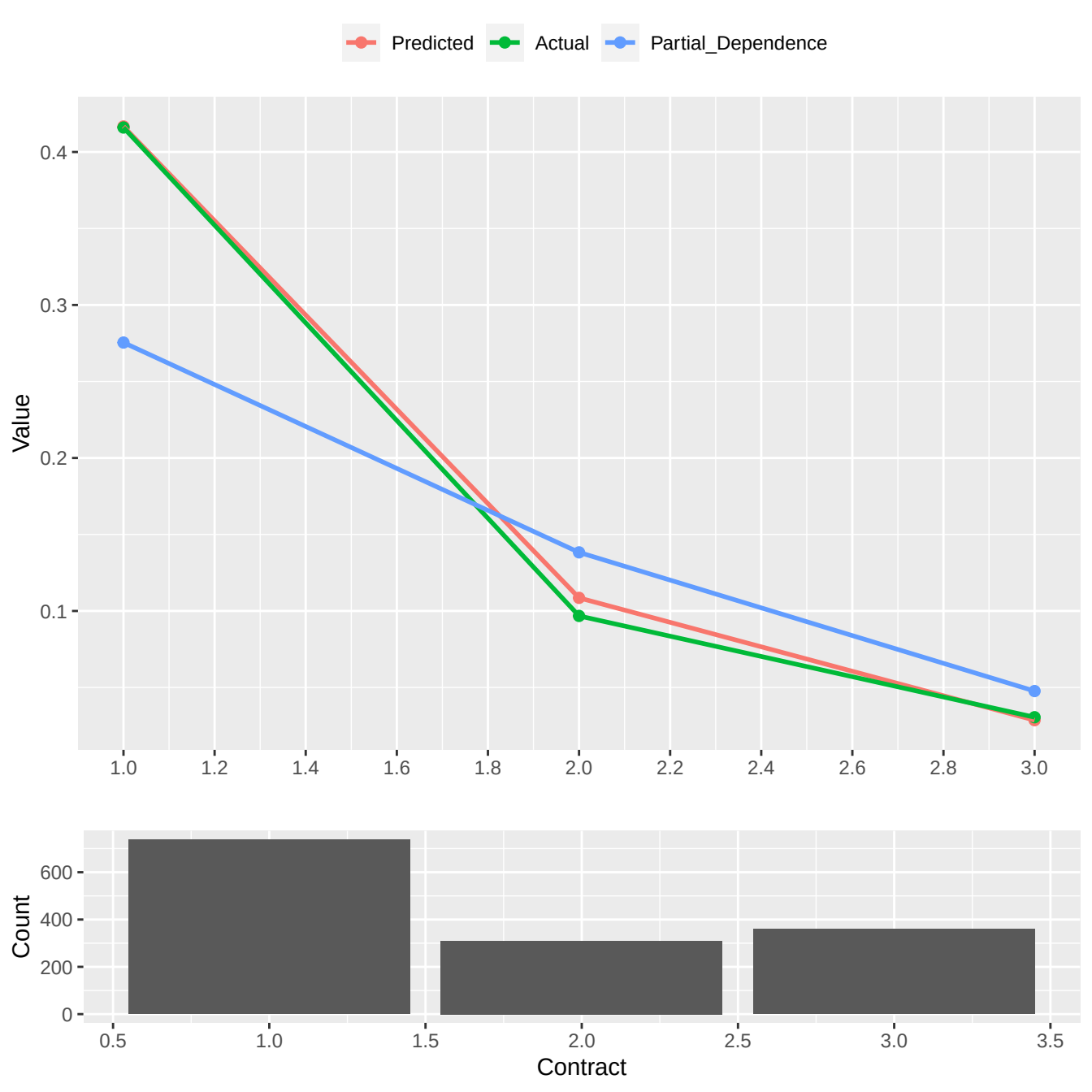

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
0
10
20
30
40
50
60
70
Count
0
50
100
150
tenure

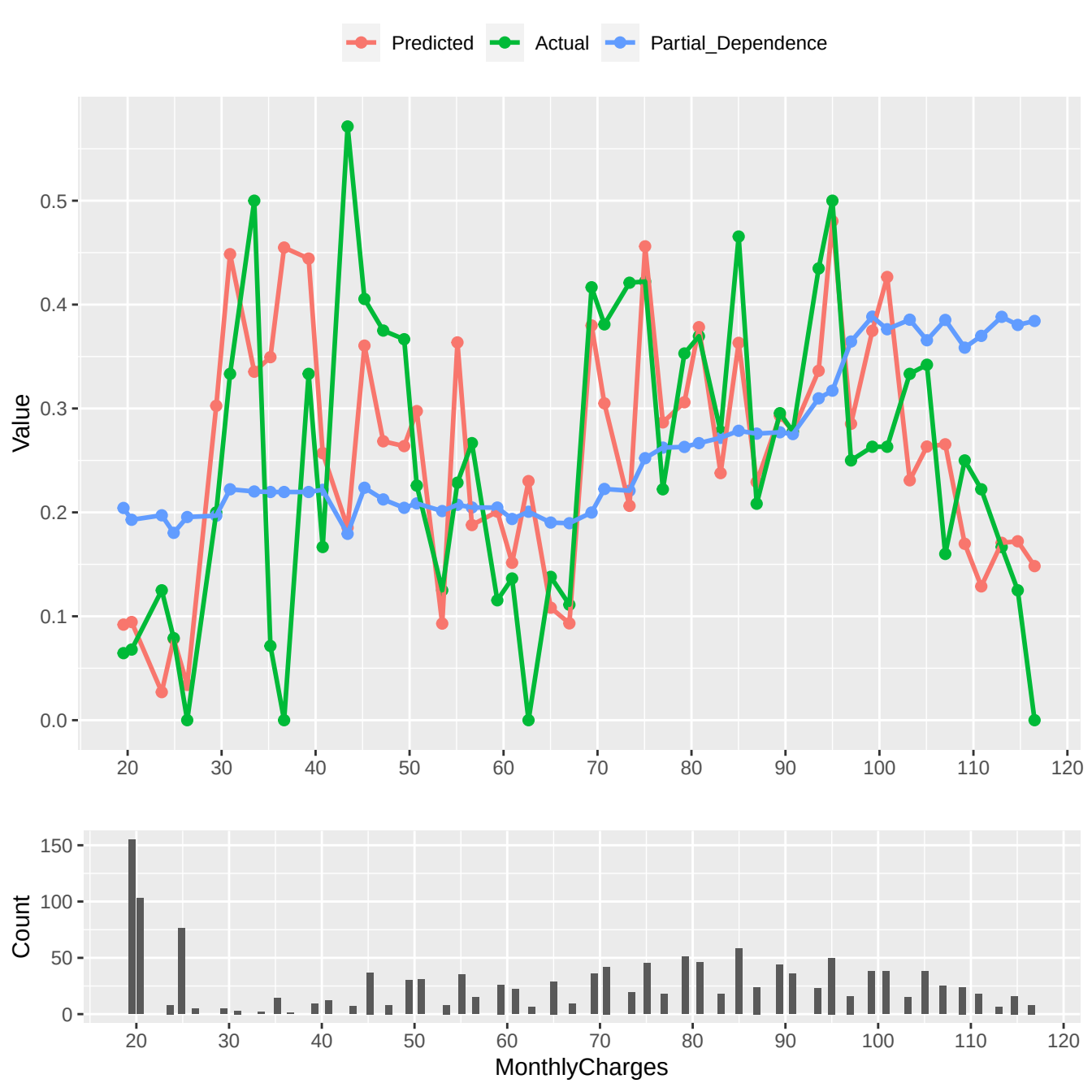

Predicted
Actual
Partial_Dependence
Value
0.5
0.4
0.3
0.2
0.1
0.0
0
1000
2000
3000
4000
5000
6000
7000
8000
Count
200
150
100
50
0
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
TotalCharges


Predicted
Actual
Partial_Dependence
Value
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
600
400
200
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
OnlineSecurity


CV error
0.841
auc
0.0
0.2
0.4
0.6
0.8
0.950
0.975
1.000
1.025
1.05
Run

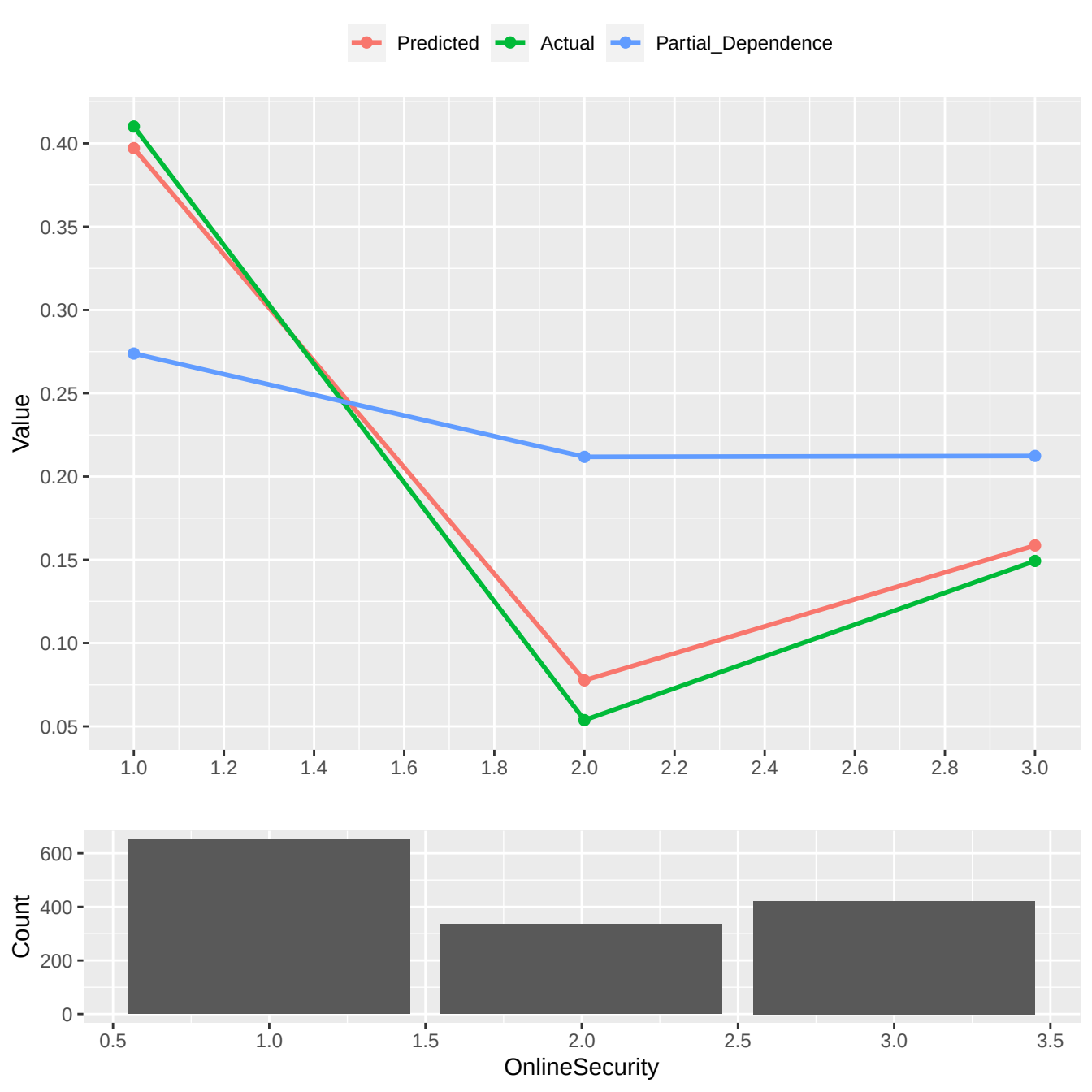

CV error
0.841
0.8
0.6
0.4
0.2
0.0
auc
1.00
1.25
1.50
1.75
2.00
Run

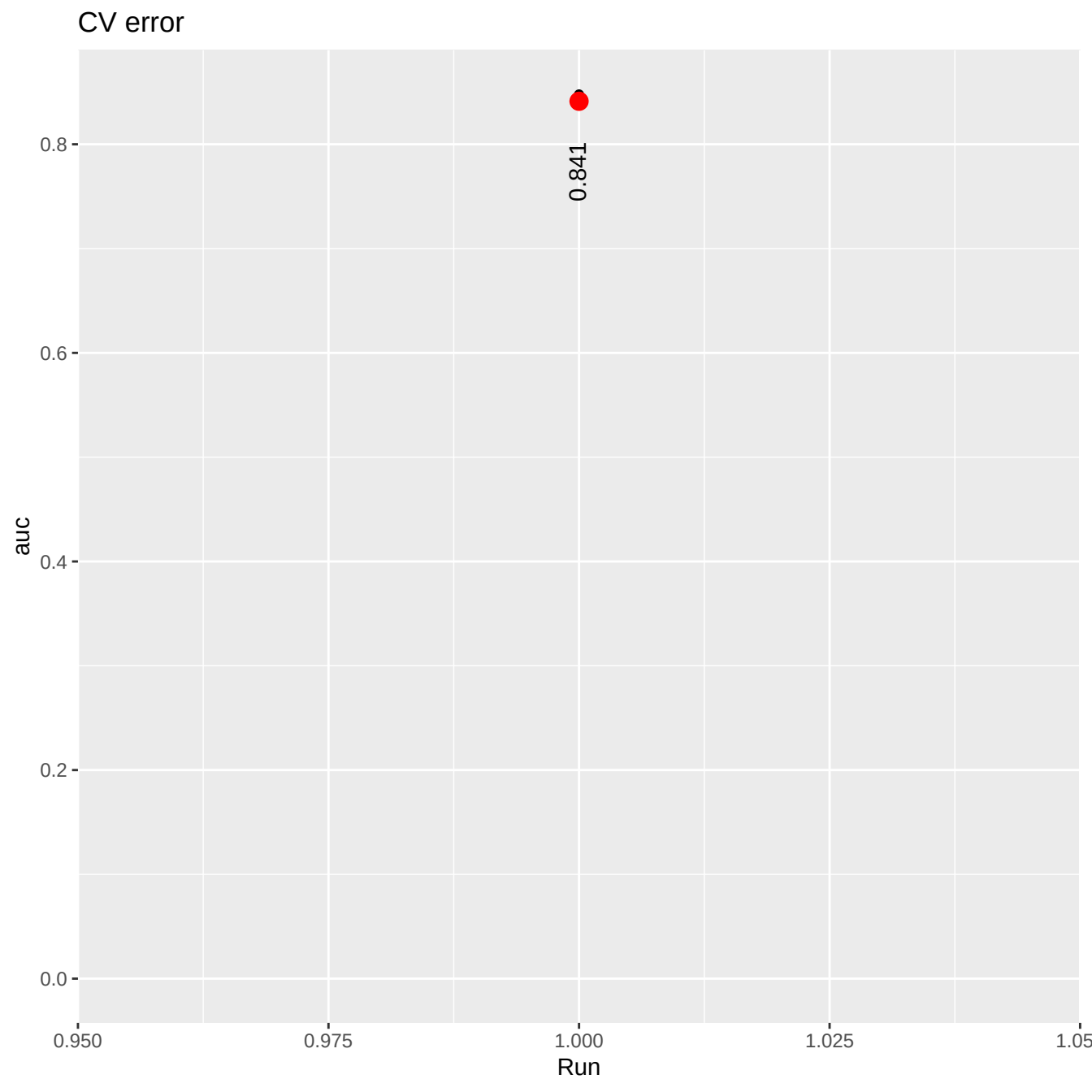

Predicted
Actual
Partial_Dependence
Value
0.4
0.3
0.2
0.1
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
600
400
200
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Contract

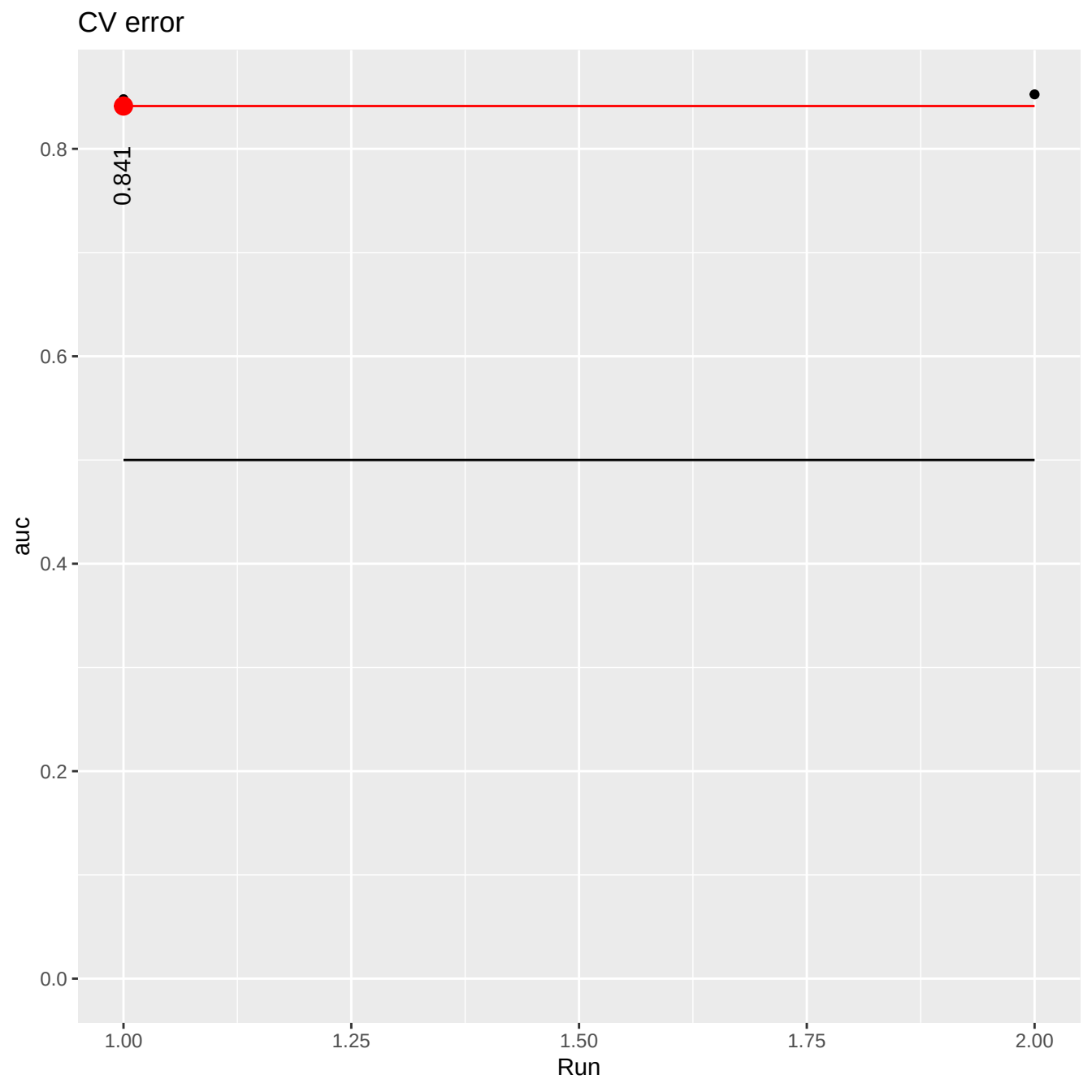

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
20
30
40
50
60
70
80
90
100
110
120
Count
0
50
100
150
MonthlyCharges

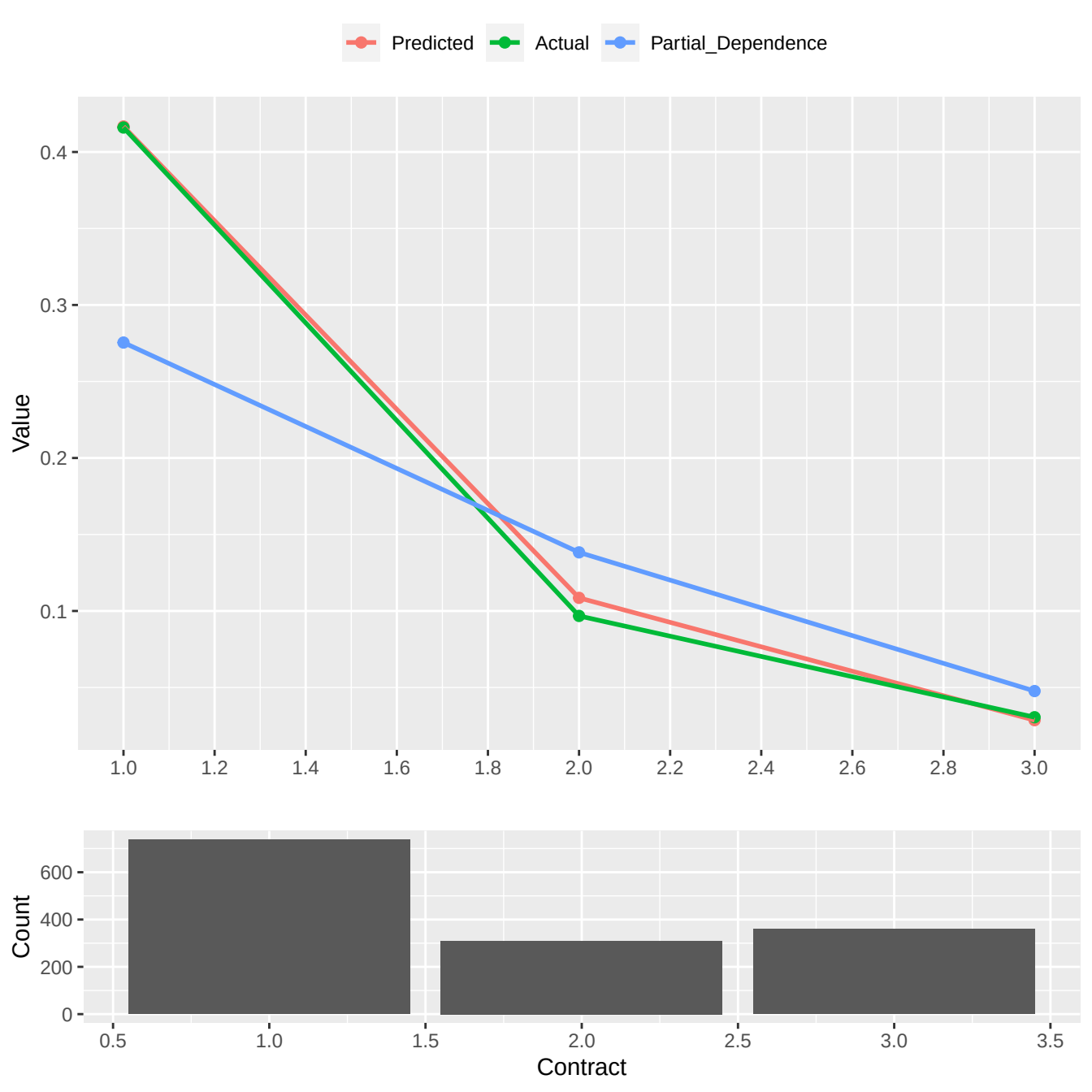

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
10
20
30
40
50
60
70
Count
0
50
100
150
tenure

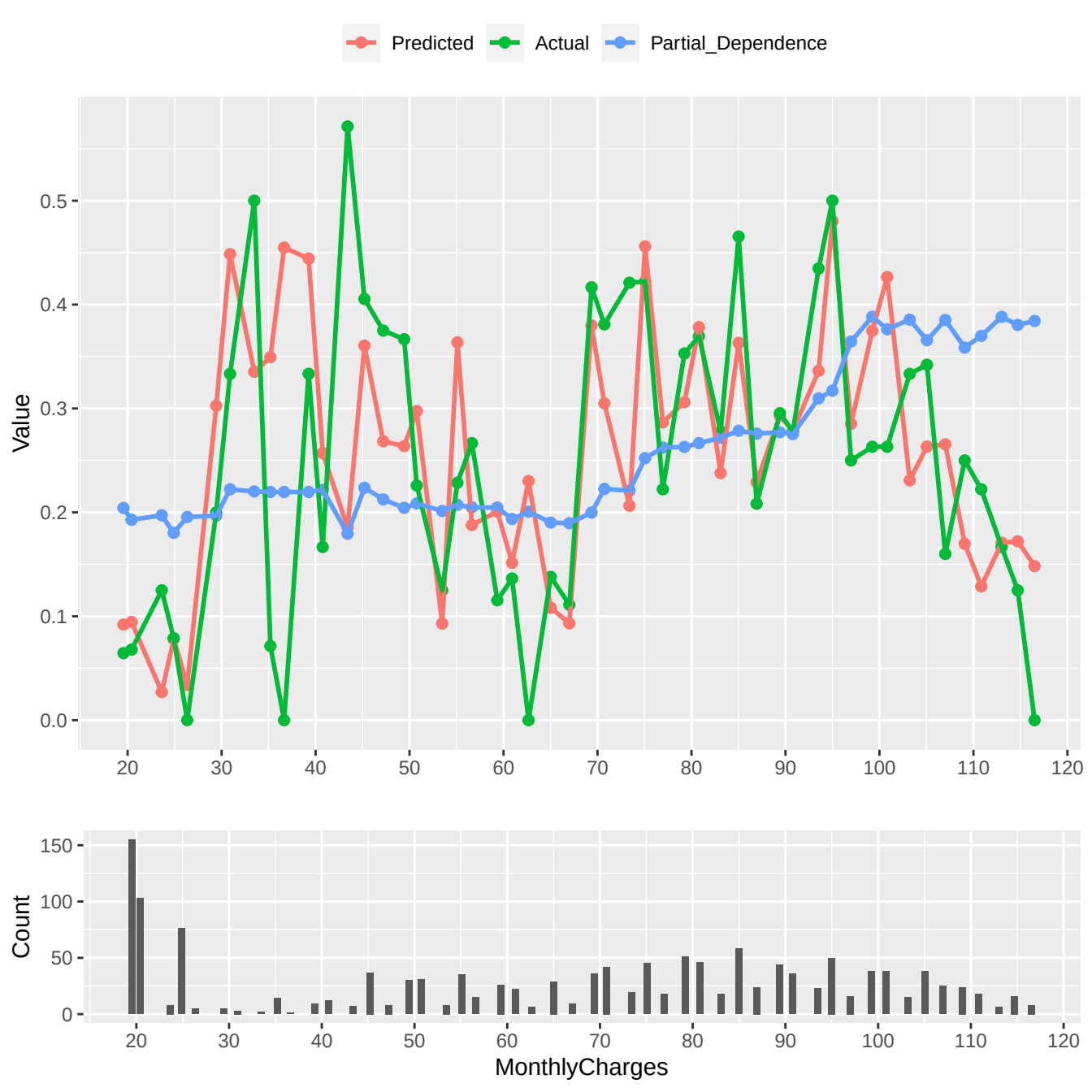

Predicted
Actual
Partial_Dependence
Value
0.5
0.4
0.3
0.2
0.1
0.0
0
1000
2000
3000
4000
5000
6000
7000
8000
Count
200
150
100
50
0
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
TotalCharges


Predicted
Actual
Partial_Dependence
Value
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
600
400
200
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
OnlineSecurity


CV error
0.841
auc
0.0
0.2
0.4
0.6
0.8
0.950
0.975
1.000
1.025
1.05
Run

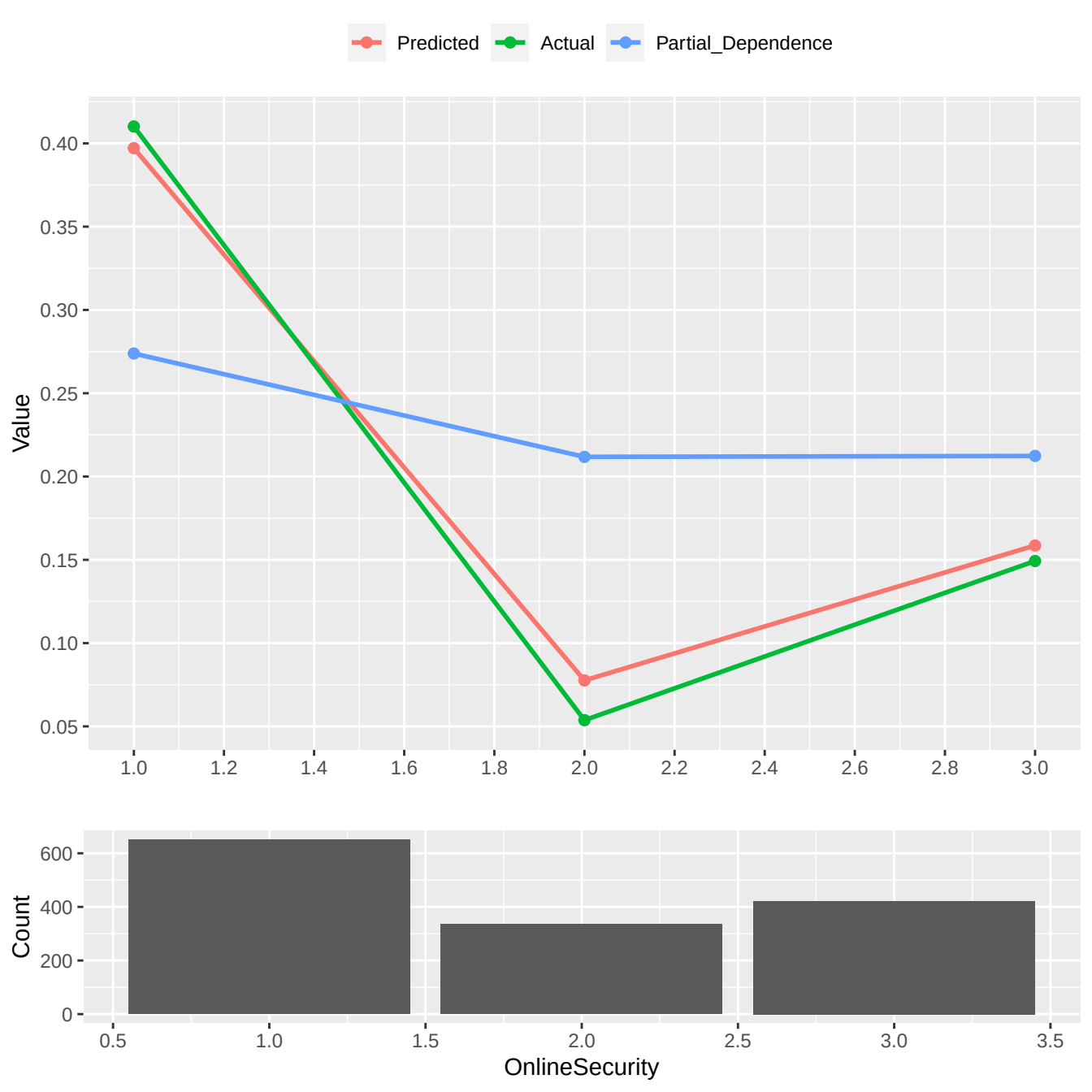

CV error
0.841
0.8
0.6
0.4
0.2
0.0
auc
1.00
1.25
1.50
1.75
2.00
Run

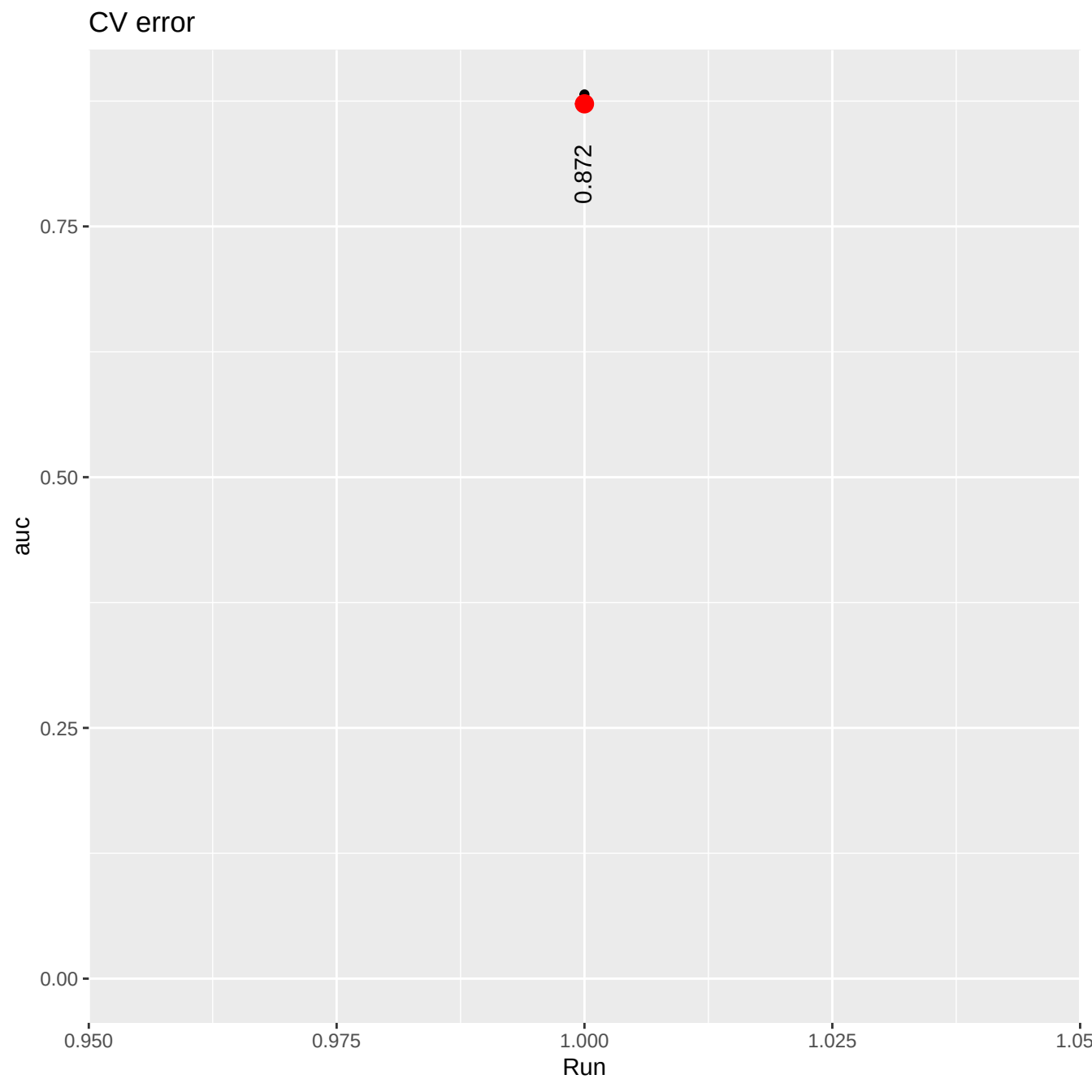

Predicted
Actual
Partial_Dependence
Value
0.4
0.3
0.2
0.1
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
600
400
200
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Contract

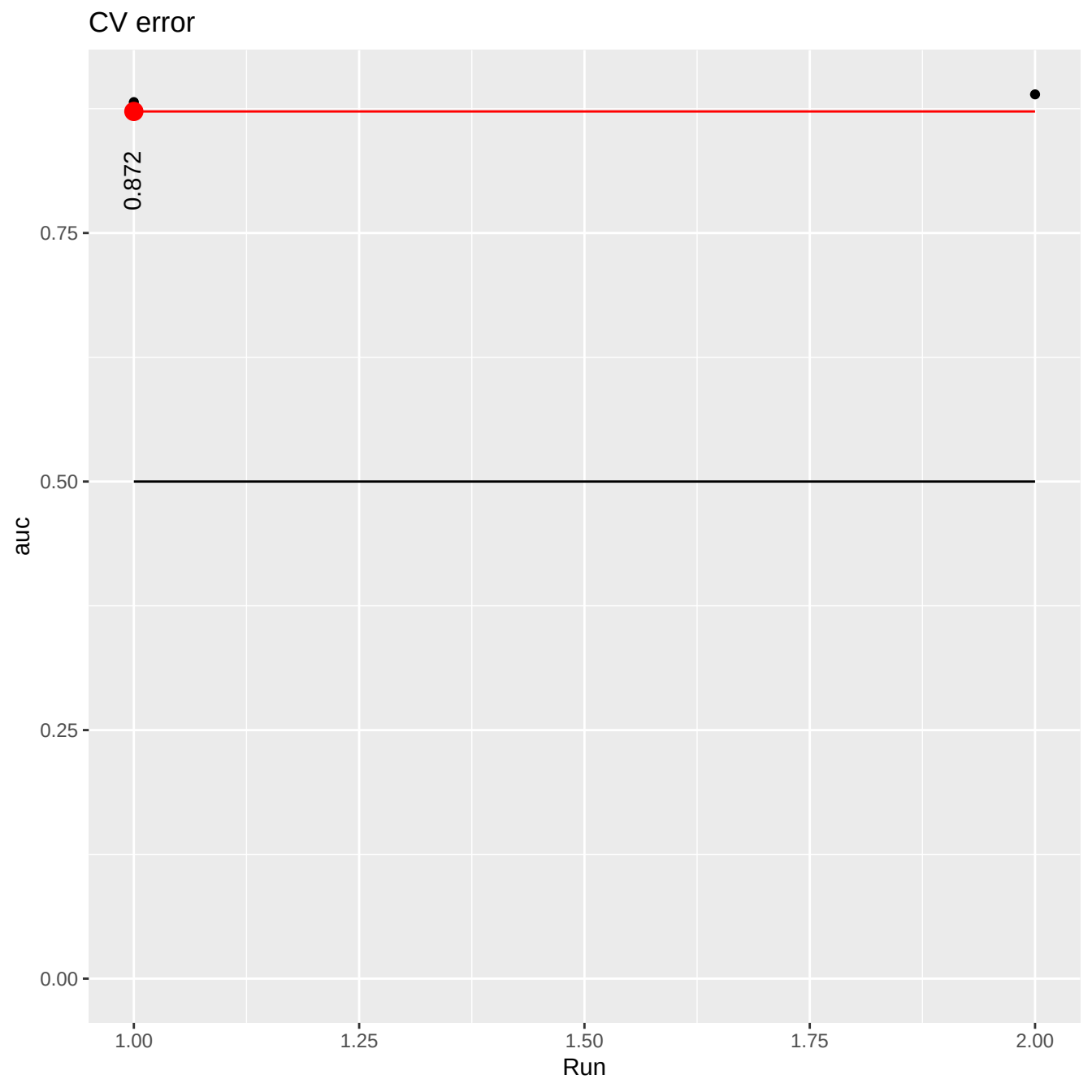

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
20
30
40
50
60
70
80
90
100
110
120
Count
0
50
100
150
MonthlyCharges

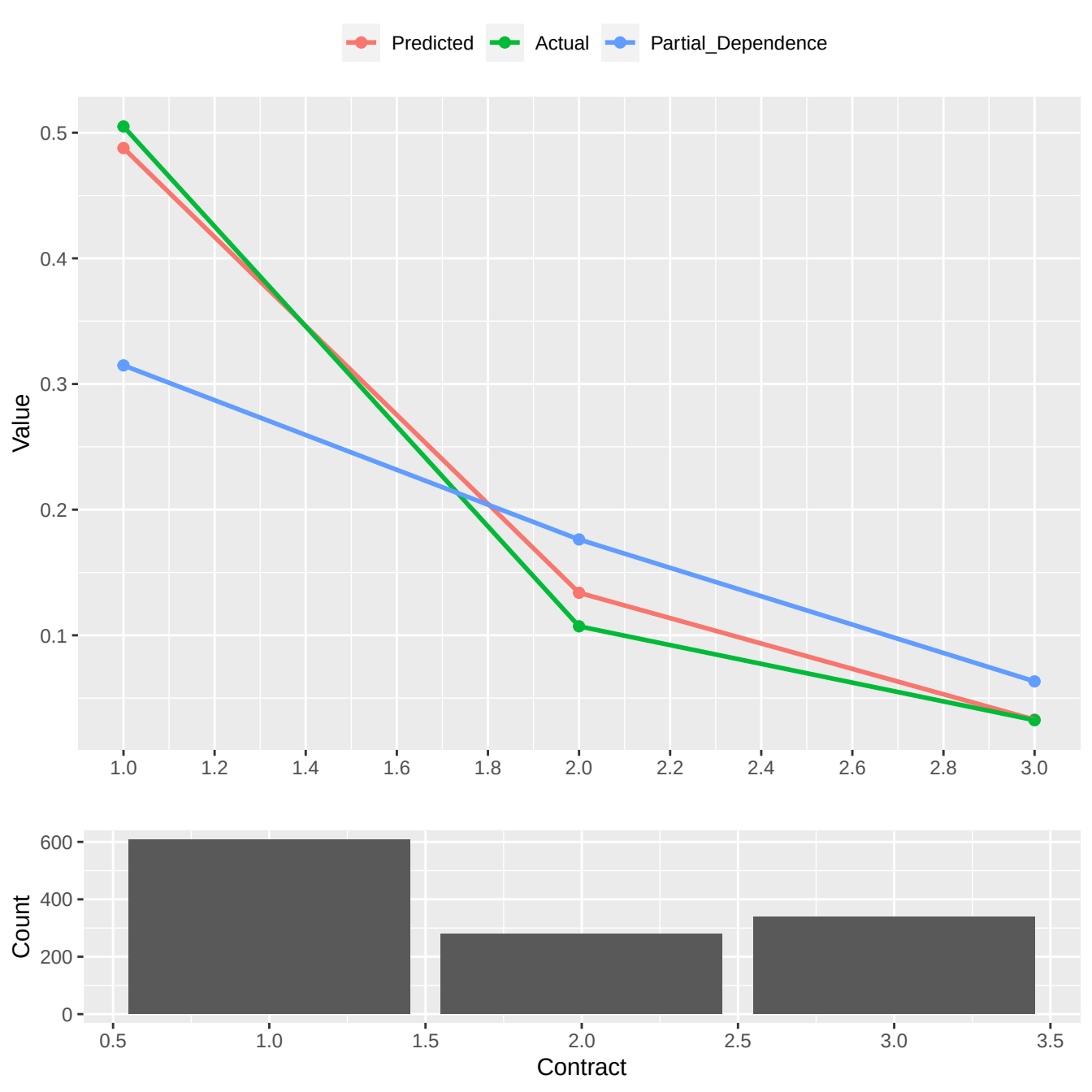

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
10
20
30
40
50
60
70
Count
0
50
100
150
tenure

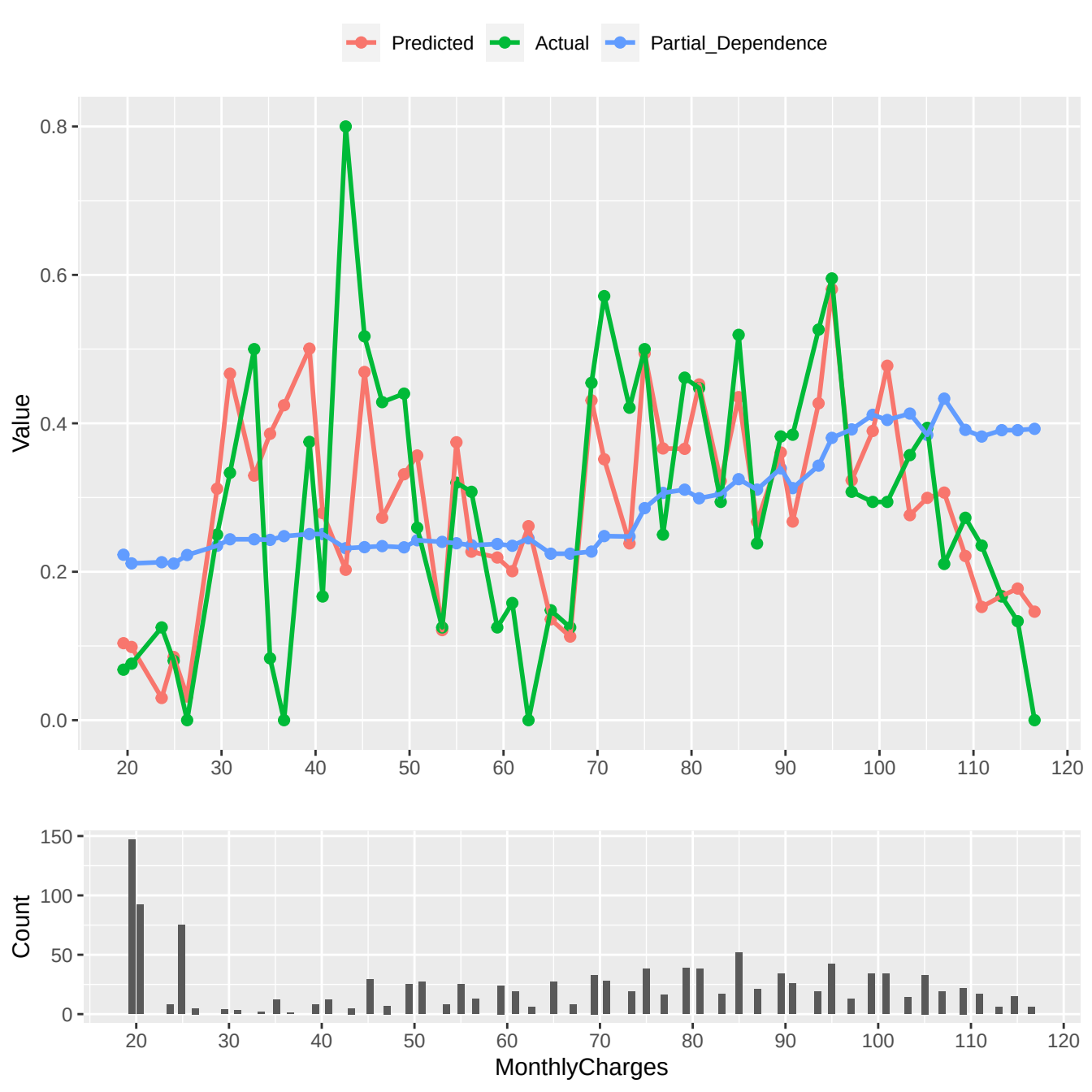

Predicted
Actual
Partial_Dependence
Value
0.5
0.4
0.3
0.2
0.1
0.0
0
1000
2000
3000
4000
5000
6000
7000
8000
Count
200
150
100
50
0
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
TotalCharges


Predicted
Actual
Partial_Dependence
Value
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
600
400
200
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
OnlineSecurity


CV error
0.872
0.75
0.50
0.25
0.00
auc
0.950
0.975
1.000
1.025
1.05
Run

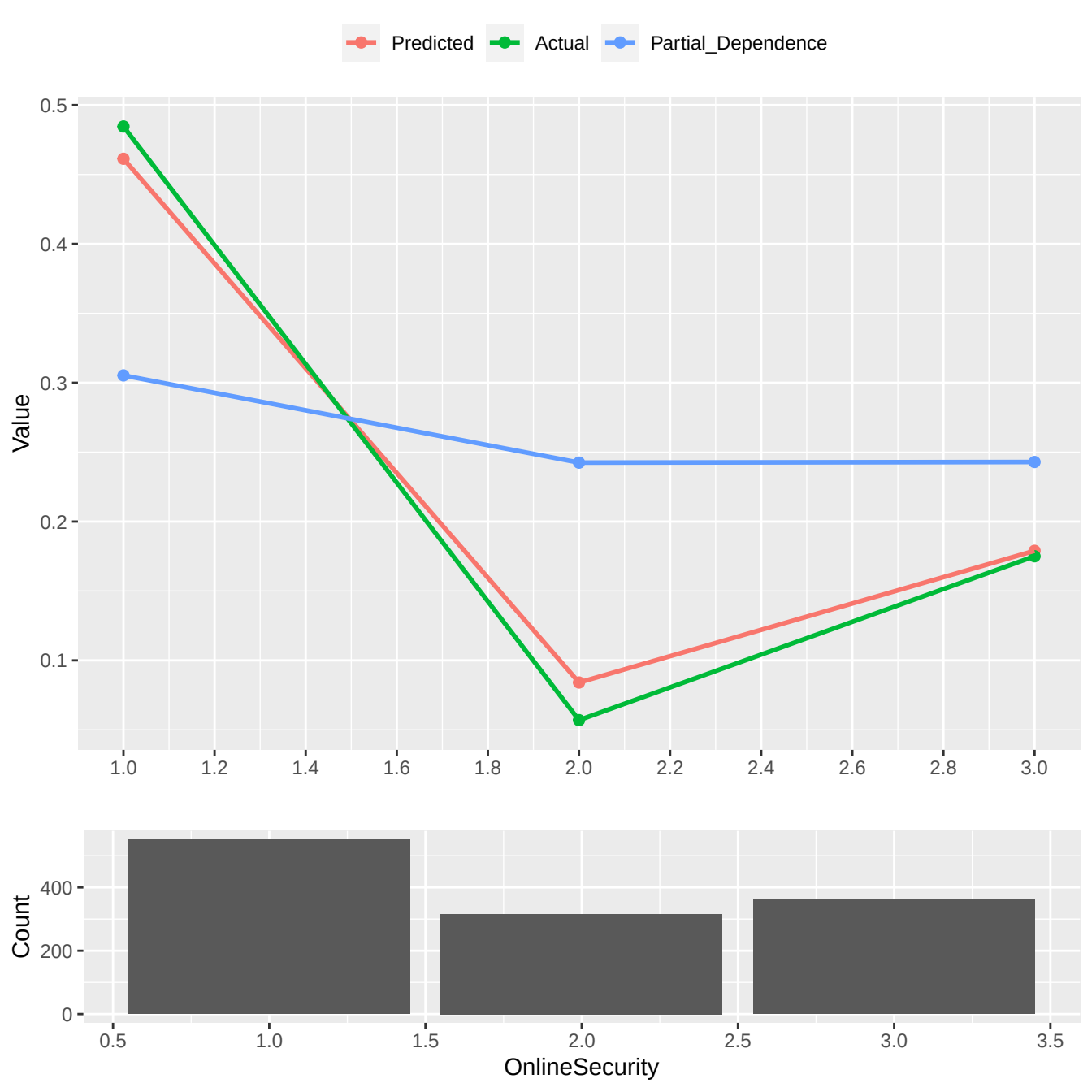

CV error
0.872
0.75
0.50
0.25
0.00
auc
1.00
1.25
1.50
1.75
2.00
Run

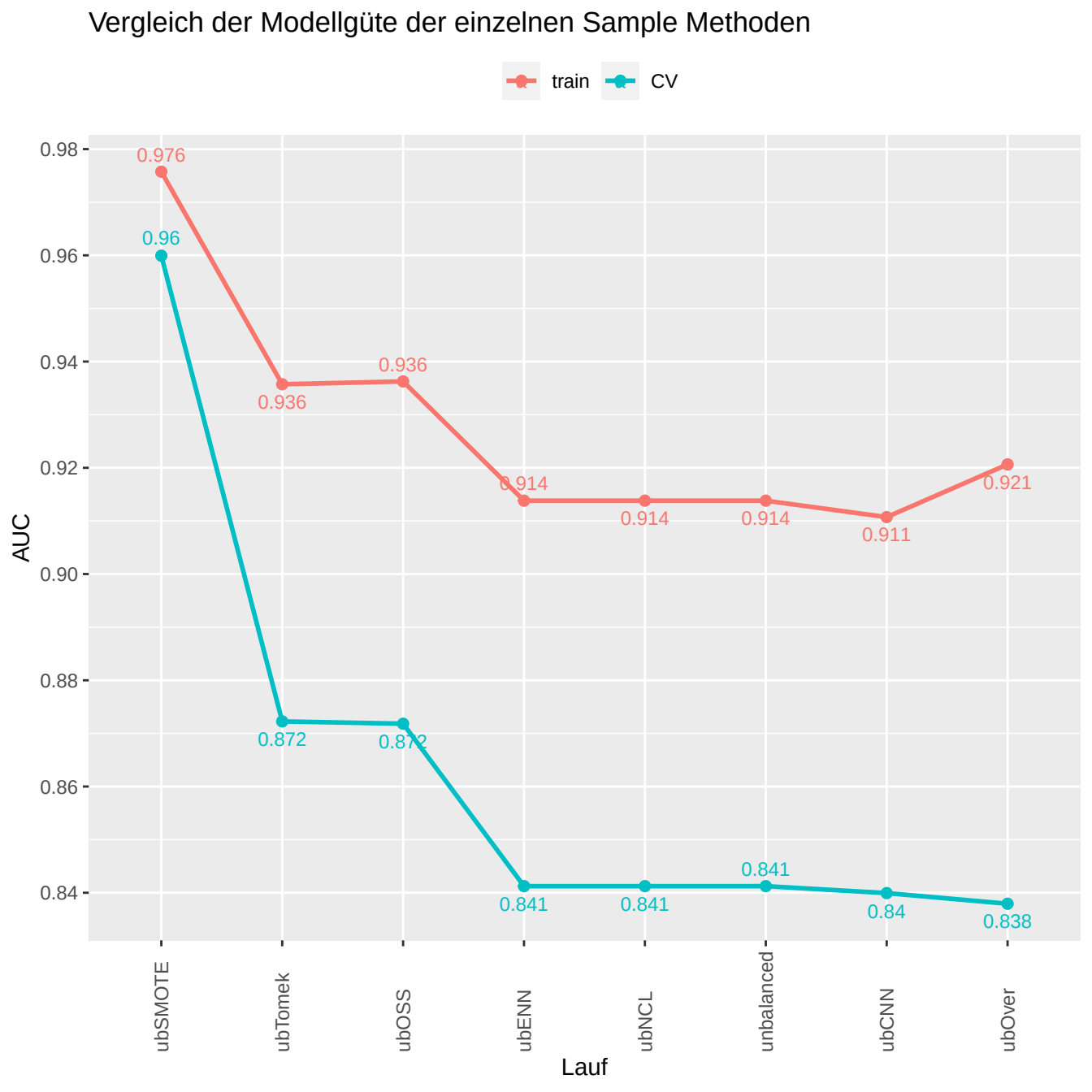

Predicted
Actual
Partial_Dependence
Value
0.5
0.4
0.3
0.2
0.1
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
600
400
200
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Contract

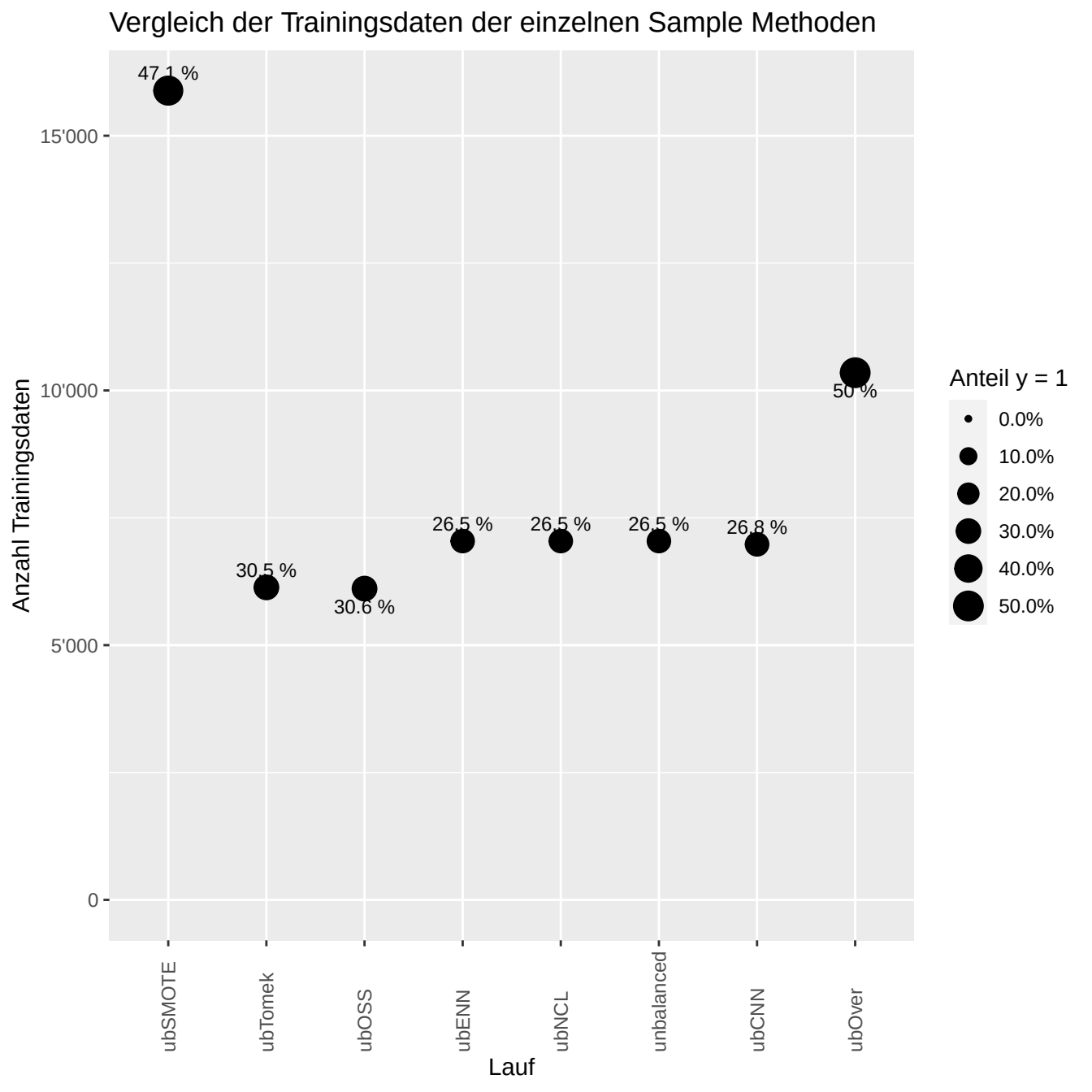

Predicted
Actual
Partial_Dependence
Value
0.0
0.2
0.4
0.6
0.8
20
30
40
50
60
70
80
90
100
110
120
Count
0
50
100
150
MonthlyCharges

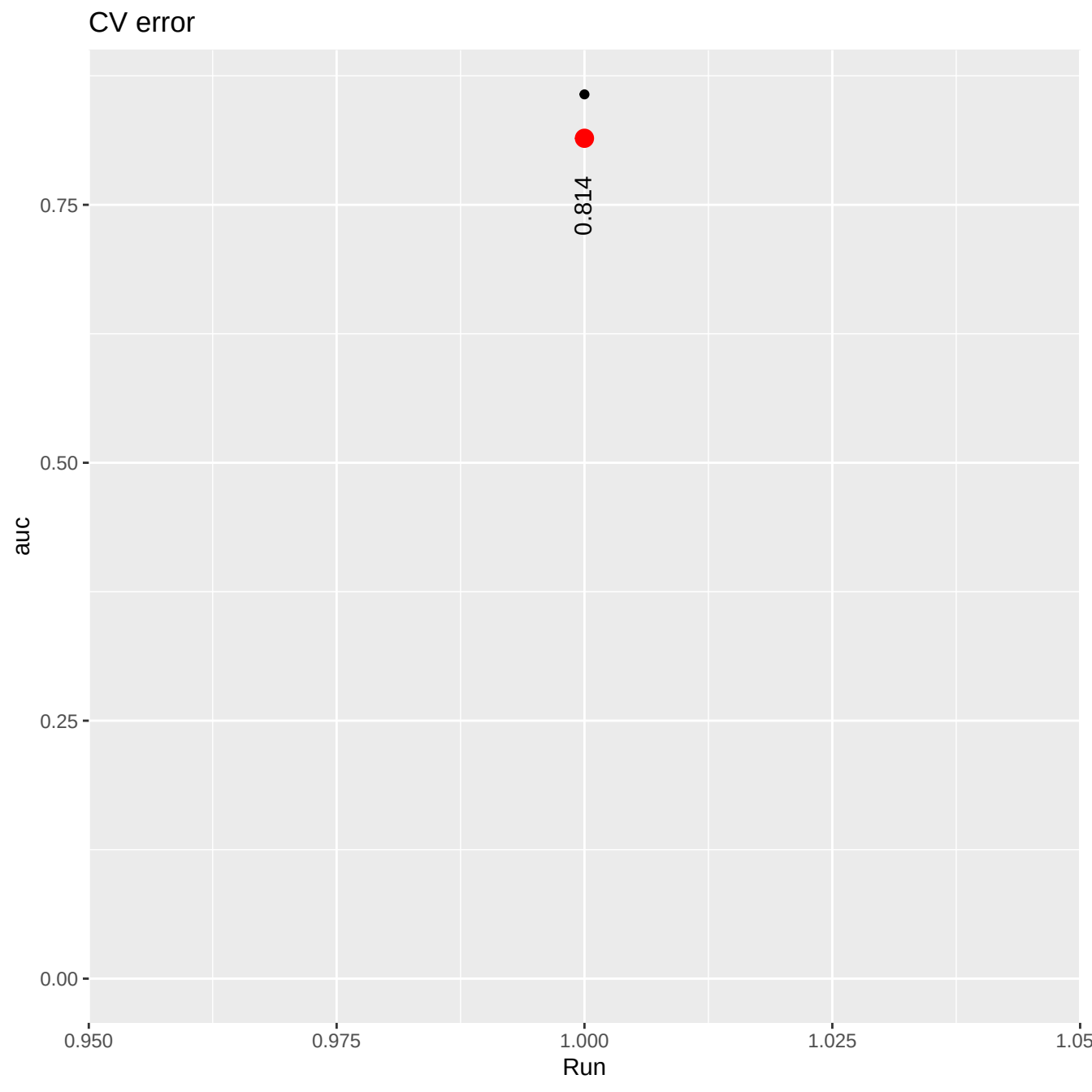

Predicted
Actual
Partial_Dependence
Value
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
10
20
30
40
50
60
70
Count
100
50
0
0
10
20
30
40
50
60
70
tenure

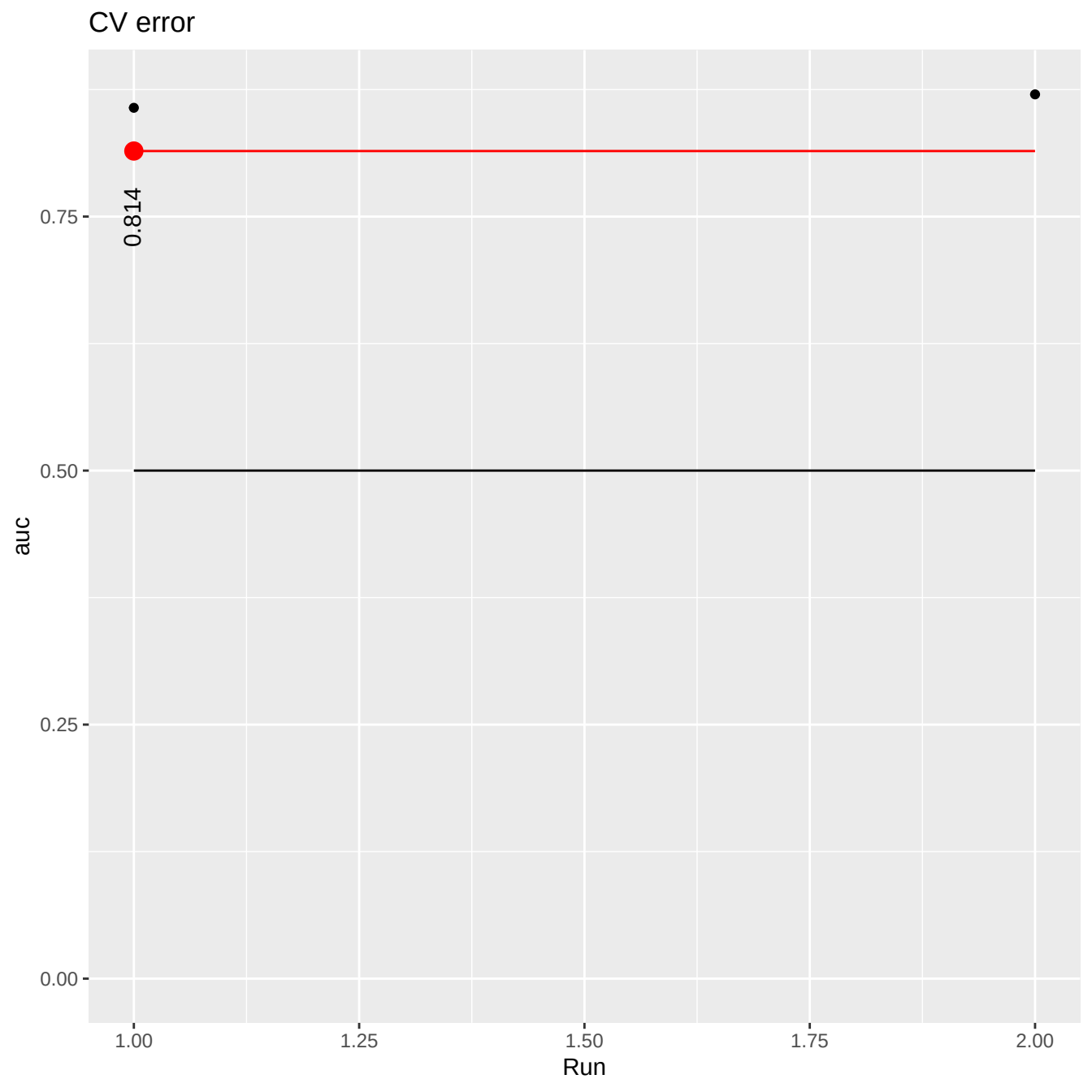

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
0
1000
2000
3000
4000
5000
6000
7000
8000
Count
0
50
100
150
200
9000
TotalCharges

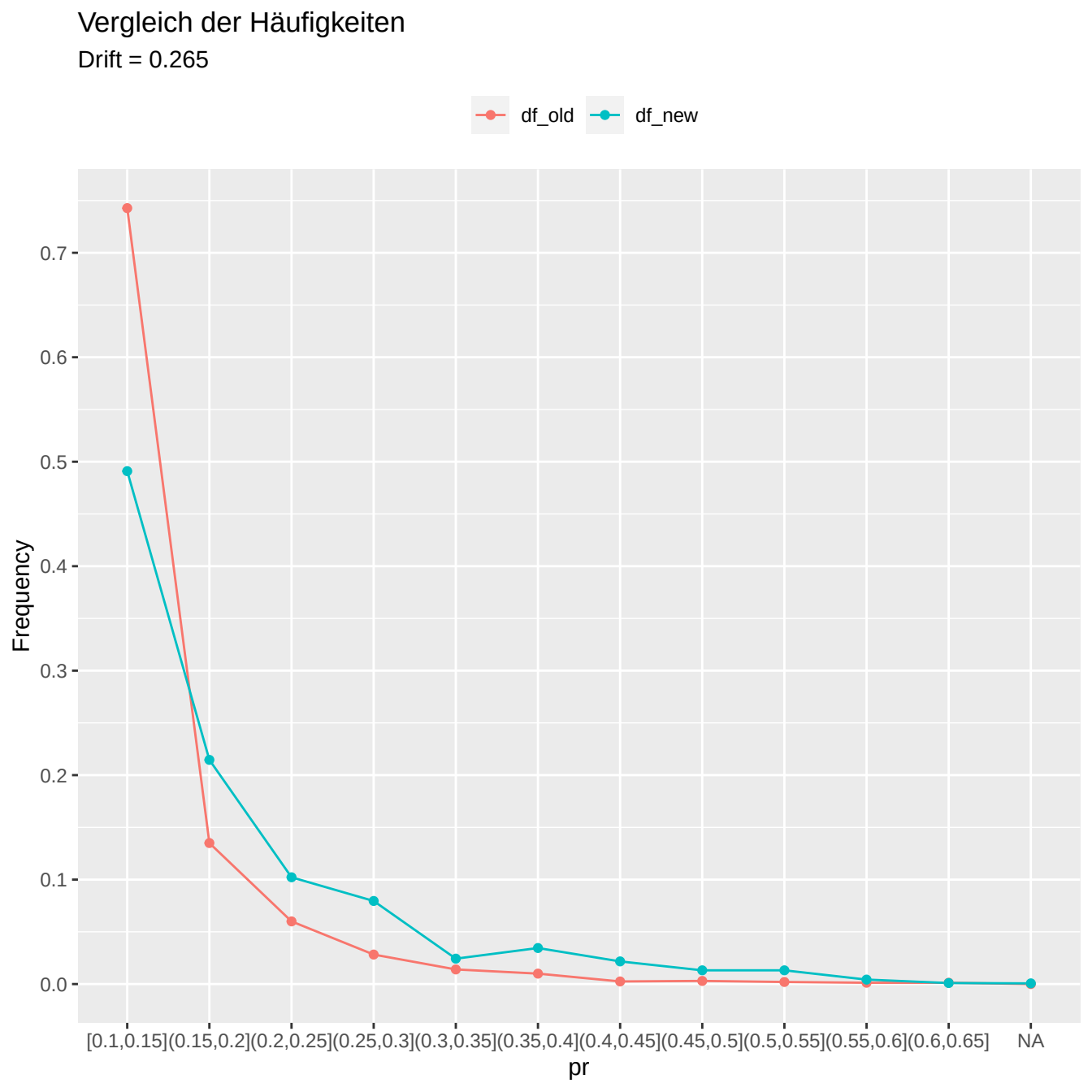

Predicted
Actual
Partial_Dependence
Value
0.5
0.4
0.3
0.2
0.1
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
400
200
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
OnlineSecurity

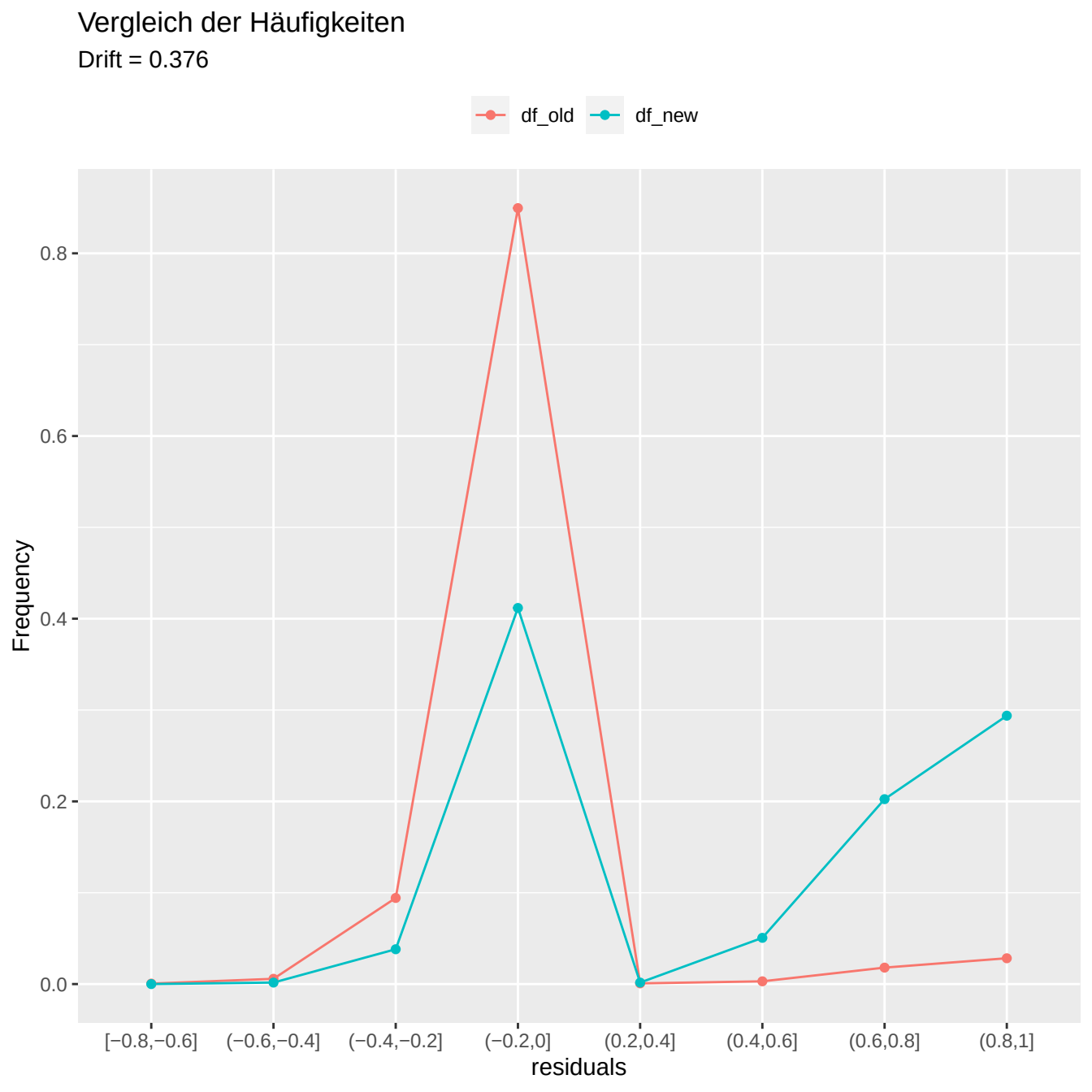

Vergleich der Modellgüte der einzelnen Sample Methoden
train
CV
AUC
Lauf
0.98
0.96
0.94
0.92
0.90
0.88
0.86
0.84
0.976
0.96
0.936
0.936
0.936
0.914
0.914
0.914
0.911
0.921
0.872
0.872
0.841
0.841
0.841
0.84
0.838
ubSMOTE
ubTomek
ubOSS
ubENN
ubNCL
unbalanced
ubCNN
ubOver

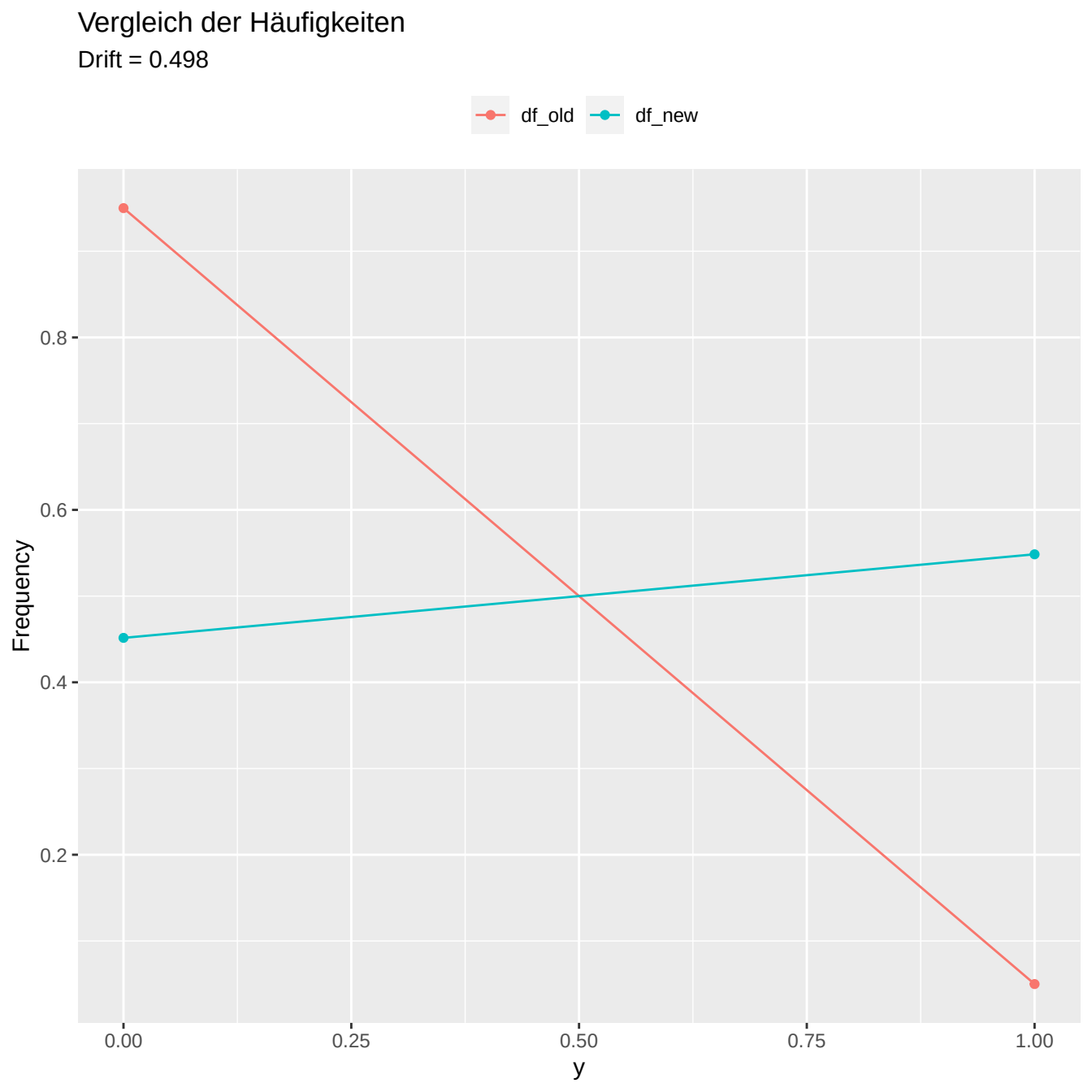

Vergleich der Trainingsdaten der einzelnen Sample Methoden
47.1 %
30.5 %
30.6 %
26.5 %
26.5 %
26.5 %
26.8 %
50 %
15'000
10'000
5'000
0
Anzahl Trainingsdaten
ubSMOTE
ubTomek
ubOSS
ubENN
ubNCL
unbalanced
ubCNN
ubOver
Lauf
Anteil y = 1
0.0%
10.0%
20.0%
30.0%
40.0%
50.0%

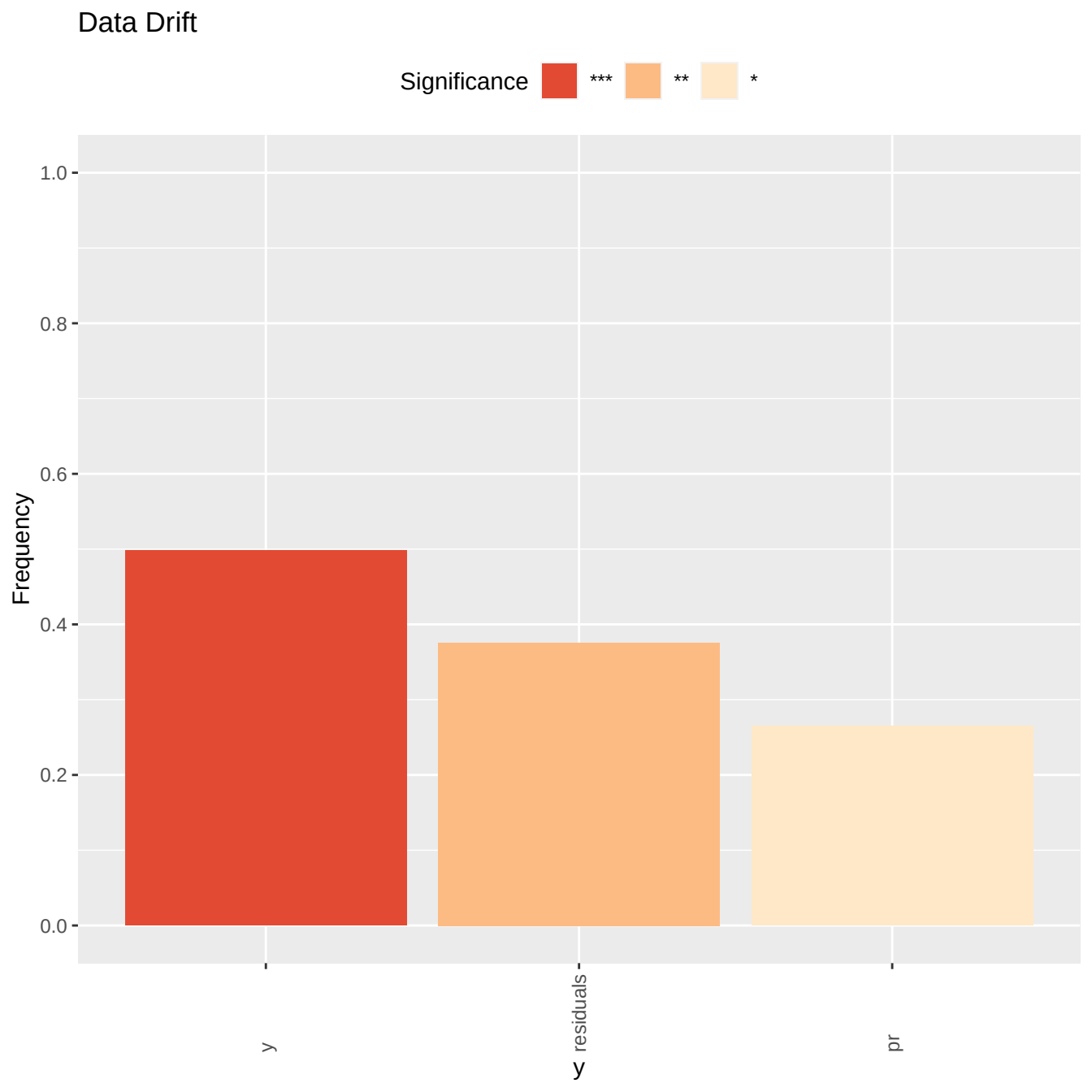

CV error
0.814
0.75
0.50
0.25
0.00
auc
0.950
0.975
1.000
1.025
1.05
Run

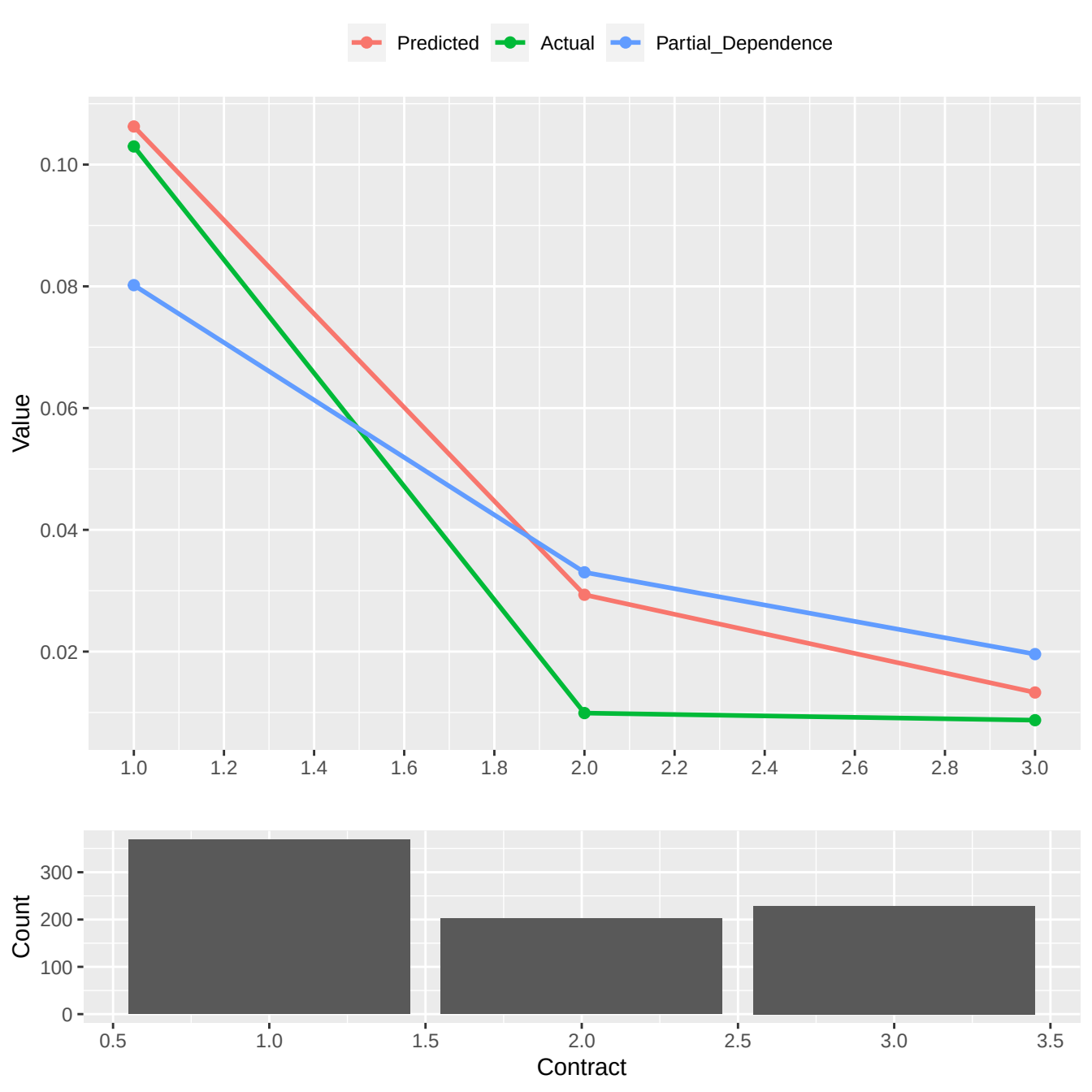

CV error
0.814
0.75
0.50
0.25
0.00
auc
1.00
1.25
1.50
1.75
2.00
Run

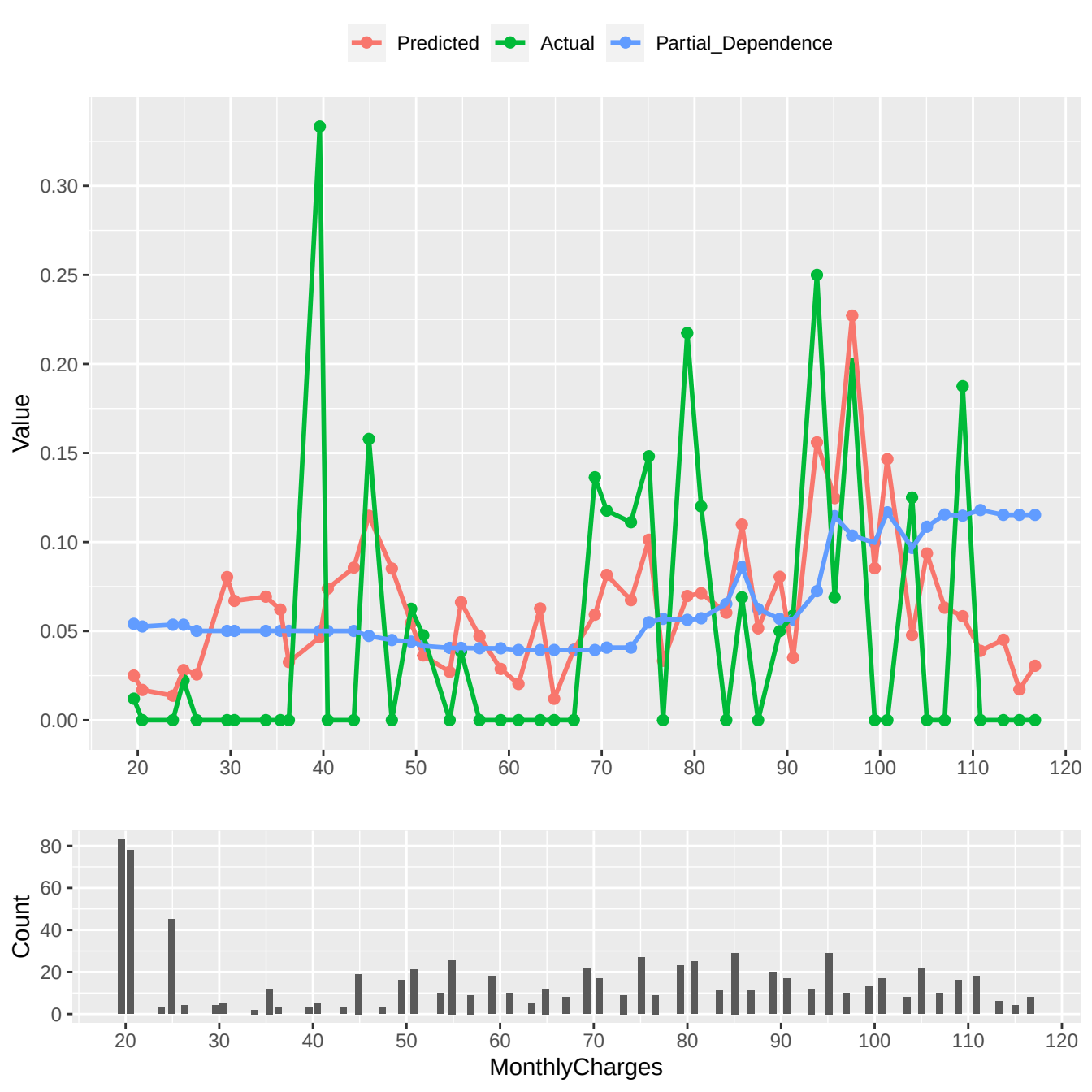

Vergleich der Häufigkeiten
Drift = 0.265
df_old
df_new
Frequency
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
[0.1,0.15]
(0.15,0.2]
(0.2,0.25]
(0.25,0.3]
(0.3,0.35]
(0.35,0.4]
(0.4,0.45]
(0.45,0.5]
(0.5,0.55]
(0.55,0.6]
(0.6,0.65]
NA
pr

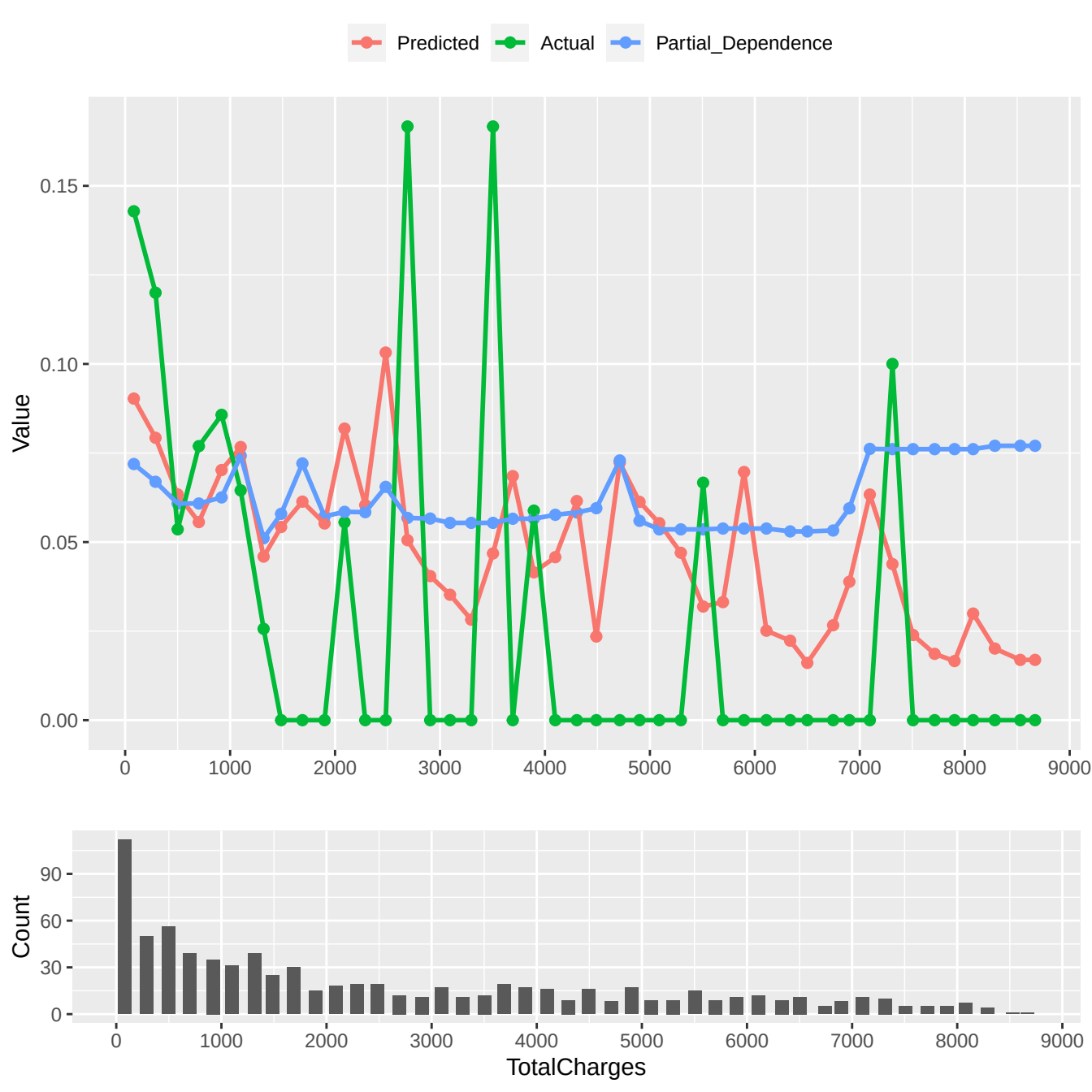

Vergleich der Häufigkeiten
Drift = 0.376
df_old
df_new
0.8
0.6
0.4
0.2
0.0
Frequency
[−0.8,−0.6]
(−0.6,−0.4]
(−0.4,−0.2]
(−0.2,0]
(0.2,0.4]
(0.4,0.6]
(0.6,0.8]
(0.8,1]
residuals


Vergleich der Häufigkeiten
Drift = 0.498
df_old
df_new
Frequency
0.8
0.6
0.4
0.2
0.00
0.25
0.50
0.75
1.00
y

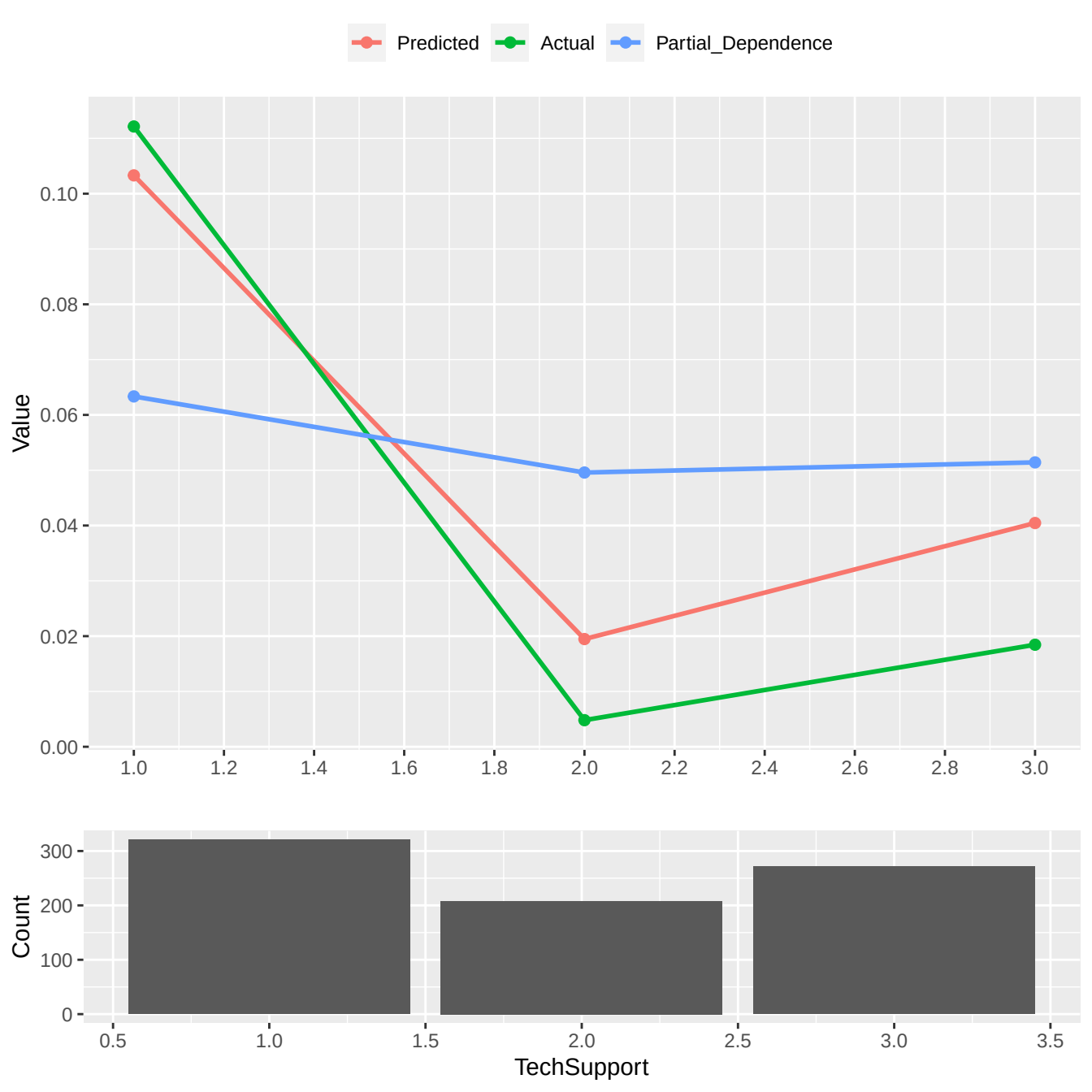

Data Drift
Significance
***
**
*
1.0
0.8
0.6
0.4
0.2
0.0
Frequency
y
residuals
pr
y

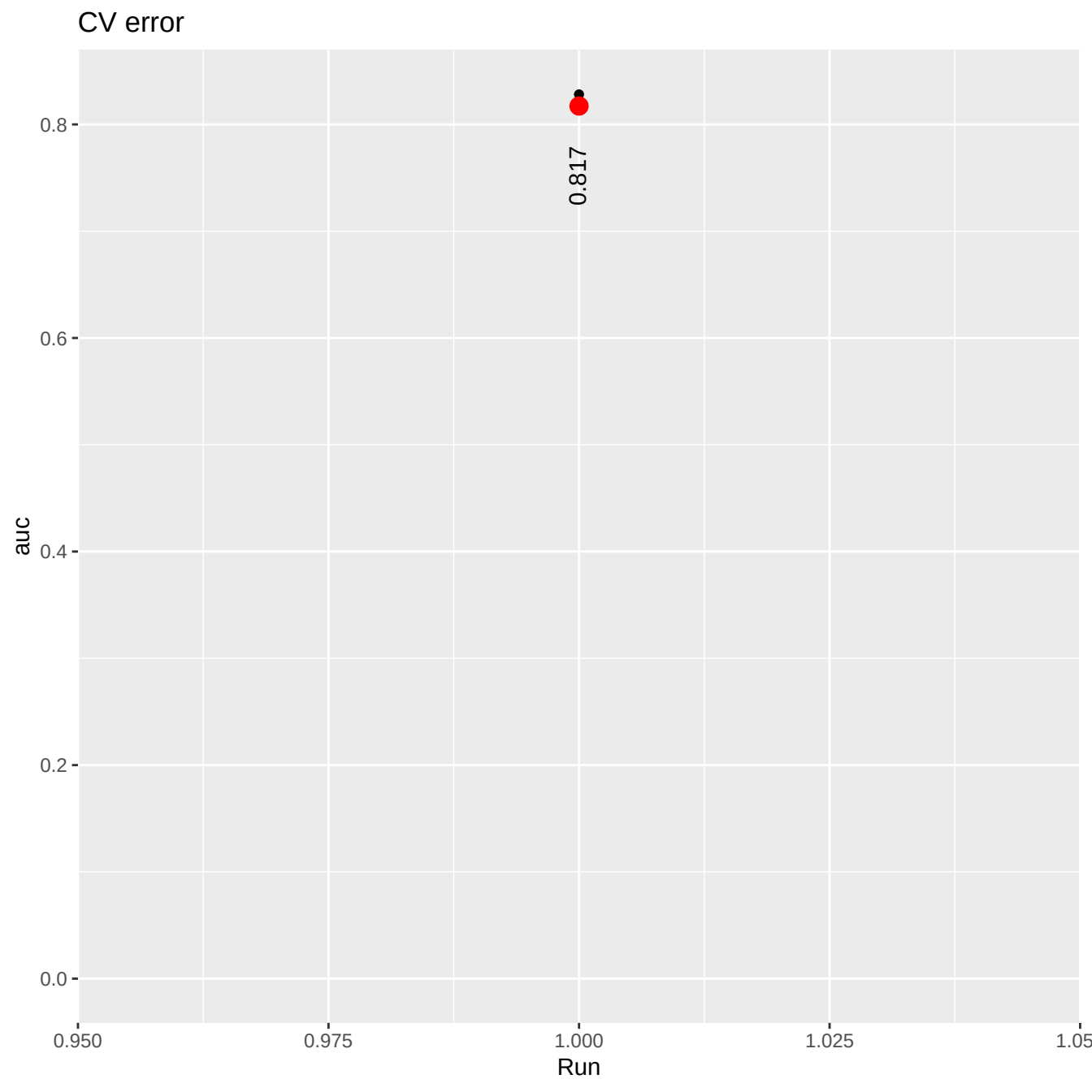

Predicted
Actual
Partial_Dependence
Value
0.10
0.08
0.06
0.04
0.02
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
300
200
100
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Contract

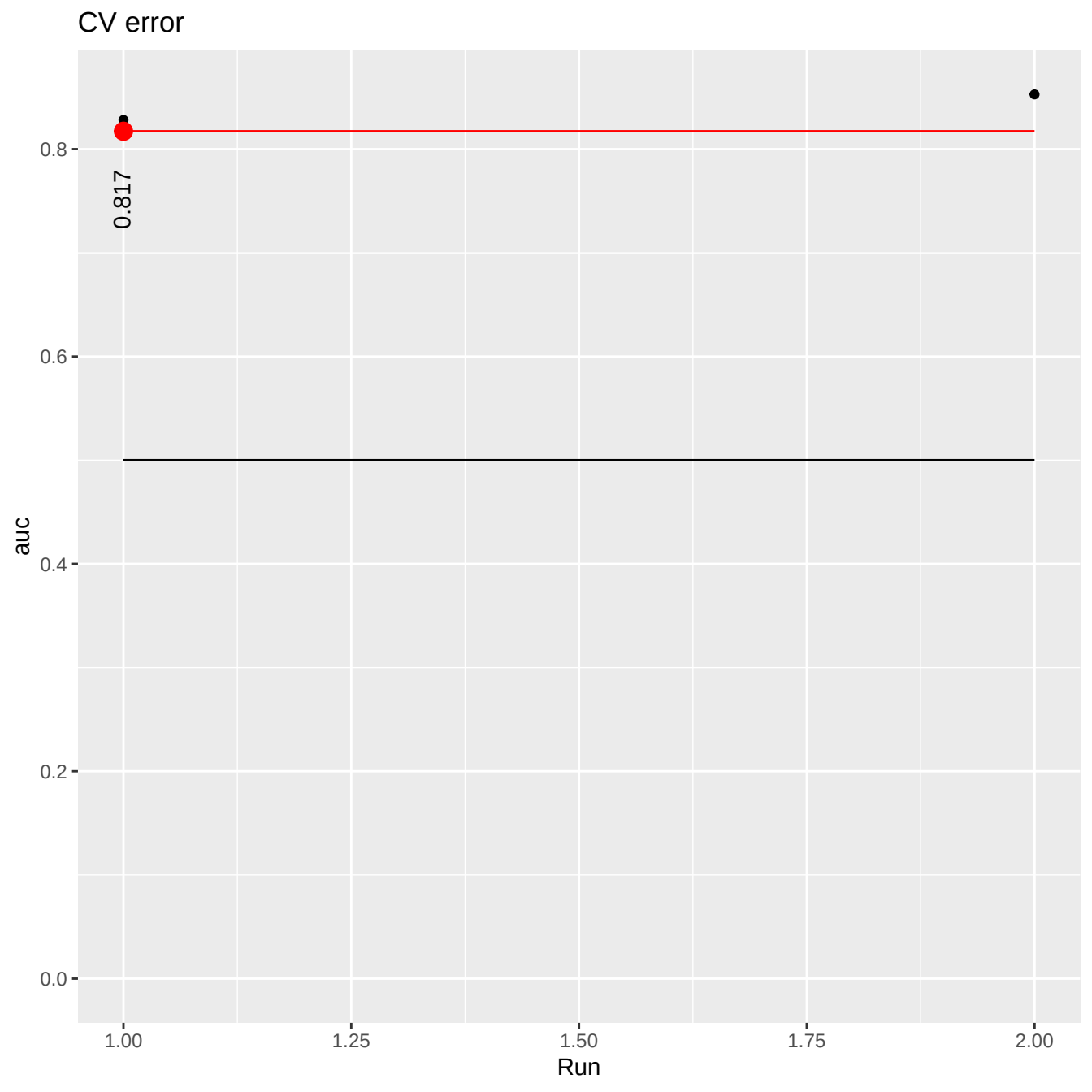

Predicted
Actual
Partial_Dependence
Value
0.30
0.25
0.20
0.15
0.10
0.05
0.00
20
30
40
50
60
70
80
90
100
110
120
Count
80
60
40
20
0
MonthlyCharges

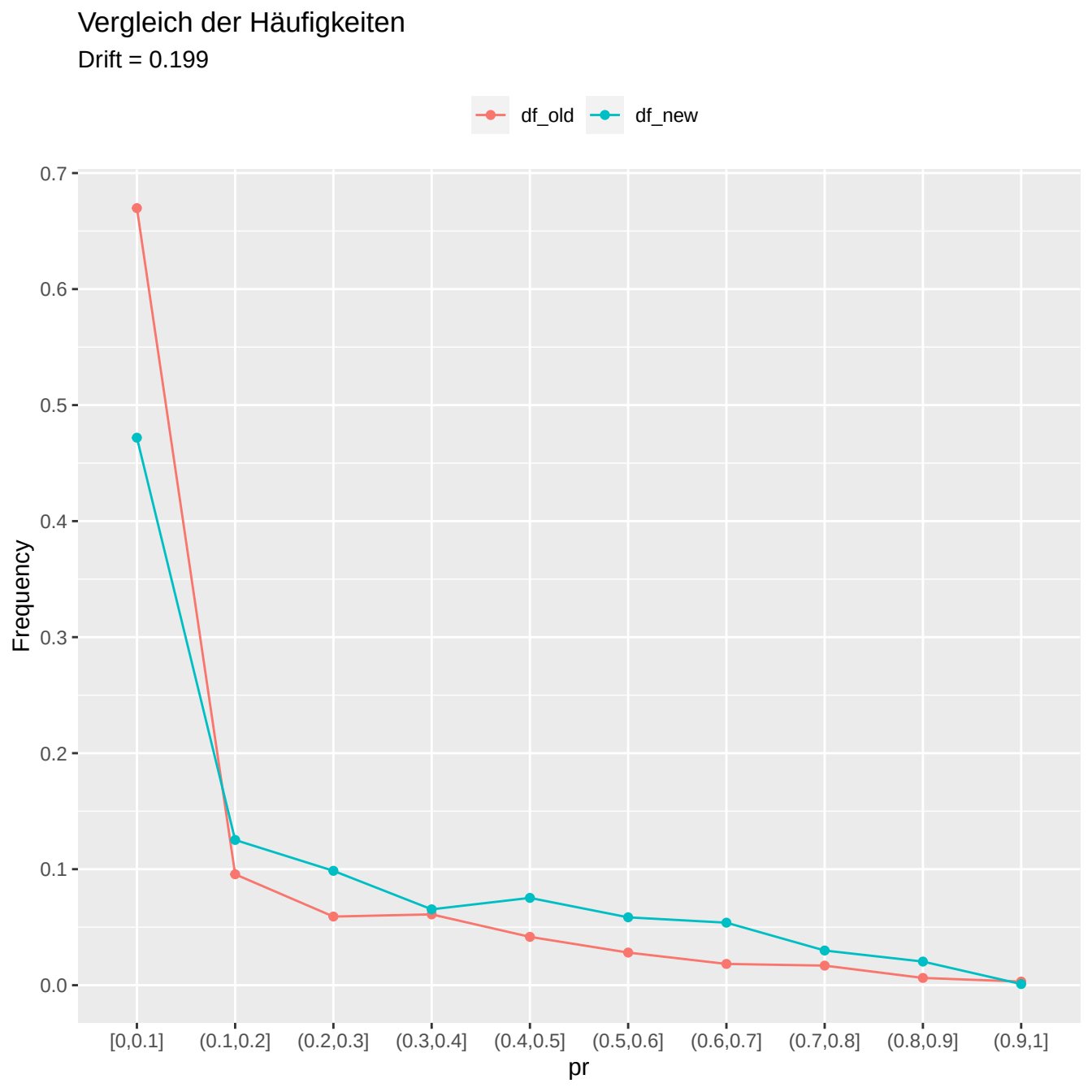

Predicted
Actual
Partial_Dependence
Value
0.15
0.10
0.05
0.00
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
Count
90
60
30
0
TotalCharges

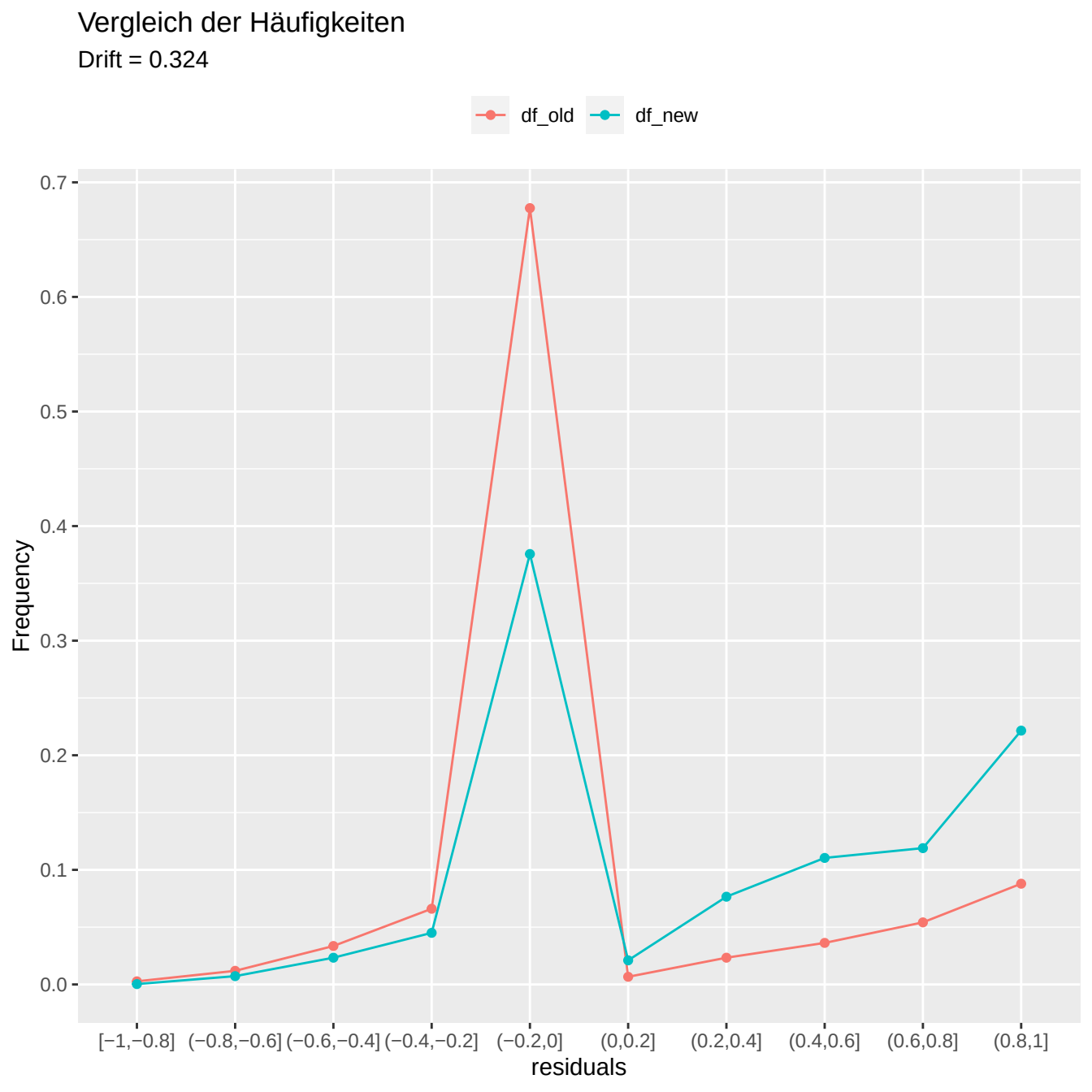

Predicted
Actual
Partial_Dependence
Value
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
50
60
70
Count
80
60
40
20
0
tenure

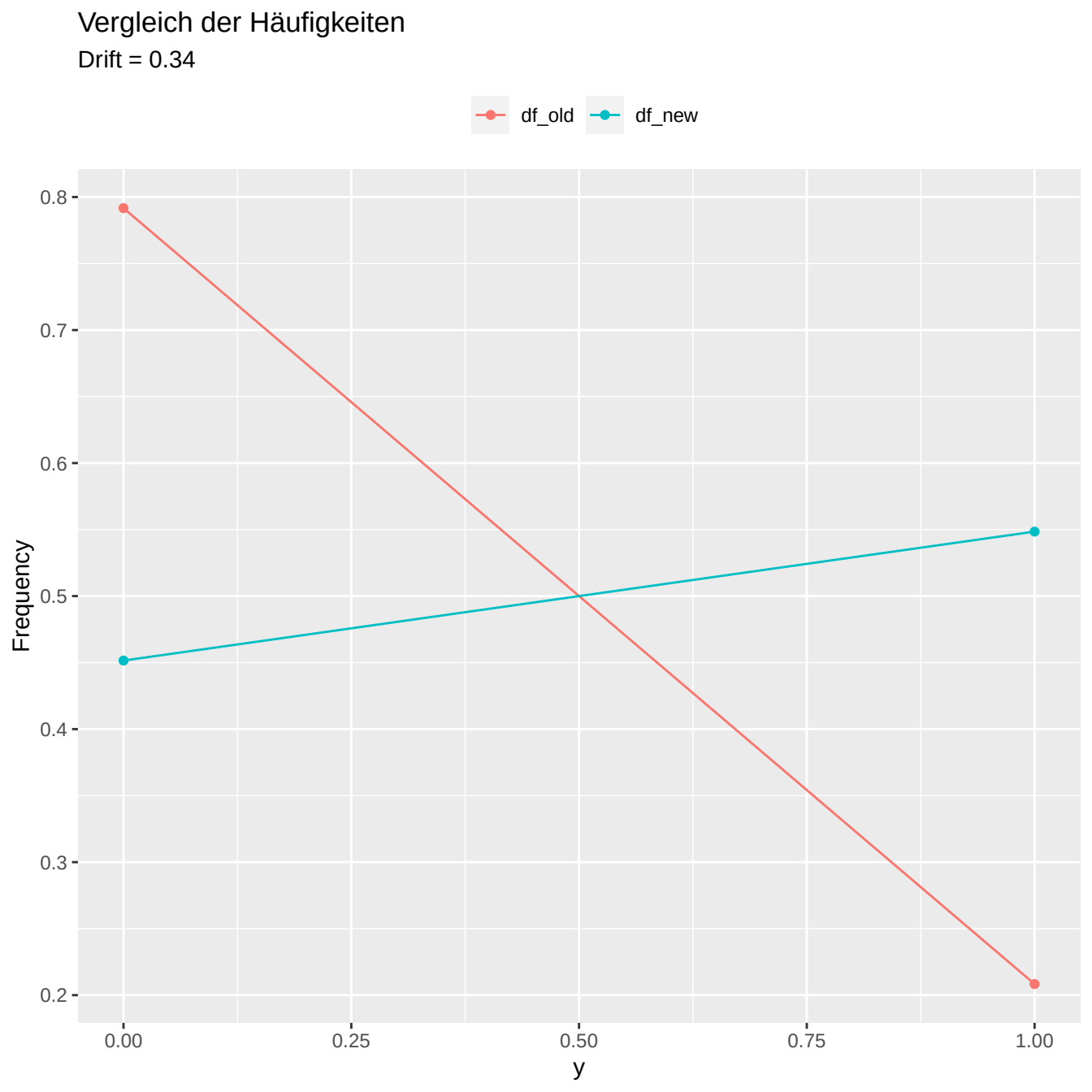

Predicted
Actual
Partial_Dependence
Value
0.10
0.08
0.06
0.04
0.02
0.00
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
300
200
100
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
TechSupport

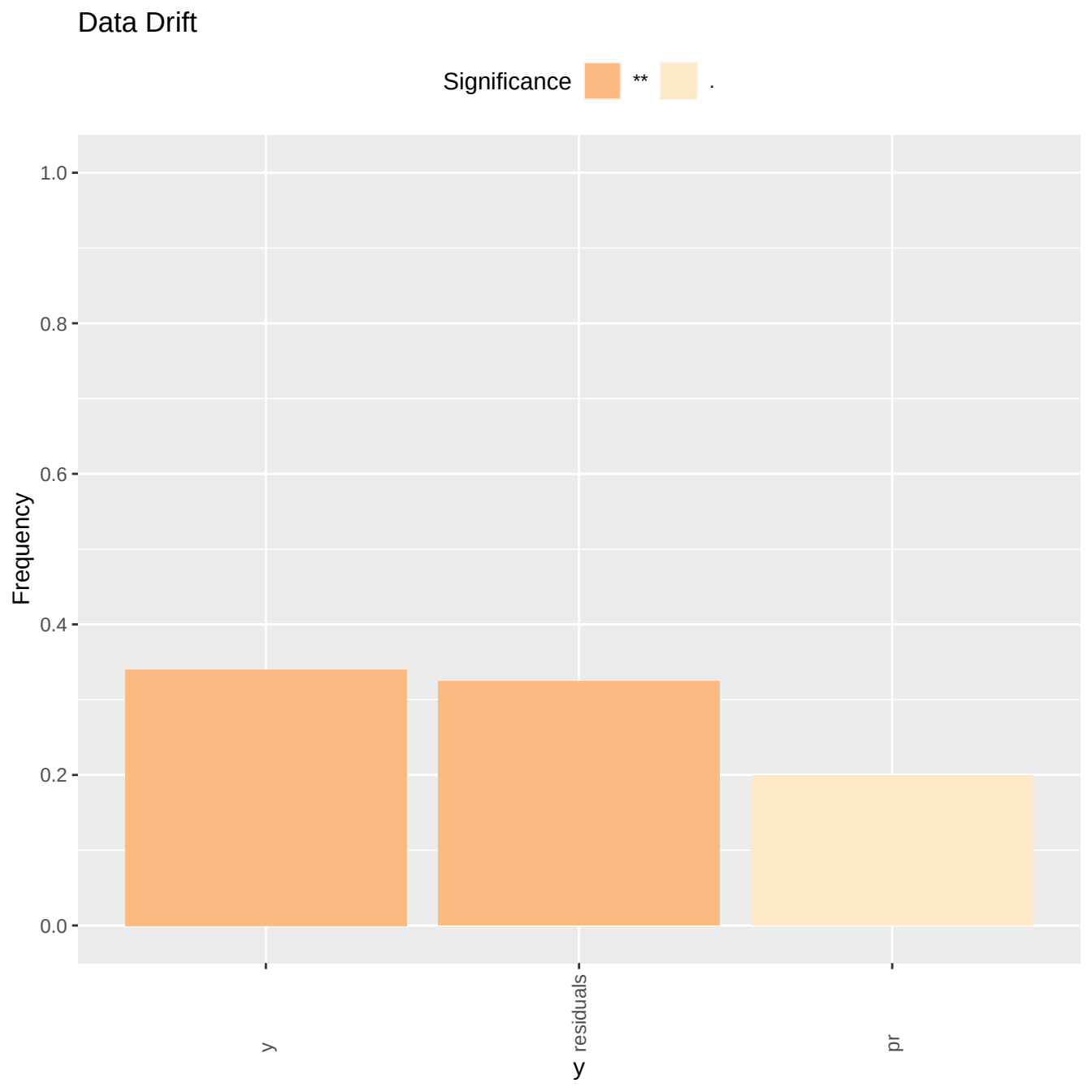

CV error
0.817
auc
0.0
0.2
0.4
0.6
0.8
0.950
0.975
1.000
1.025
1.05
Run


CV error
0.817
0.8
0.6
0.4
0.2
0.0
auc
1.00
1.25
1.50
1.75
2.00
Run

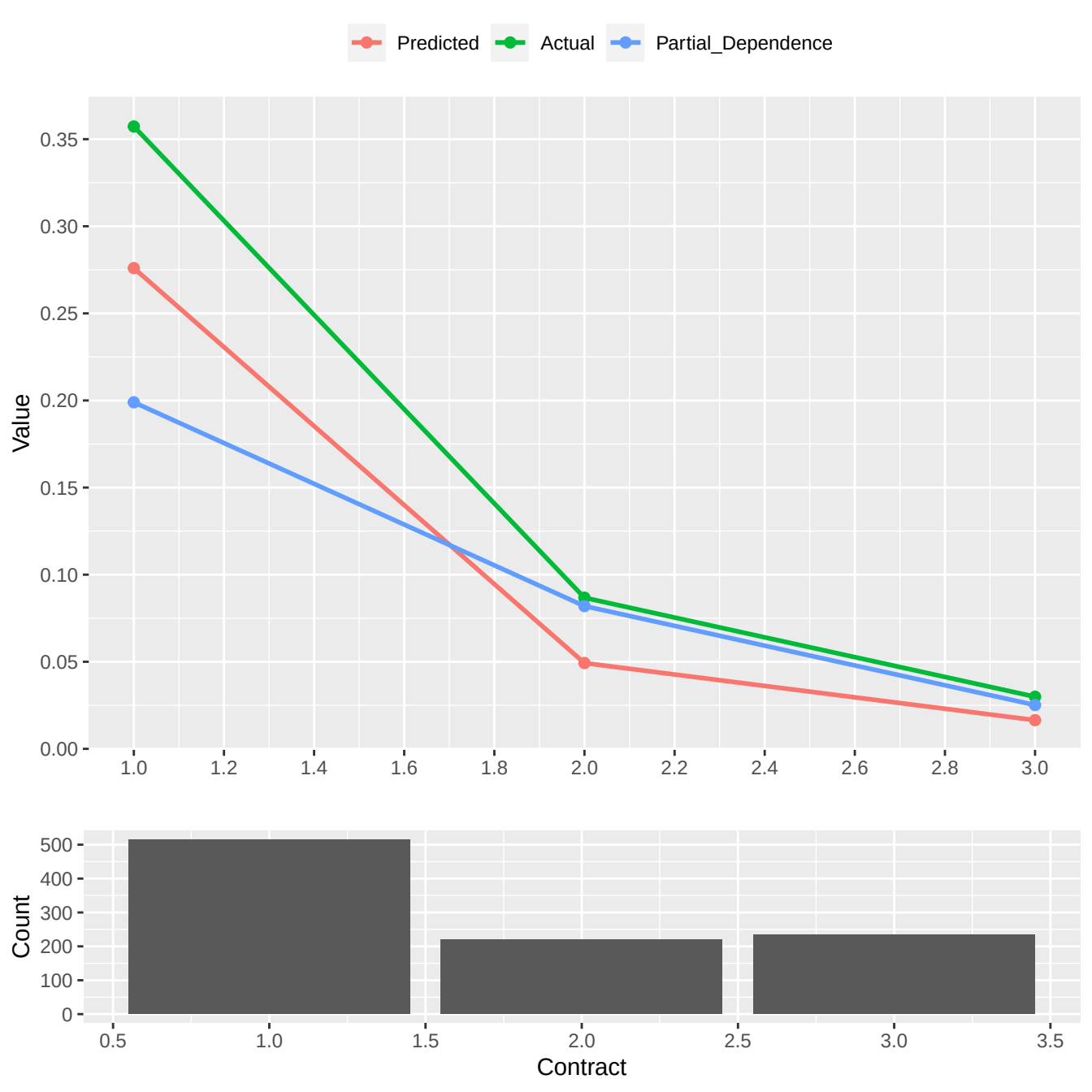

Vergleich der Häufigkeiten
Drift = 0.199
df_old
df_new
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Frequency
[0,0.1]
(0.1,0.2]
(0.2,0.3]
(0.3,0.4]
(0.4,0.5]
(0.5,0.6]
(0.6,0.7]
(0.7,0.8]
(0.8,0.9]
(0.9,1]
pr

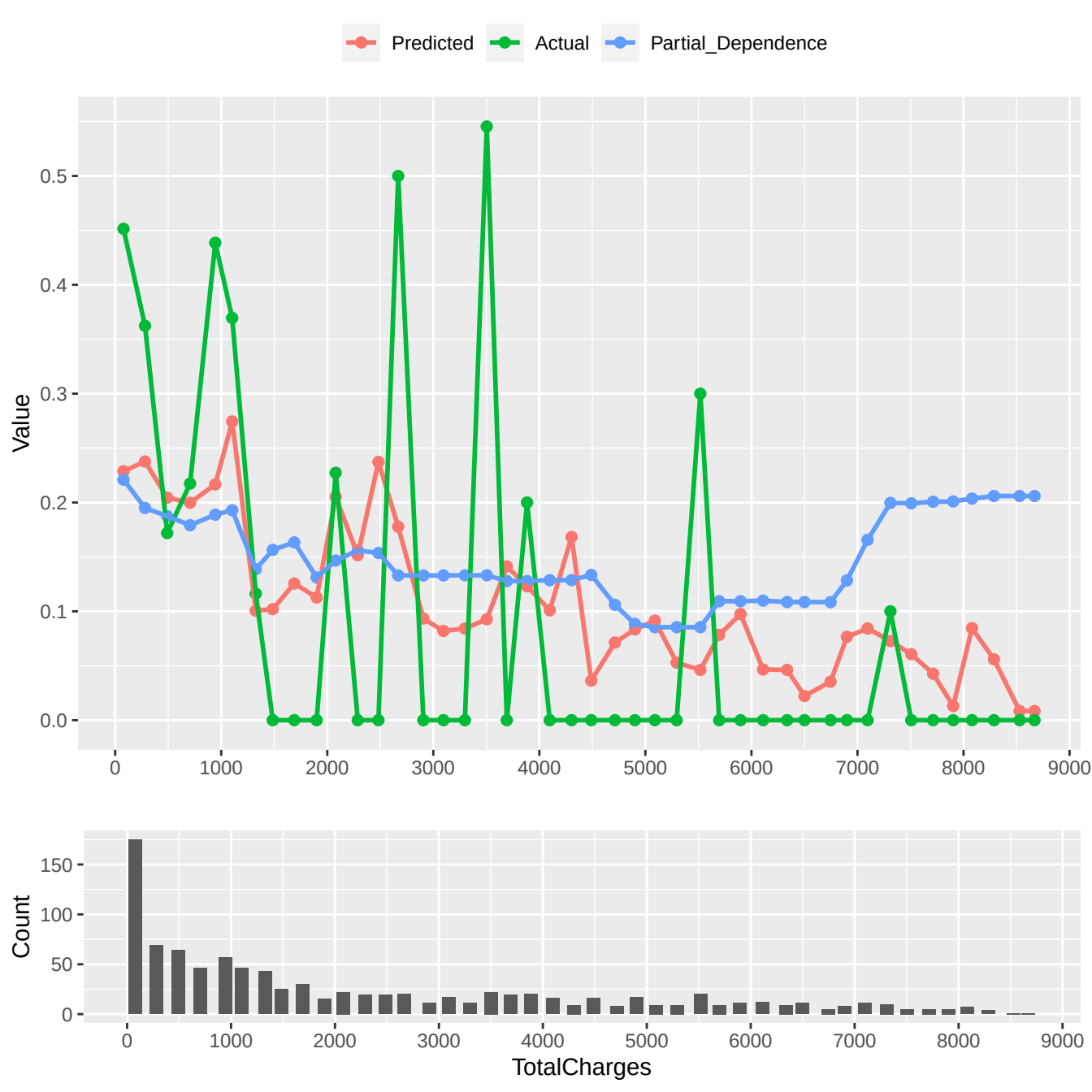

Vergleich der Häufigkeiten
Drift = 0.324
df_old
df_new
Frequency
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
[−1,−0.8]
(−0.8,−0.6]
(−0.6,−0.4]
(−0.4,−0.2]
(−0.2,0]
(0,0.2]
(0.2,0.4]
(0.4,0.6]
(0.6,0.8]
(0.8,1]
residuals


Vergleich der Häufigkeiten
Drift = 0.34
df_old
df_new
0.8
0.7
0.6
0.5
0.4
0.3
0.2
Frequency
0.00
0.25
0.50
0.75
1.00
y

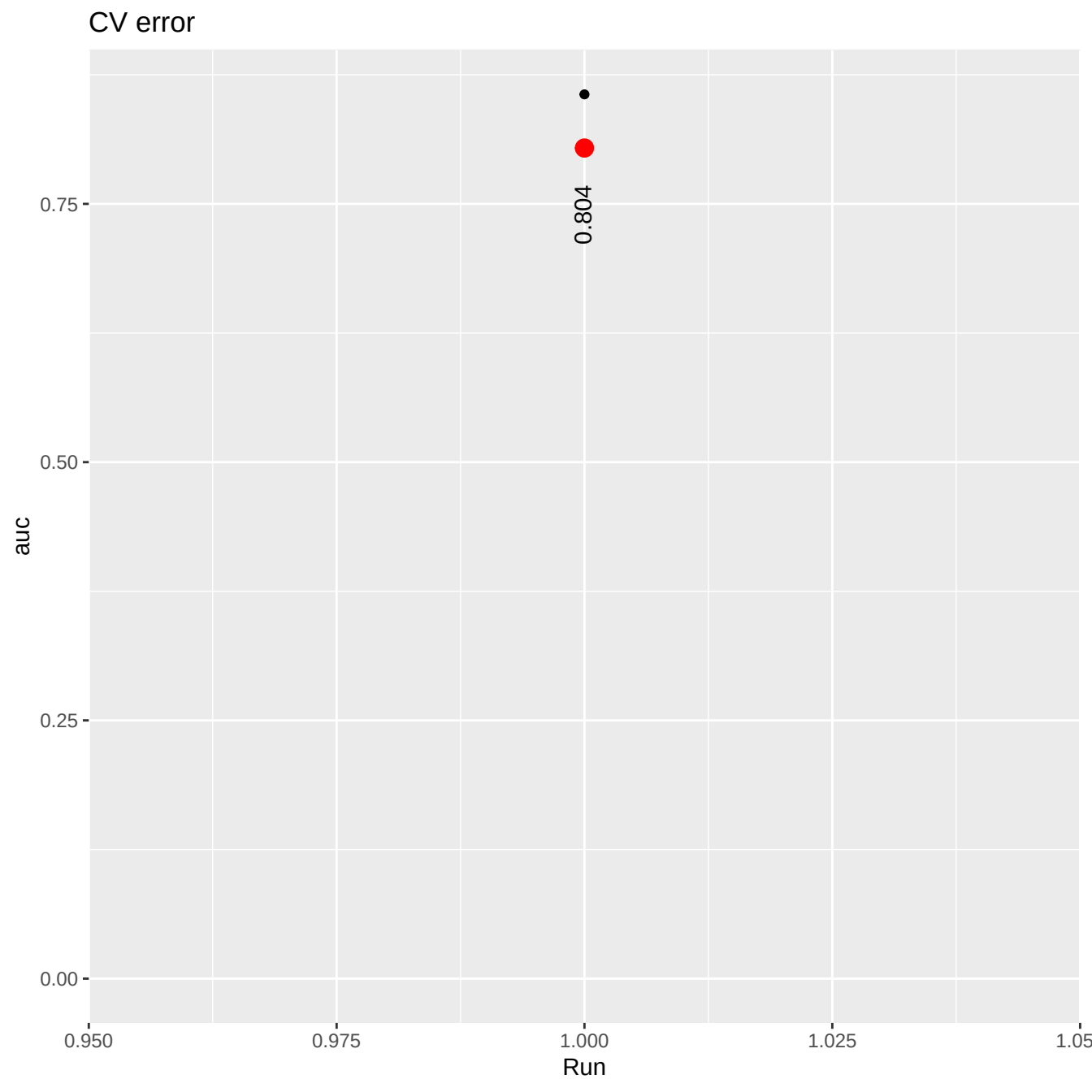

Data Drift
Significance
**
.
1.0
0.8
0.6
0.4
0.2
0.0
Frequency
y
residuals
pr
y

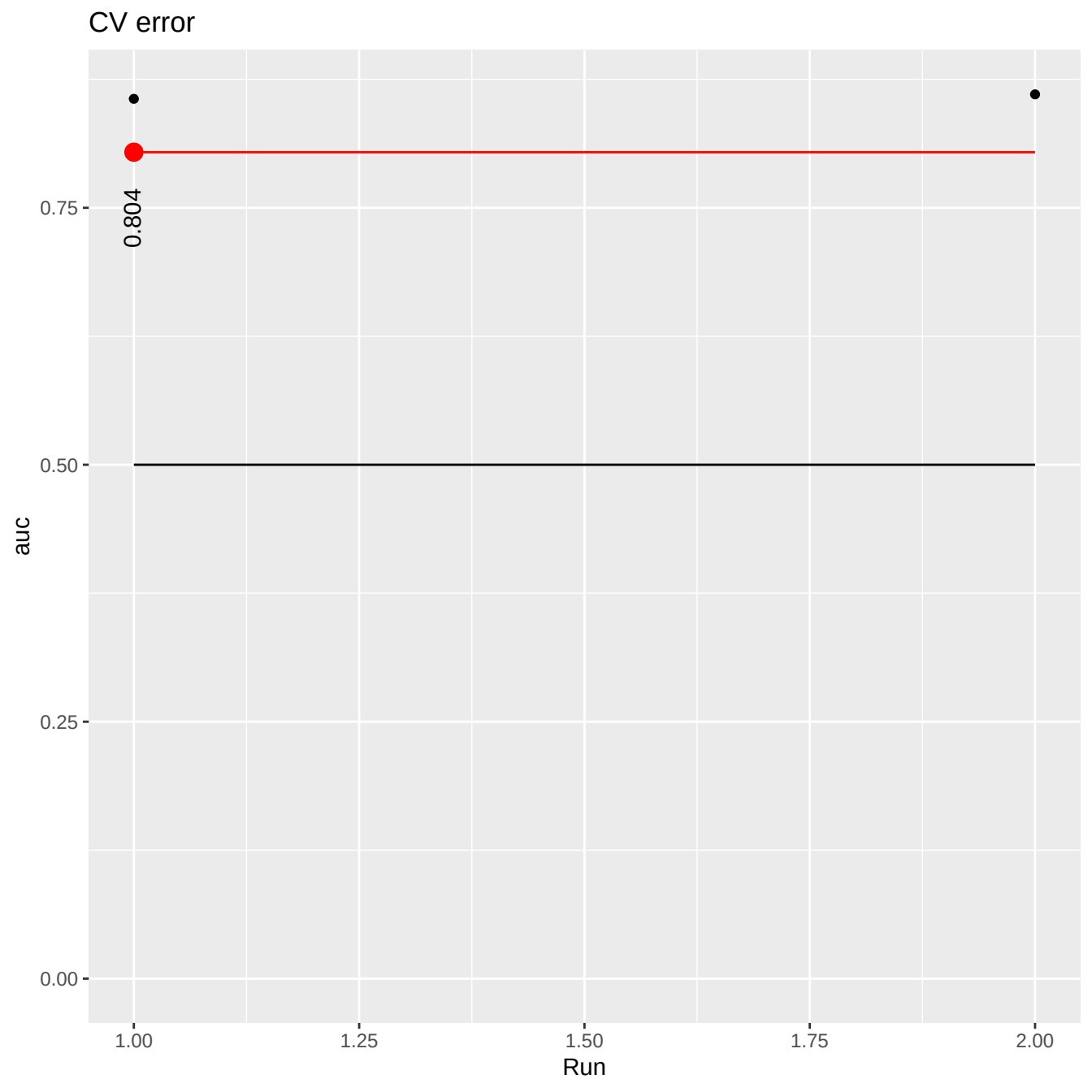

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
0.6
20
30
40
50
60
70
80
90
100
110
120
Count
0
25
50
75
MonthlyCharges

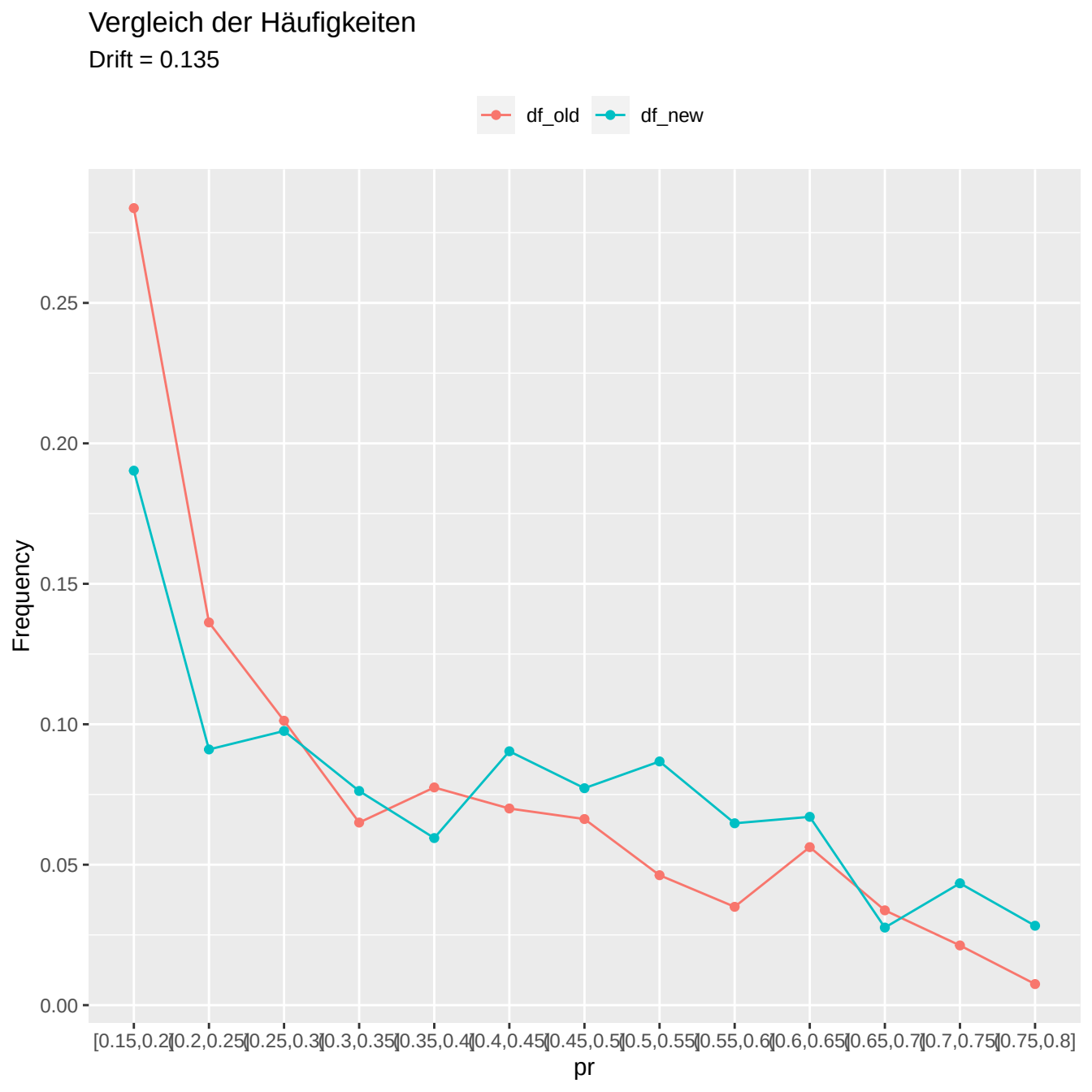

Predicted
Actual
Partial_Dependence
Value
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
500
400
300
200
100
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Contract

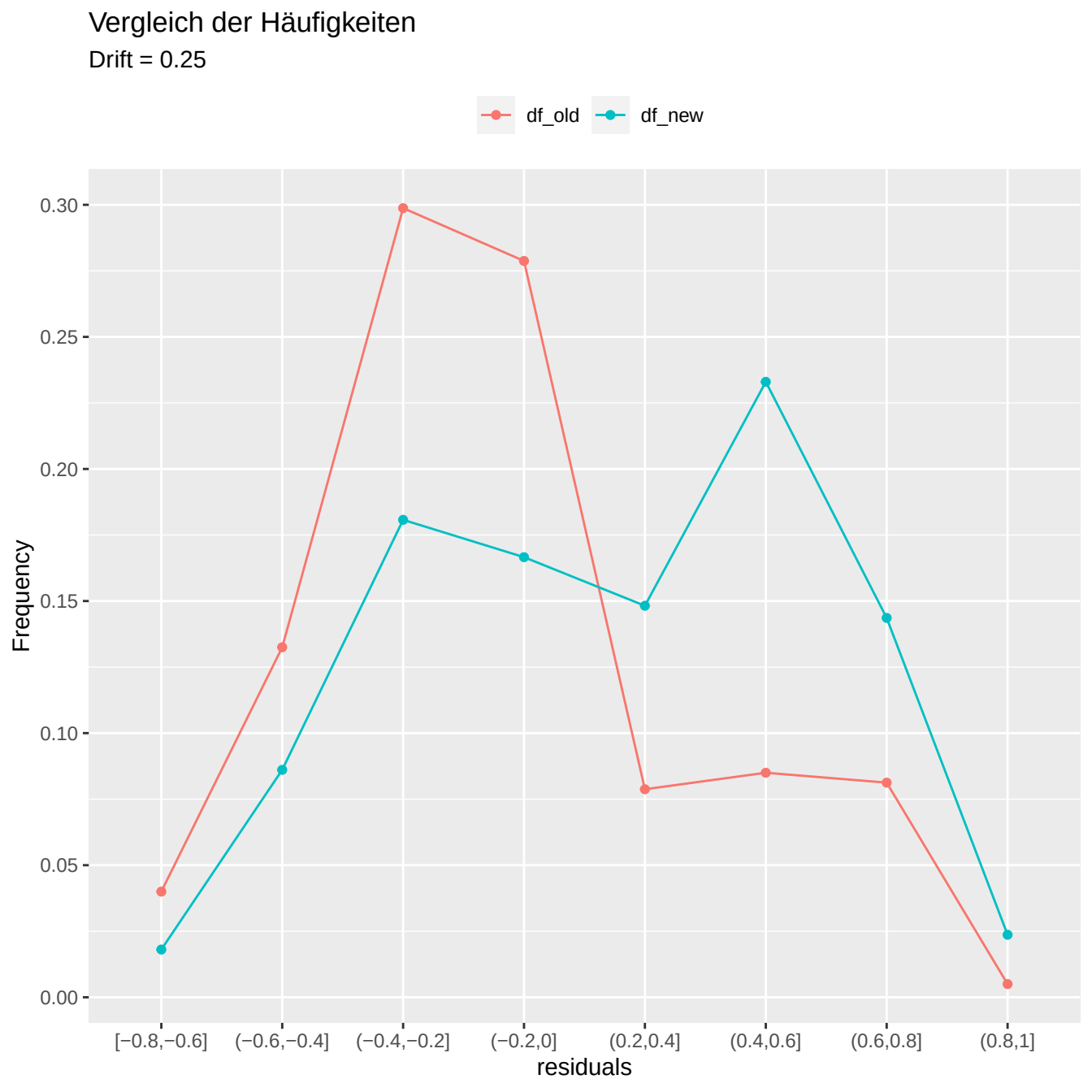

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
Count
0
50
100
150
TotalCharges

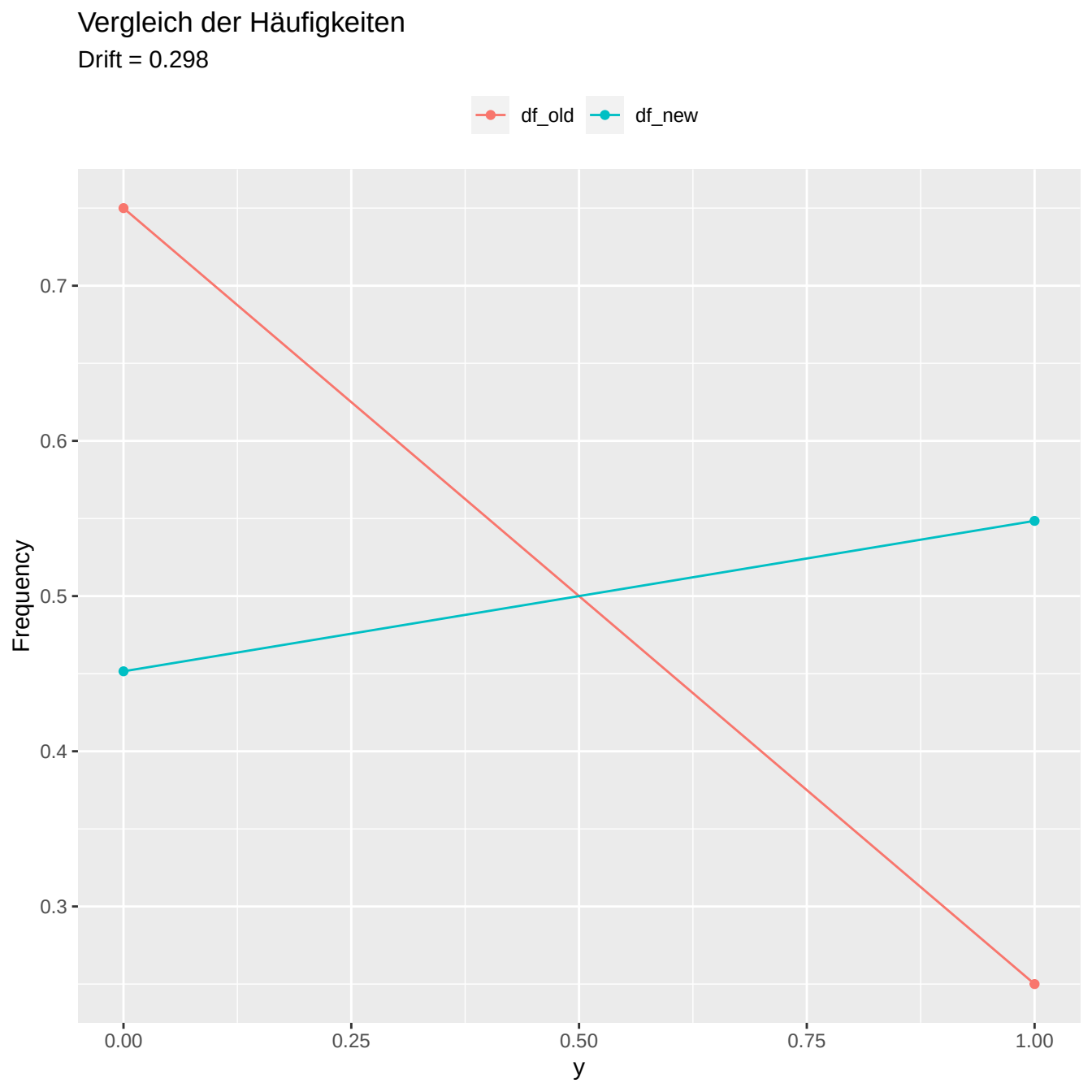

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
10
20
30
40
50
60
70
Count
0
50
100
tenure

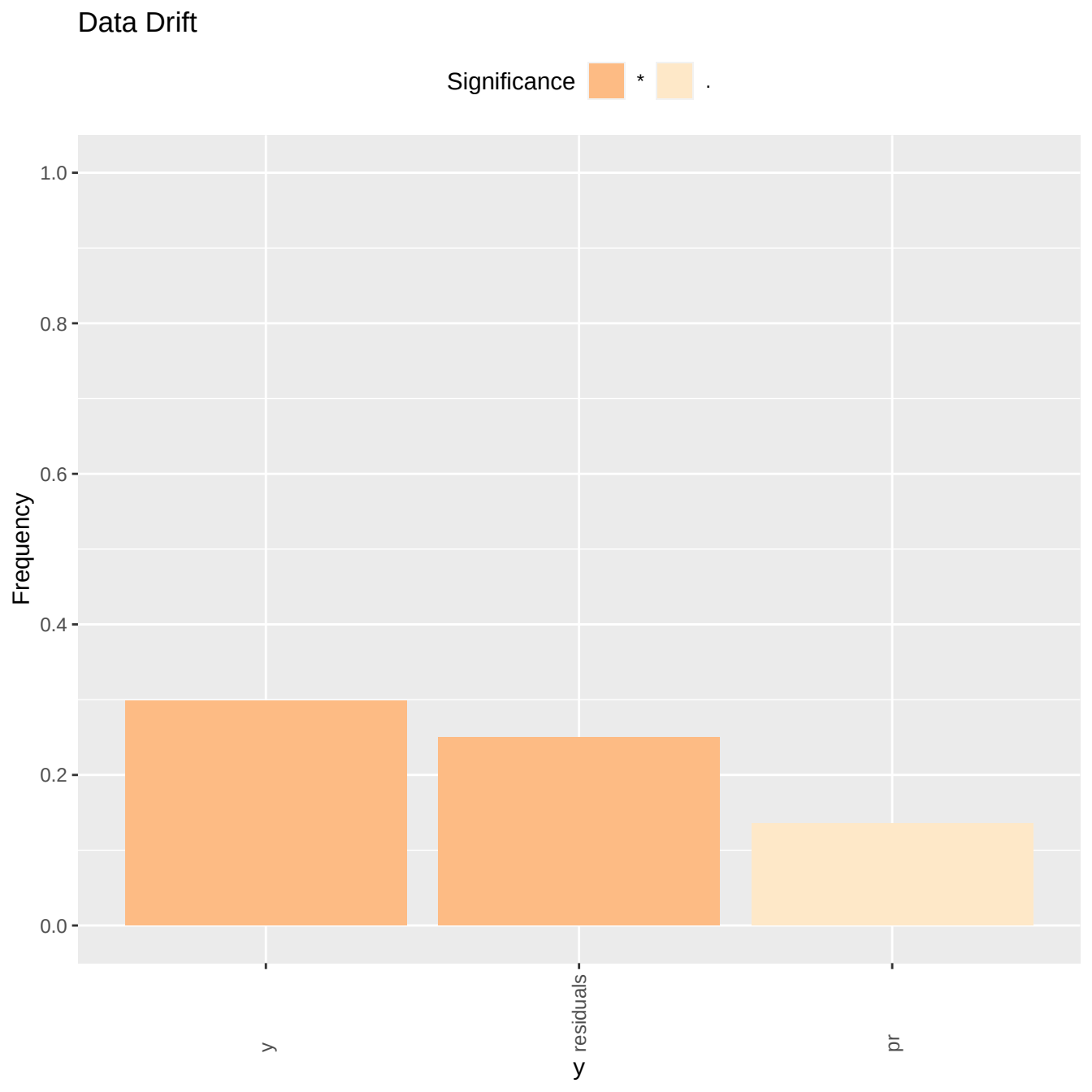

CV error
0.804
0.75
0.50
0.25
0.00
auc
0.950
0.975
1.000
1.025
1.05
Run

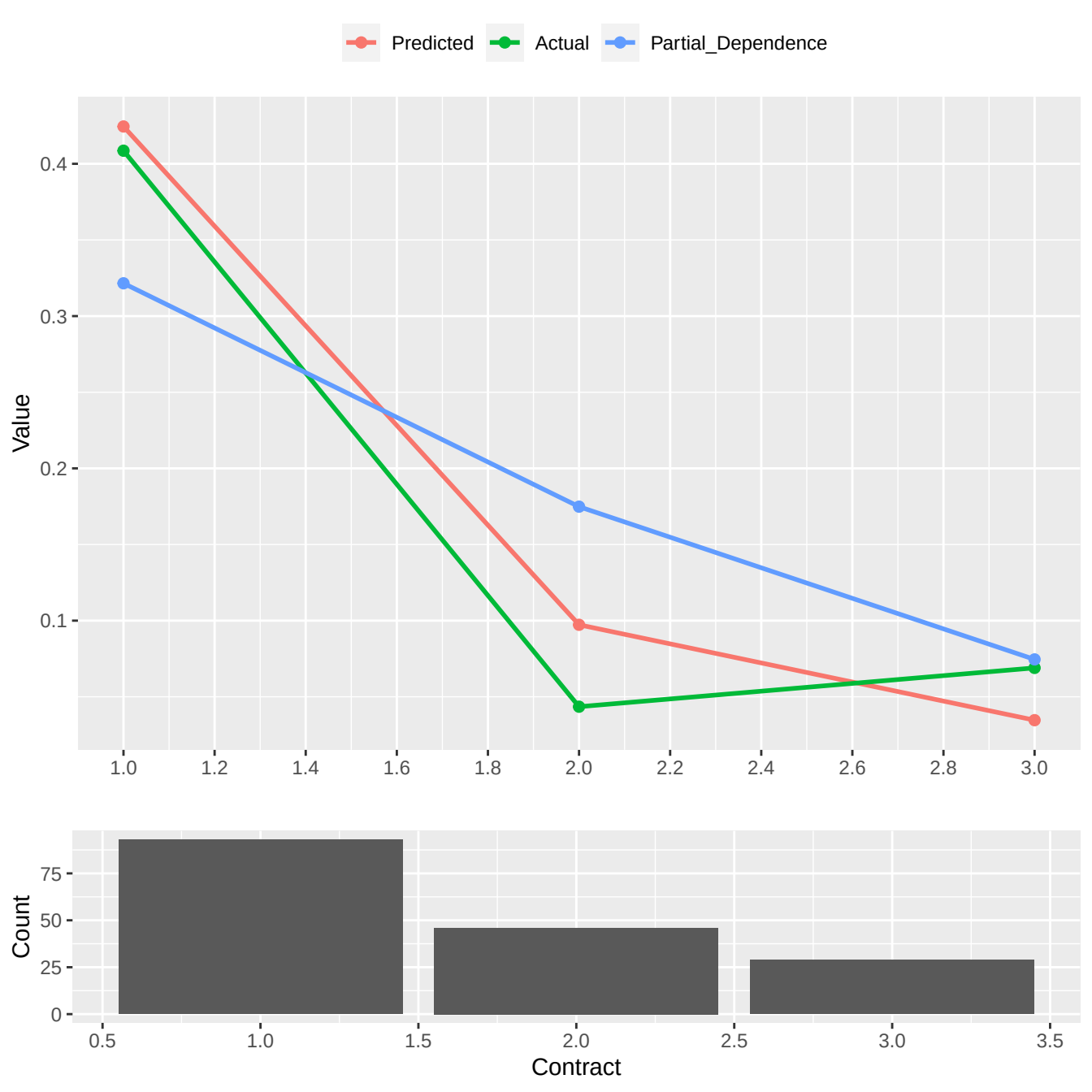

CV error
0.804
0.75
0.50
0.25
0.00
auc
1.00
1.25
1.50
1.75
2.00
Run

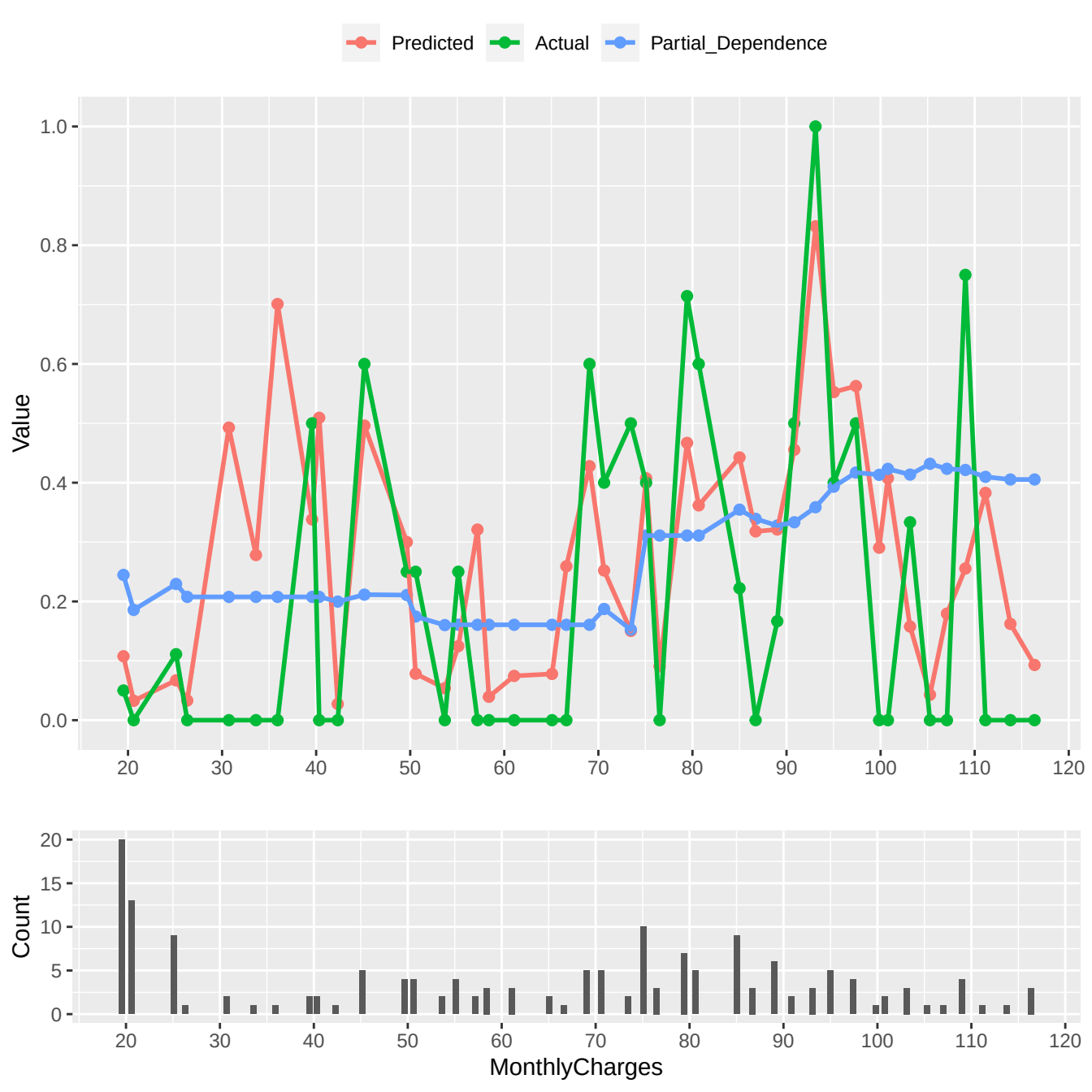

Vergleich der Häufigkeiten
Drift = 0.135
df_old
df_new
Frequency
0.25
0.20
0.15
0.10
0.05
0.00
[0.15,0.2]
[0.2,0.25]
[0.25,0.3]
[0.3,0.35]
[0.35,0.4]
[0.4,0.45]
[0.45,0.5]
[0.5,0.55]
[0.55,0.6]
[0.6,0.65]
[0.65,0.7]
[0.7,0.75]
[0.75,0.8]
pr

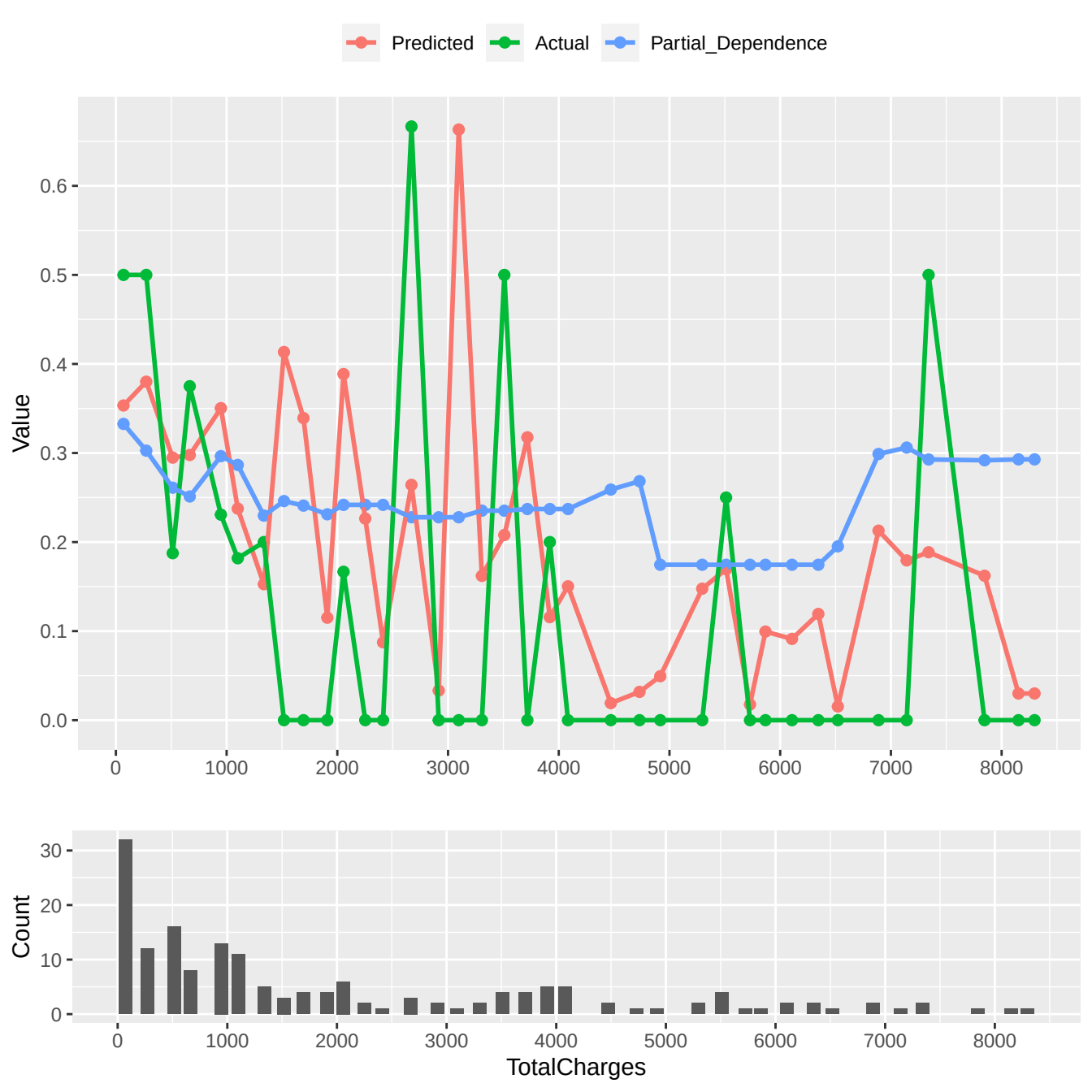

Vergleich der Häufigkeiten
Drift = 0.25
df_old
df_new
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Frequency
[−0.8,−0.6]
(−0.6,−0.4]
(−0.4,−0.2]
(−0.2,0]
(0.2,0.4]
(0.4,0.6]
(0.6,0.8]
(0.8,1]
residuals

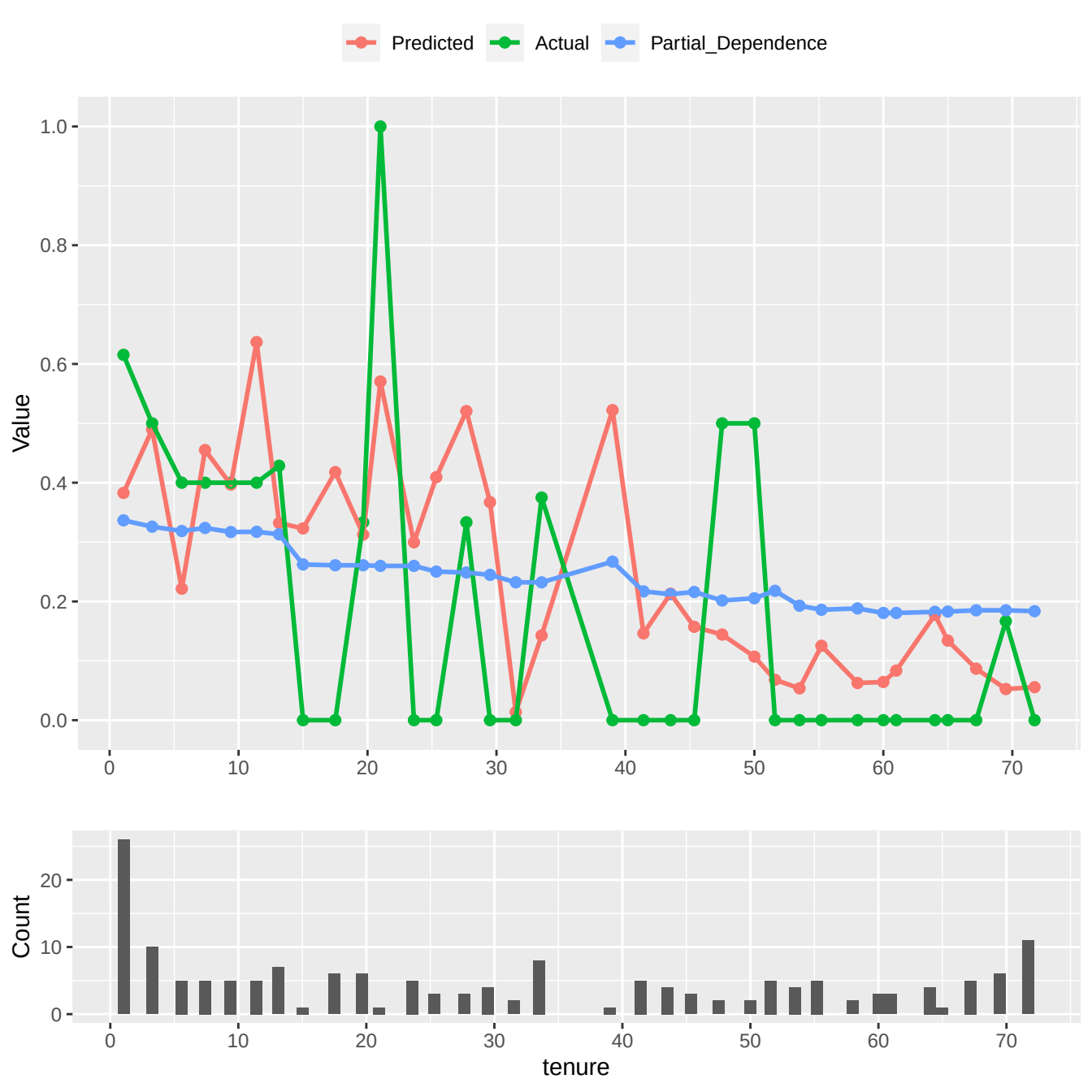

Vergleich der Häufigkeiten
Drift = 0.298
df_old
df_new
Frequency
0.7
0.6
0.5
0.4
0.3
0.00
0.25
0.50
0.75
1.00
y

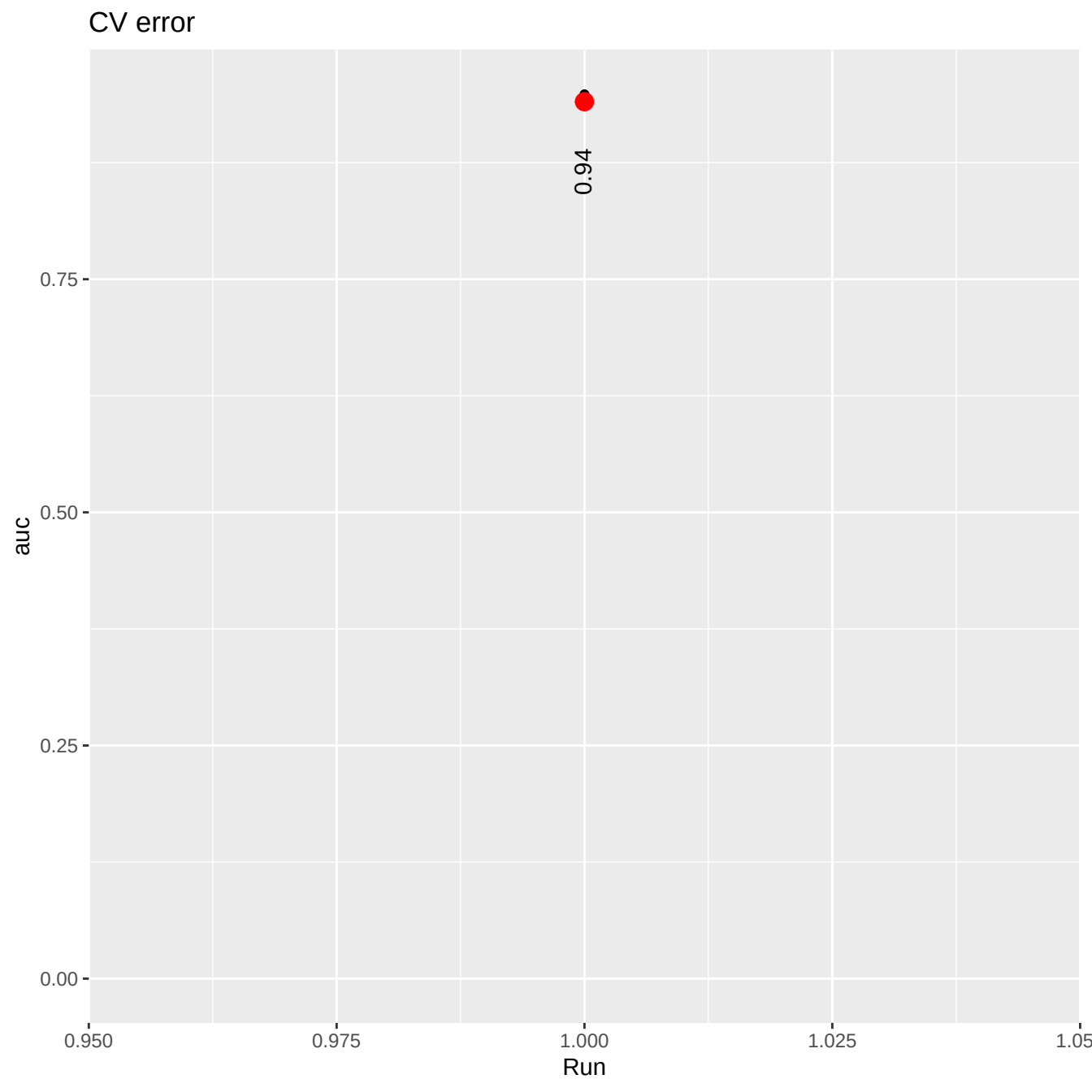

Data Drift
Significance
*
.
1.0
0.8
0.6
0.4
0.2
0.0
Frequency
y
residuals
pr
y

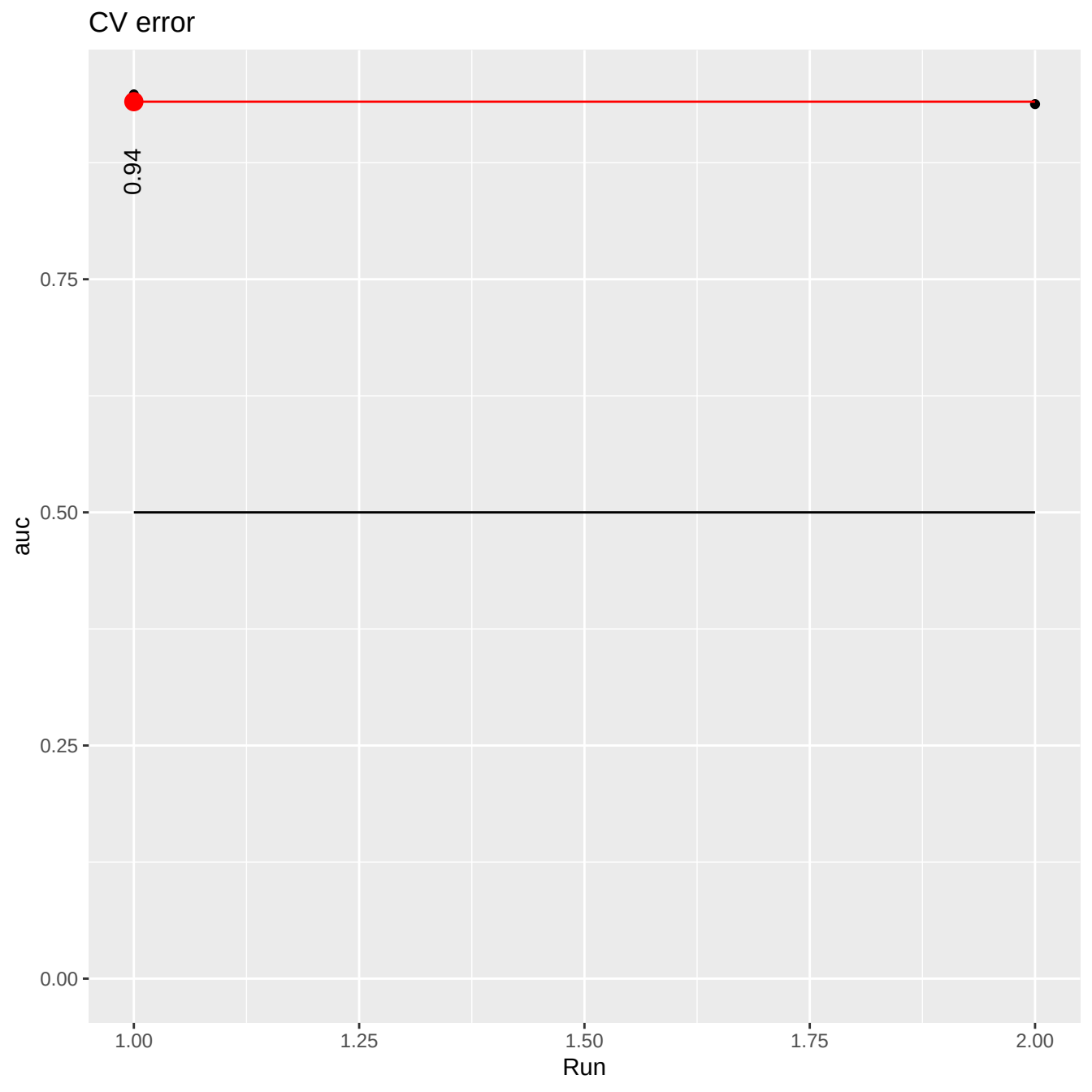

Predicted
Actual
Partial_Dependence
Value
0.4
0.3
0.2
0.1
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
75
50
25
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Contract

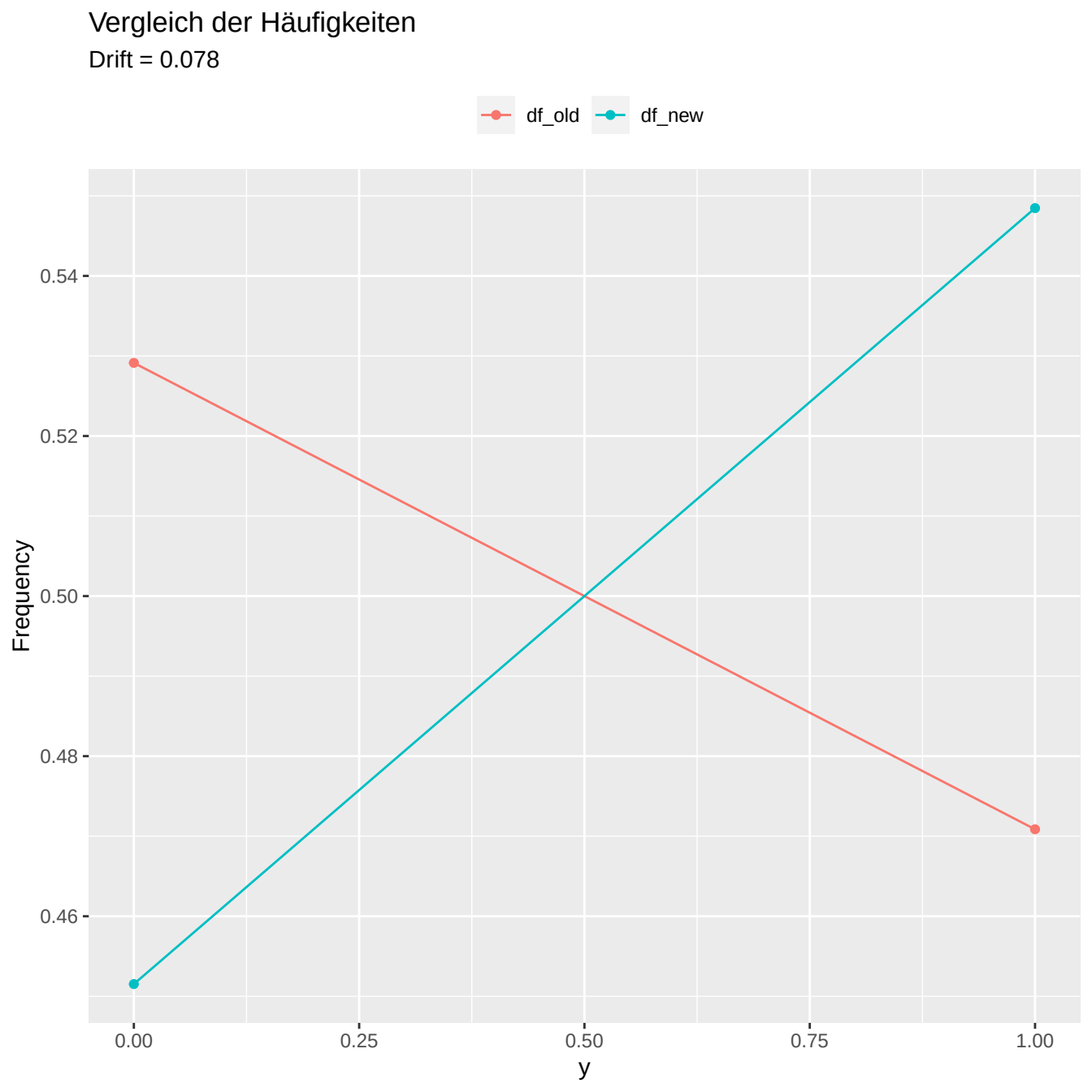

Predicted
Actual
Partial_Dependence
Value
1.0
0.8
0.6
0.4
0.2
0.0
20
30
40
50
60
70
80
90
100
110
120
Count
20
15
10
5
0
MonthlyCharges

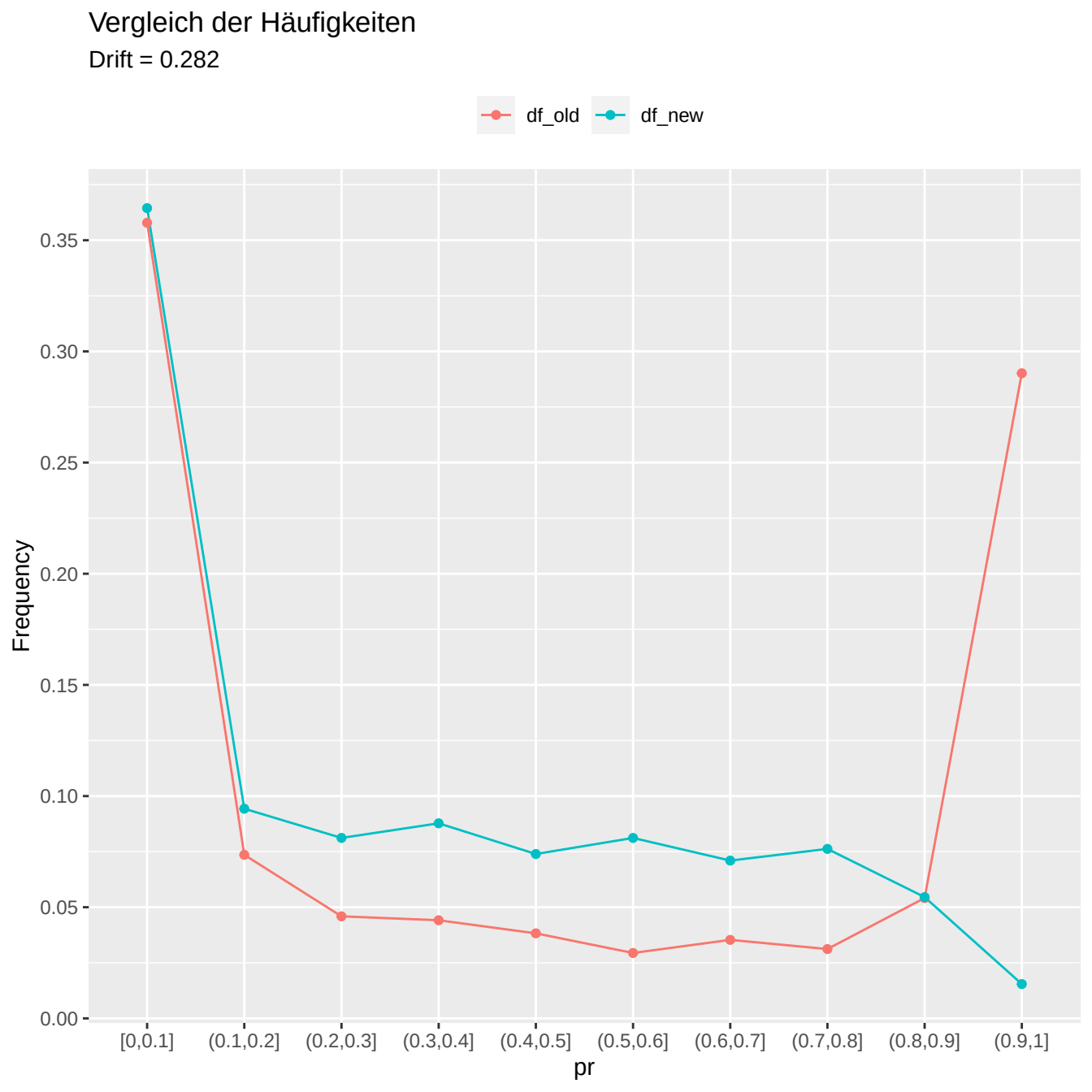

Predicted
Actual
Partial_Dependence
Value
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
1000
2000
3000
4000
5000
6000
7000
8000
Count
0
10
20
30
TotalCharges

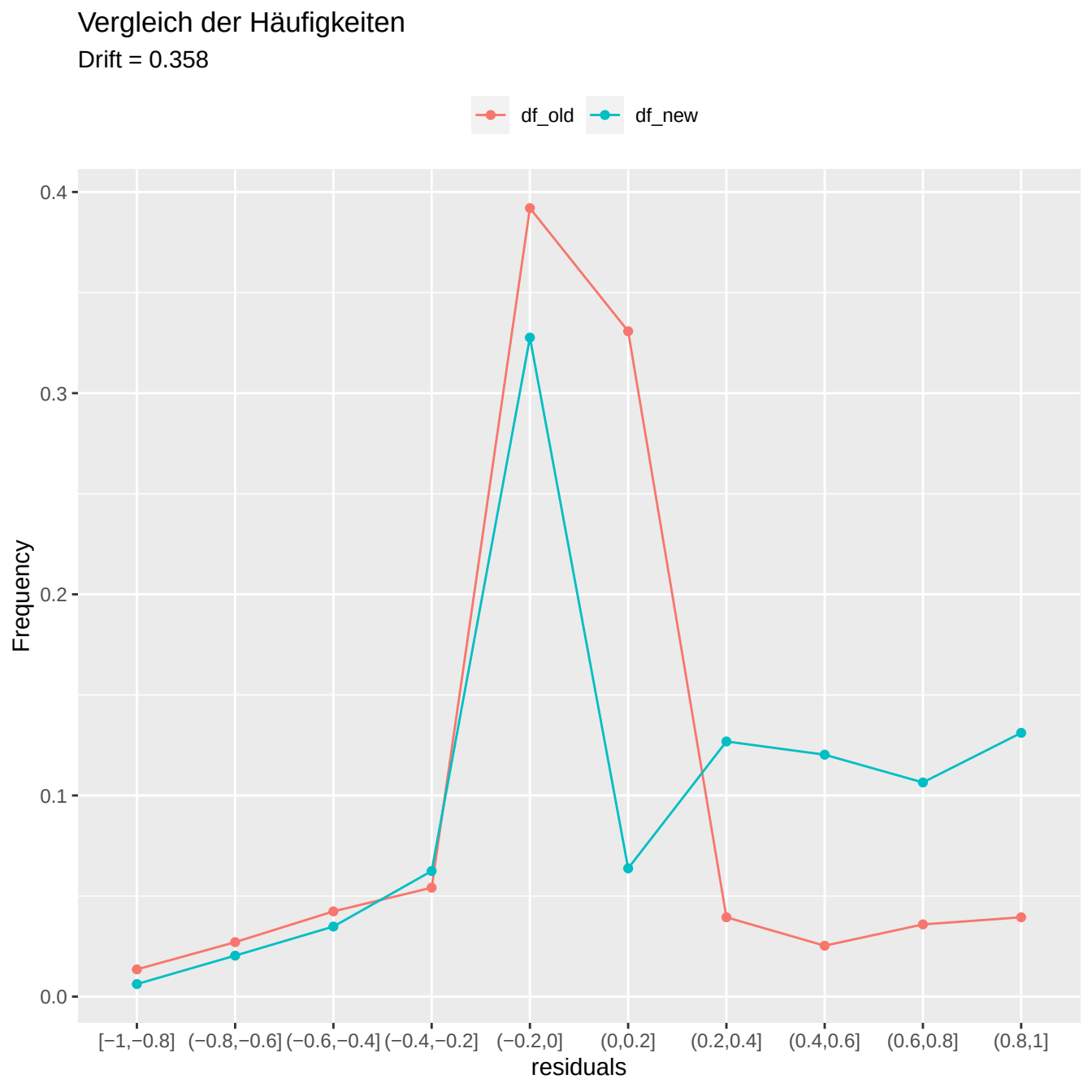

Predicted
Actual
Partial_Dependence
Value
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
60
70
Count
20
10
0
tenure

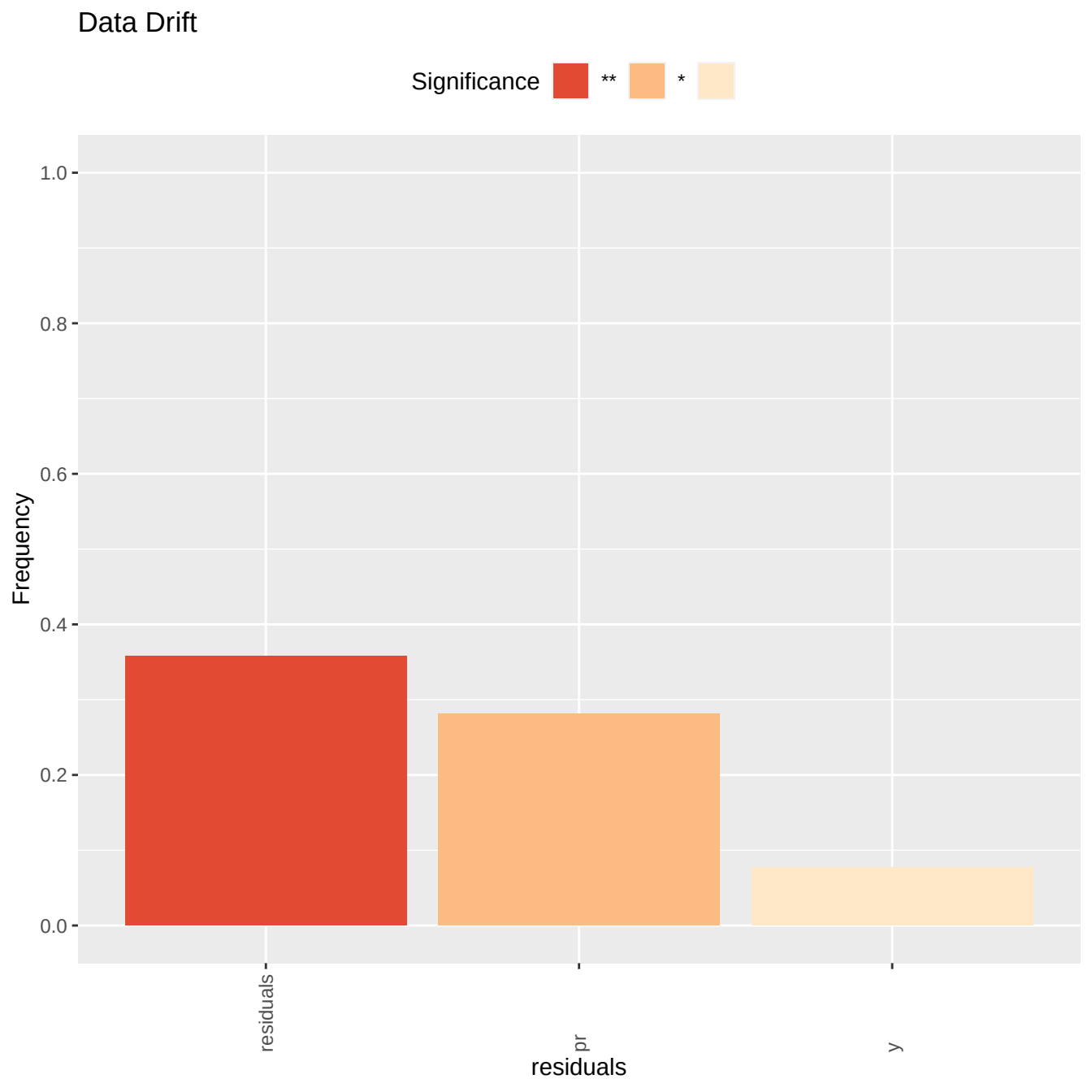

CV error
0.94
0.75
0.50
0.25
0.00
auc
0.950
0.975
1.000
1.025
1.05
Run

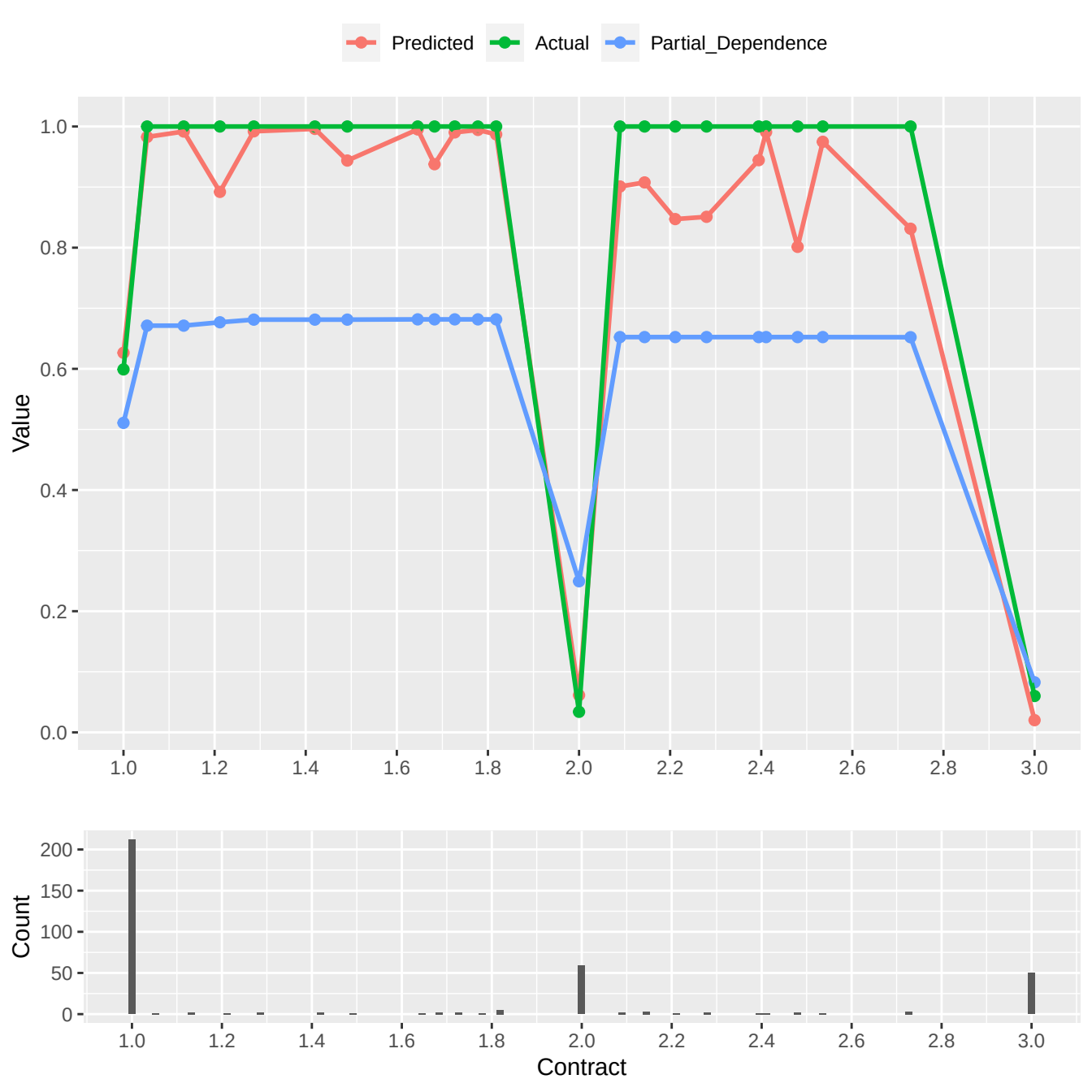

CV error
0.94
0.75
0.50
0.25
0.00
auc
1.00
1.25
1.50
1.75
2.00
Run

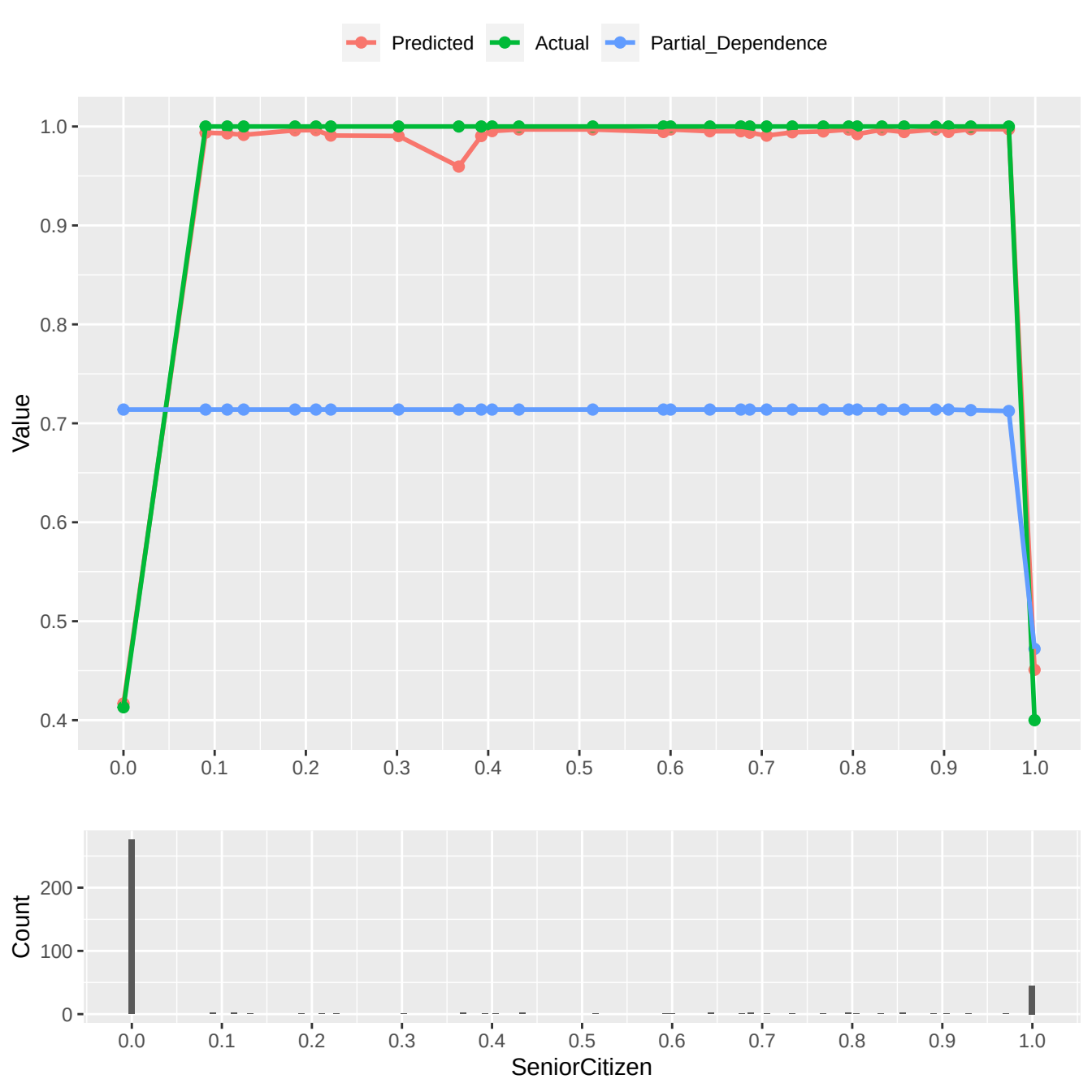

Vergleich der Häufigkeiten
Drift = 0.078
df_old
df_new
0.54
0.52
0.50
0.48
0.46
Frequency
0.00
0.25
0.50
0.75
1.00
y

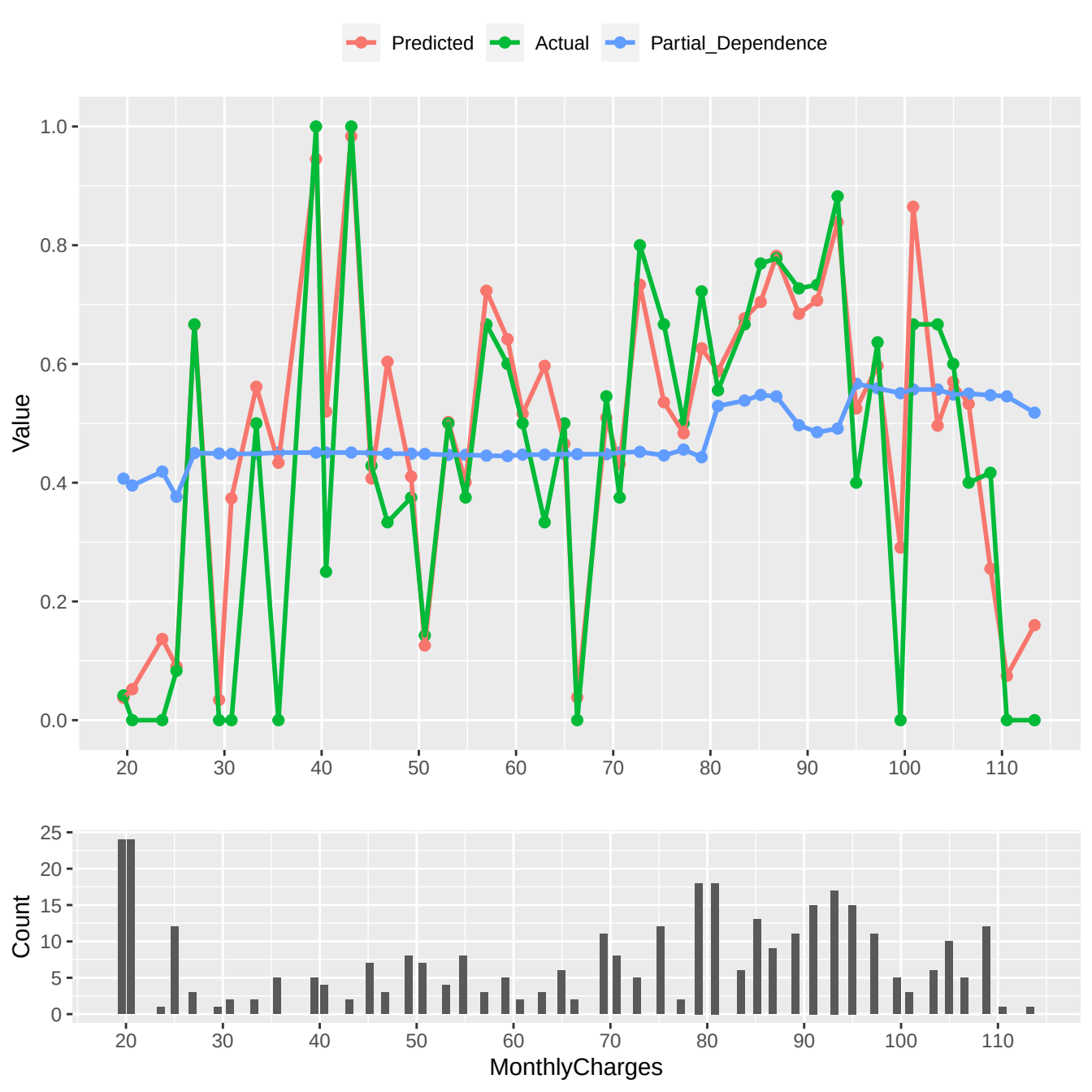

Vergleich der Häufigkeiten
Drift = 0.282
df_old
df_new
Frequency
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
[0,0.1]
(0.1,0.2]
(0.2,0.3]
(0.3,0.4]
(0.4,0.5]
(0.5,0.6]
(0.6,0.7]
(0.7,0.8]
(0.8,0.9]
(0.9,1]
pr

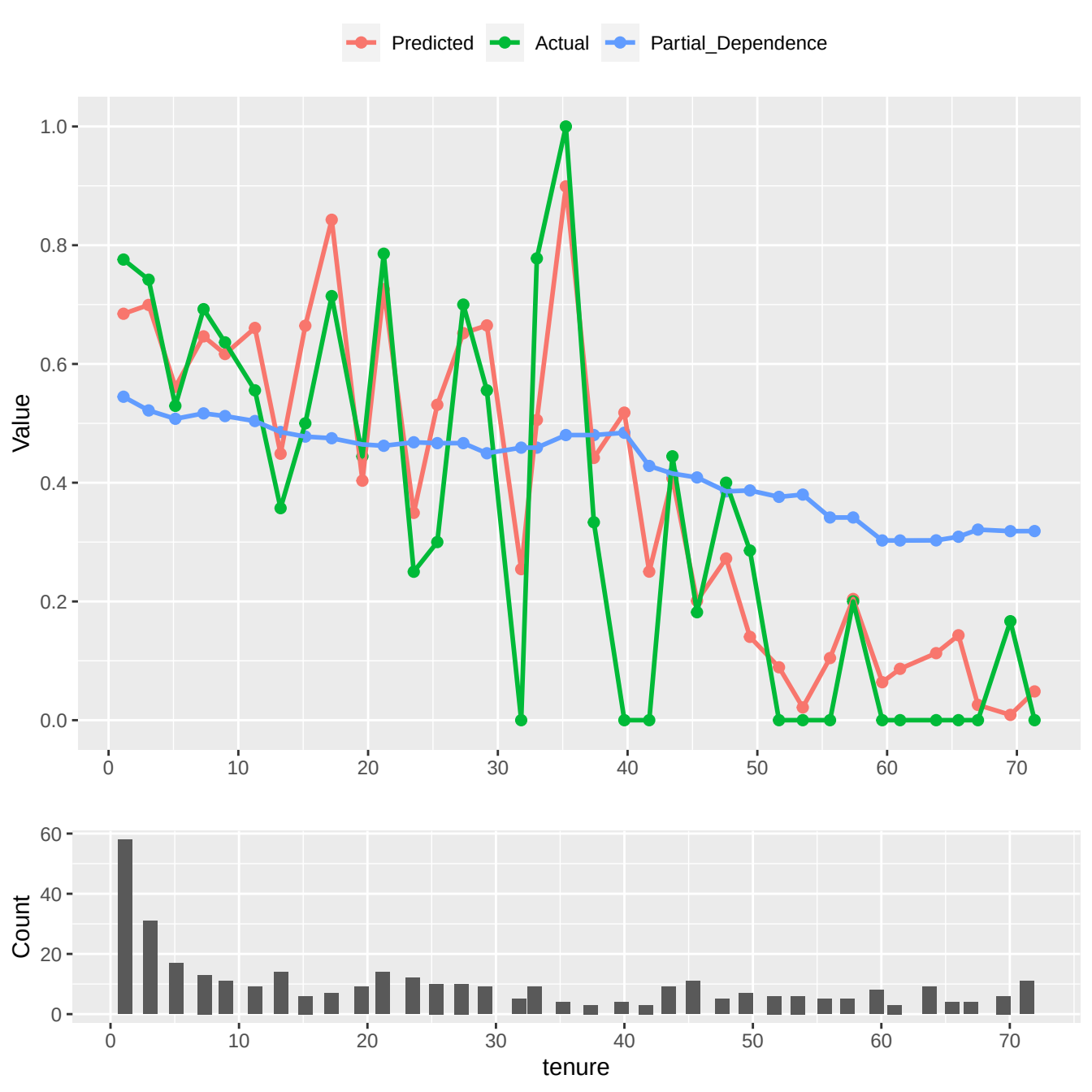

Vergleich der Häufigkeiten
Drift = 0.358
df_old
df_new
0.4
0.3
0.2
0.1
0.0
Frequency
[−1,−0.8]
(−0.8,−0.6]
(−0.6,−0.4]
(−0.4,−0.2]
(−0.2,0]
(0,0.2]
(0.2,0.4]
(0.4,0.6]
(0.6,0.8]
(0.8,1]
residuals

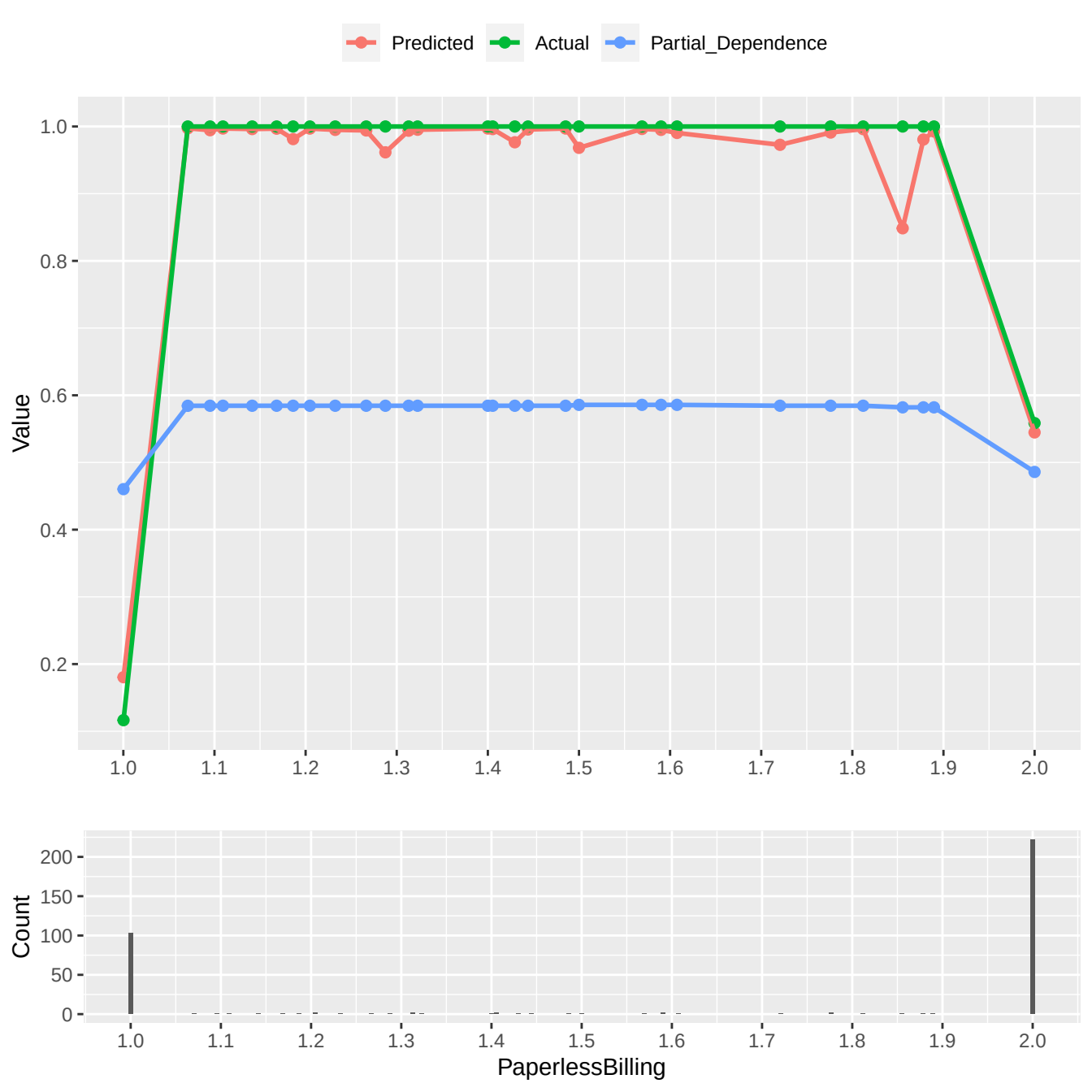

Data Drift
Significance
**
*
1.0
0.8
0.6
0.4
0.2
0.0
Frequency
residuals
pr
y
residuals

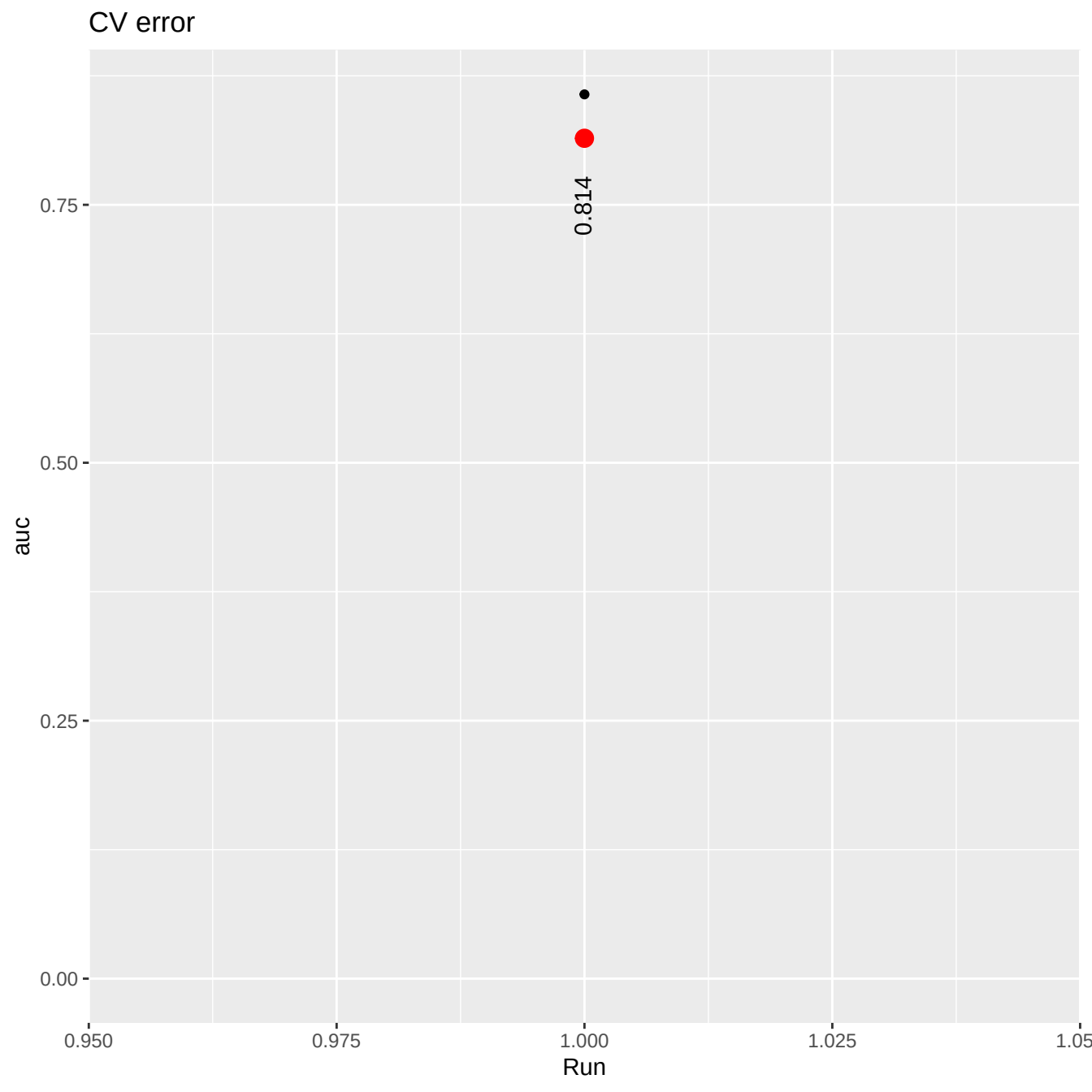

Predicted
Actual
Partial_Dependence
Value
1.0
0.8
0.6
0.4
0.2
0.0
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
200
150
100
50
0
Contract

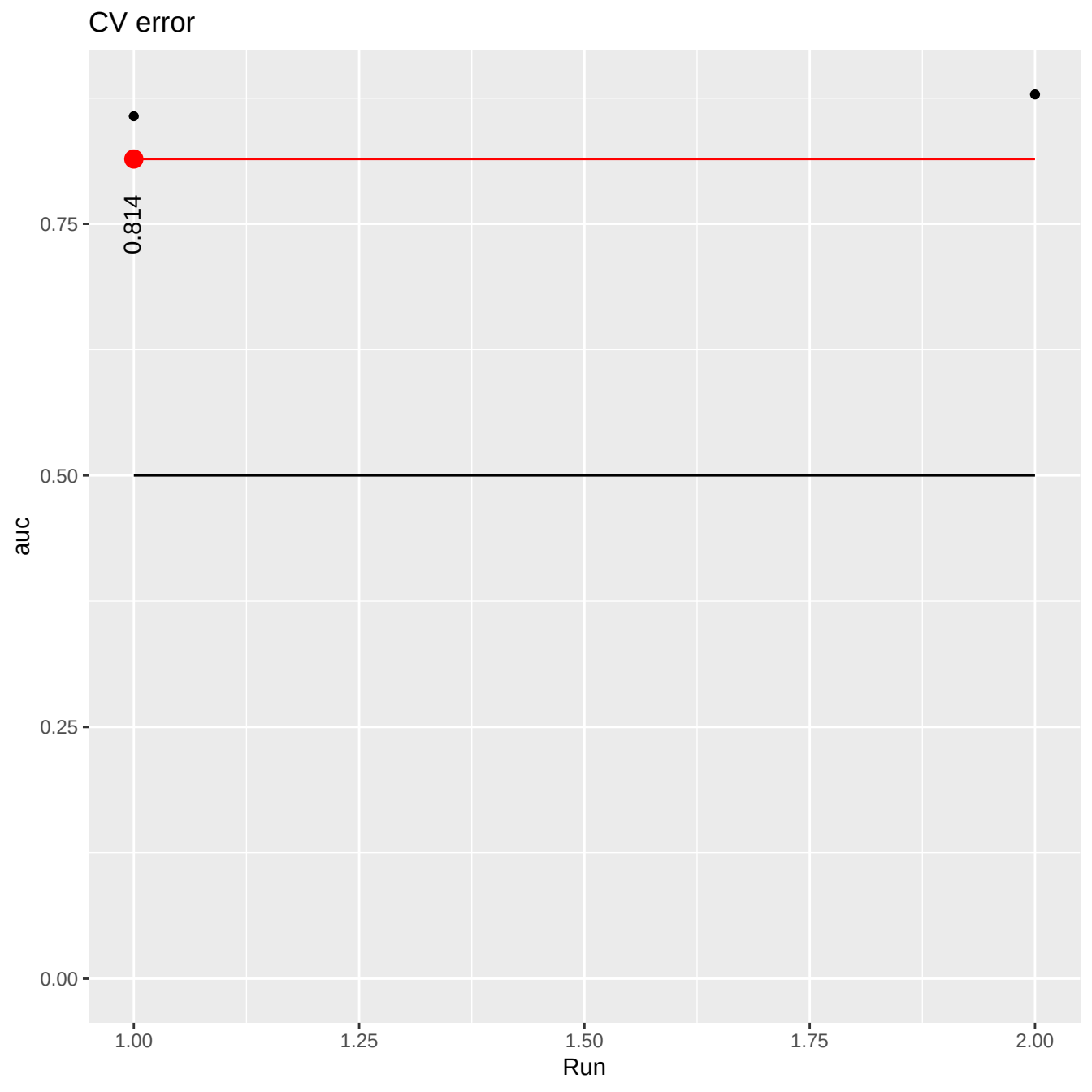

Predicted
Actual
Partial_Dependence
Value
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Count
200
100
0
SeniorCitizen

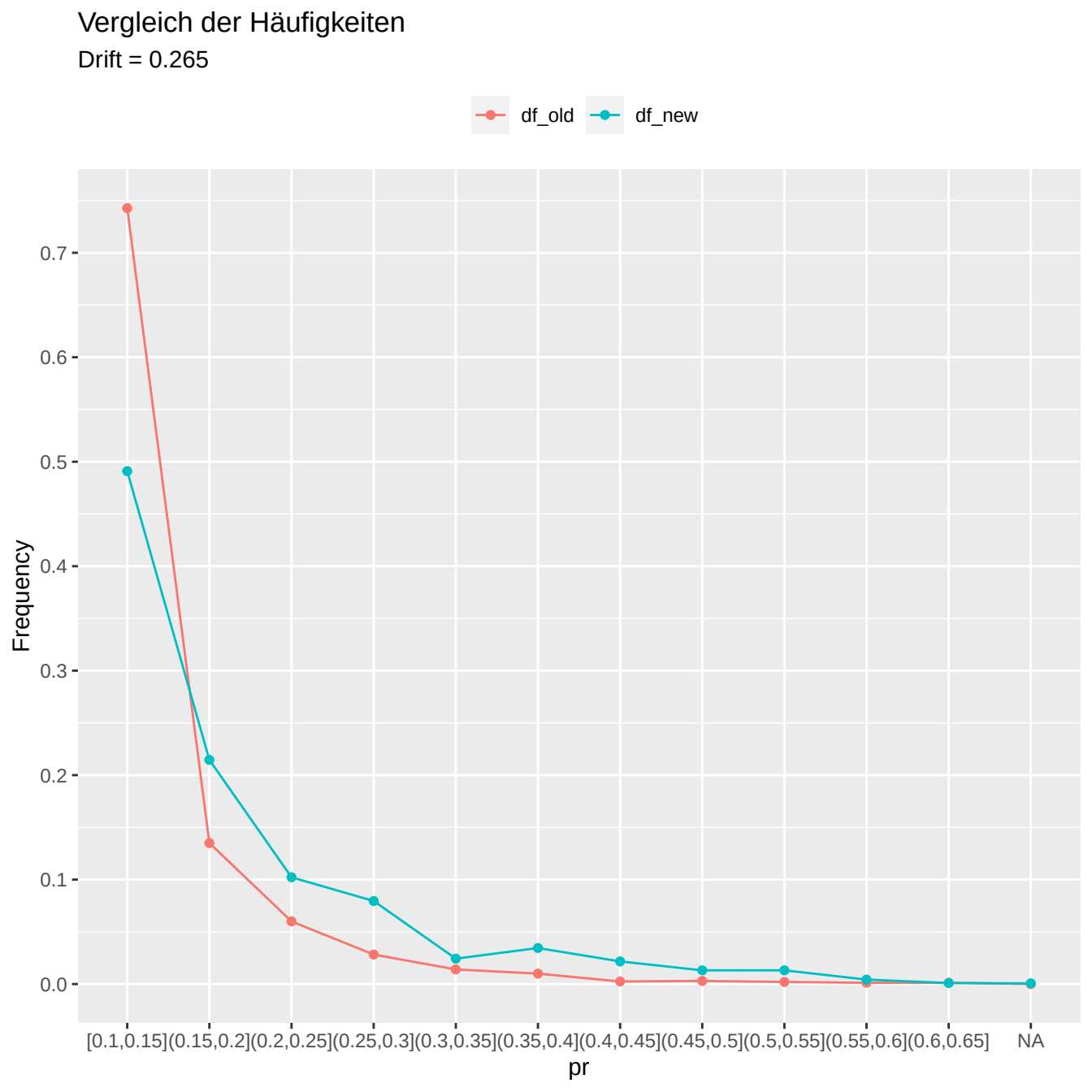

Predicted
Actual
Partial_Dependence
Value
1.0
0.8
0.6
0.4
0.2
0.0
20
30
40
50
60
70
80
90
100
110
Count
25
20
15
10
5
0
MonthlyCharges

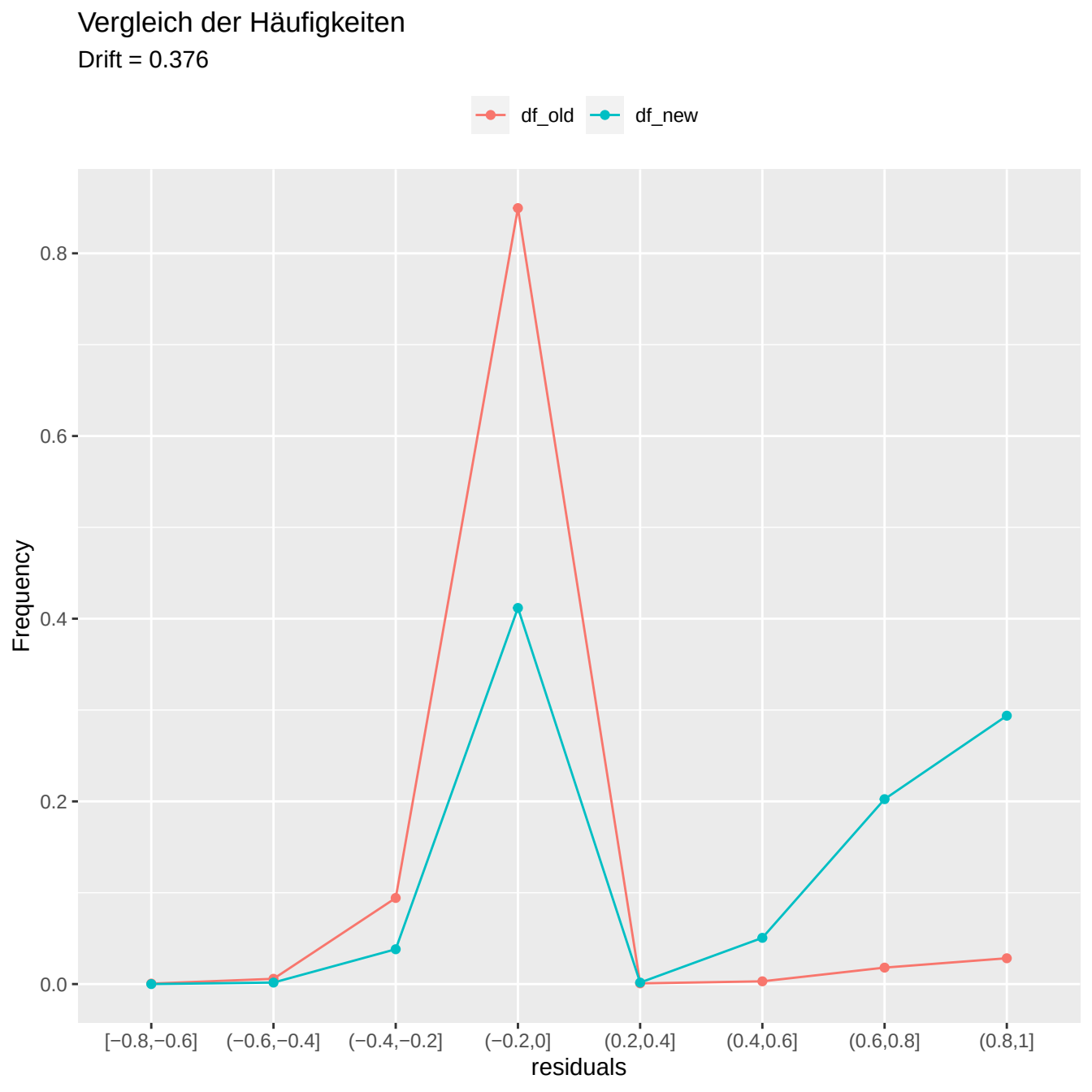

Predicted
Actual
Partial_Dependence
Value
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
60
70
Count
60
40
20
0
tenure

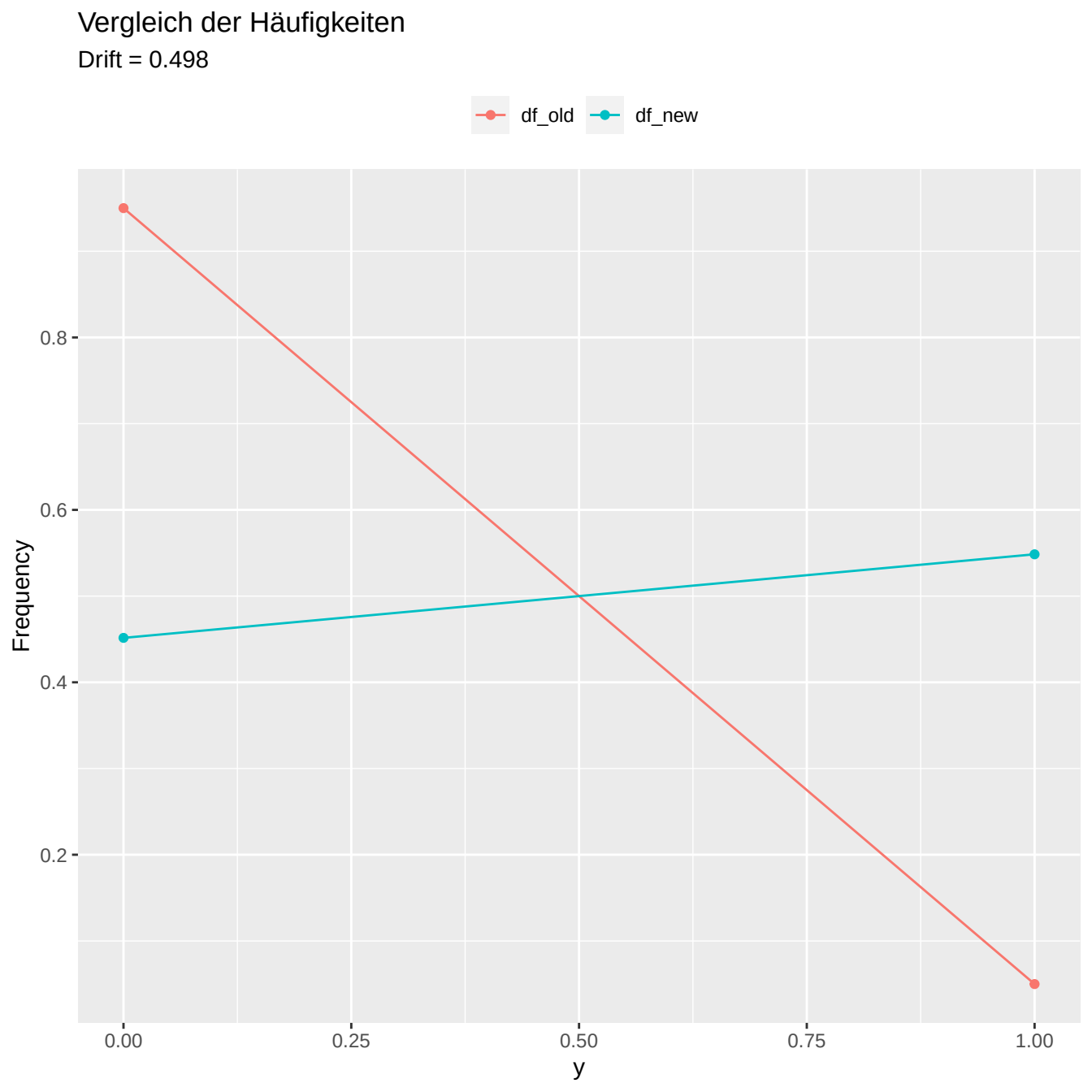

Predicted
Actual
Partial_Dependence
Value
1.0
0.8
0.6
0.4
0.2
1.0
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
2.0
Count
200
150
100
50
0
PaperlessBilling

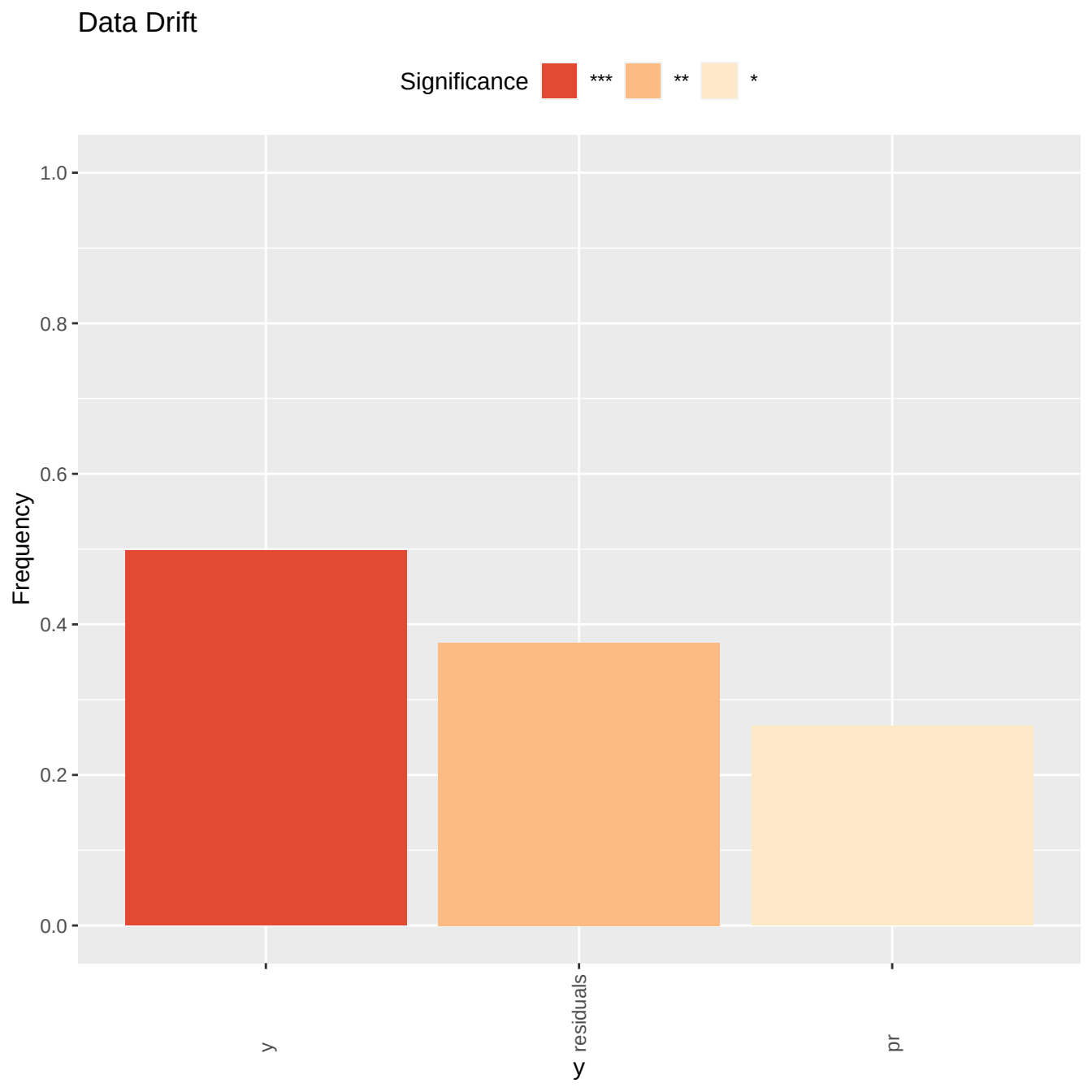

CV error
0.814
0.75
0.50
0.25
0.00
auc
0.950
0.975
1.000
1.025
1.05
Run

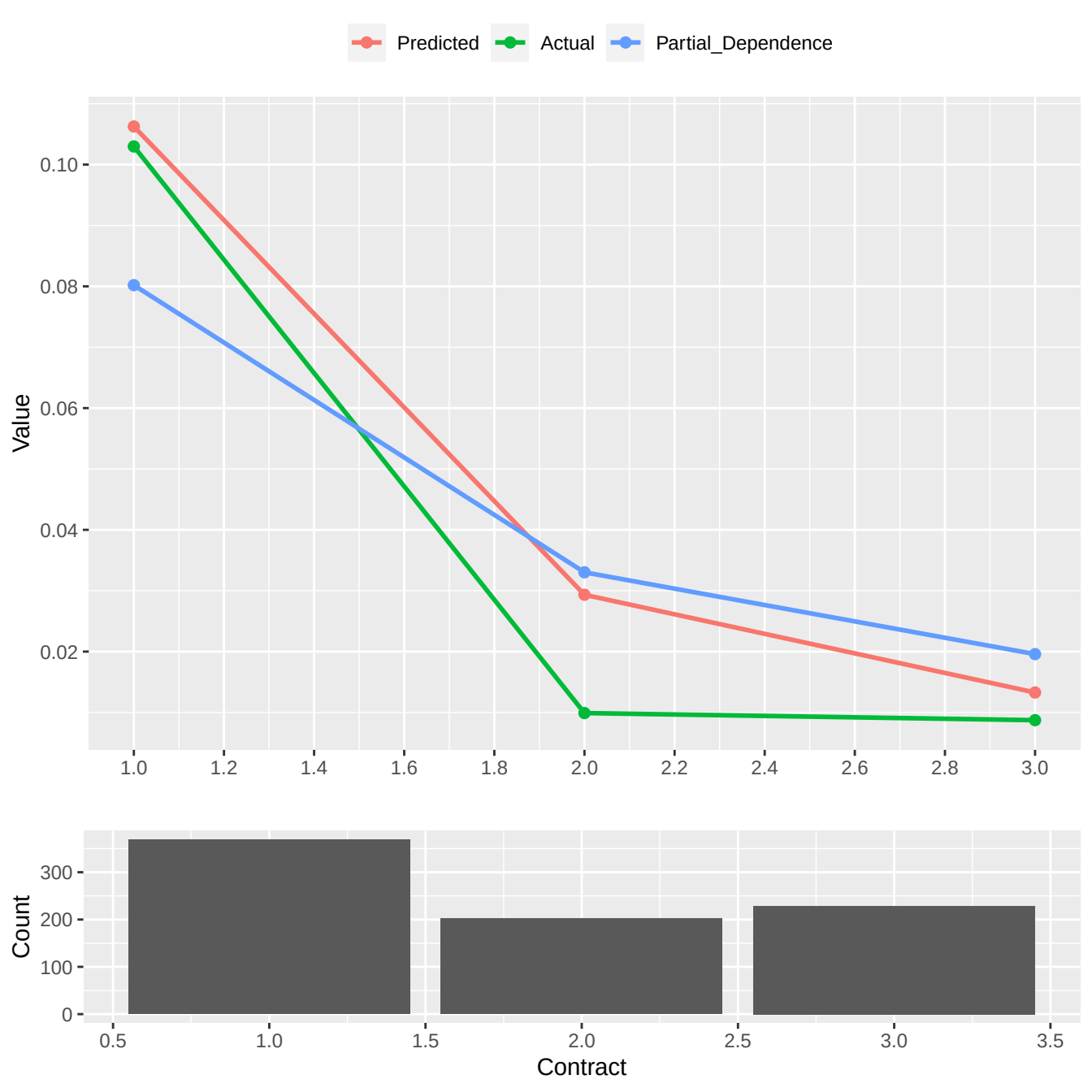

CV error
0.814
0.75
0.50
0.25
0.00
auc
1.00
1.25
1.50
1.75
2.00
Run

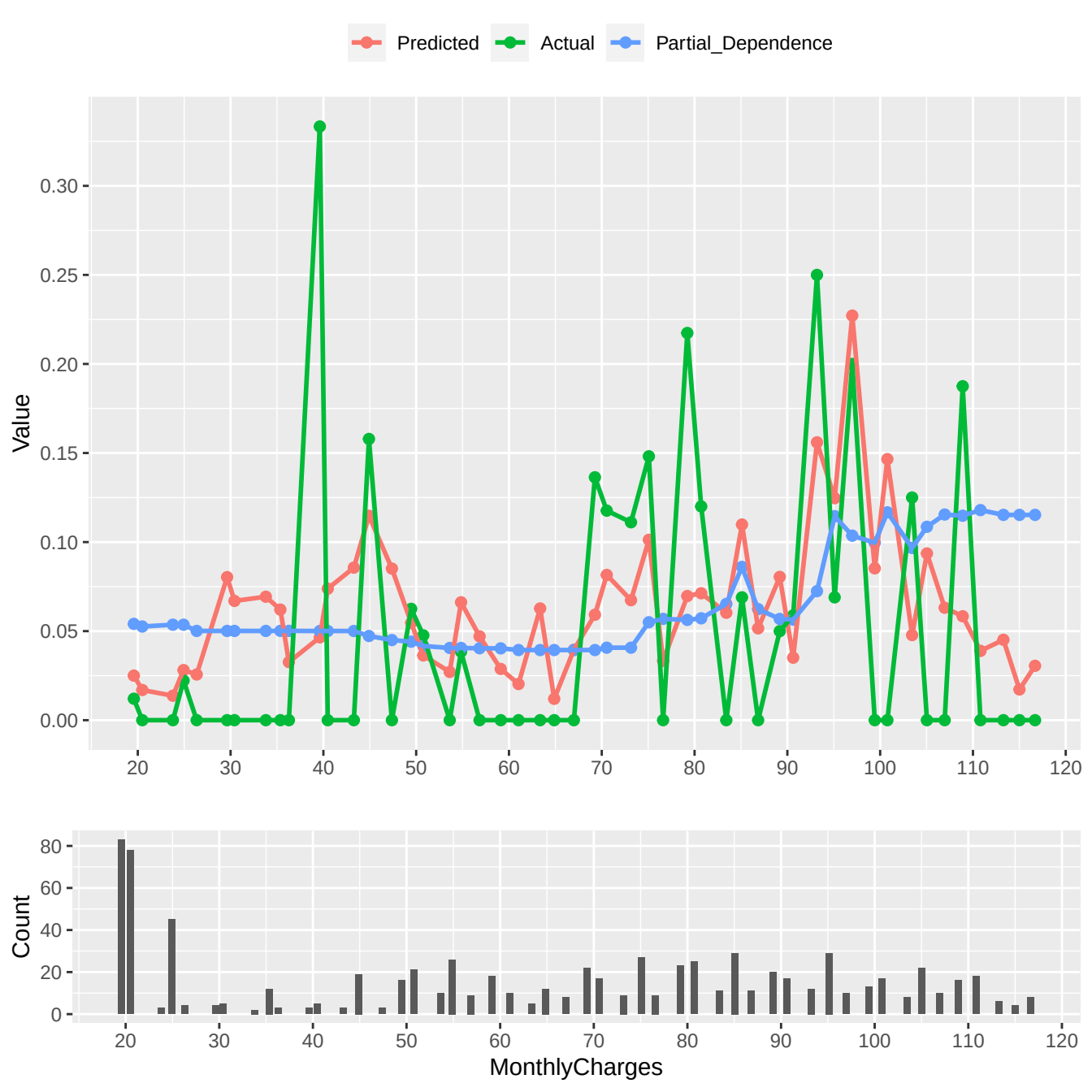

Vergleich der Häufigkeiten
Drift = 0.265
df_old
df_new
Frequency
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
[0.1,0.15]
(0.15,0.2]
(0.2,0.25]
(0.25,0.3]
(0.3,0.35]
(0.35,0.4]
(0.4,0.45]
(0.45,0.5]
(0.5,0.55]
(0.55,0.6]
(0.6,0.65]
NA
pr

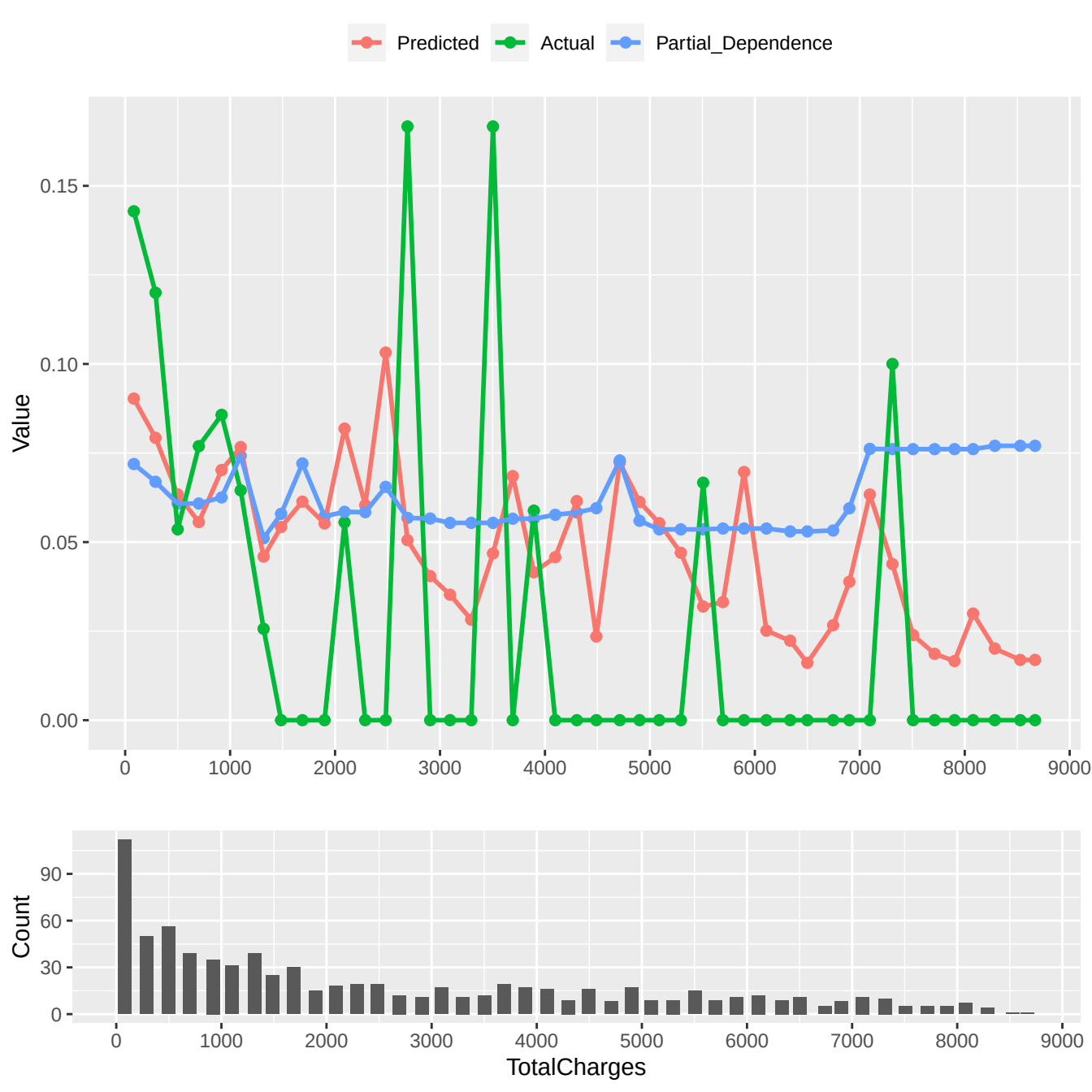

Vergleich der Häufigkeiten
Drift = 0.376
df_old
df_new
0.8
0.6
0.4
0.2
0.0
Frequency
[−0.8,−0.6]
(−0.6,−0.4]
(−0.4,−0.2]
(−0.2,0]
(0.2,0.4]
(0.4,0.6]
(0.6,0.8]
(0.8,1]
residuals


Vergleich der Häufigkeiten
Drift = 0.498
df_old
df_new
Frequency
0.8
0.6
0.4
0.2
0.00
0.25
0.50
0.75
1.00
y

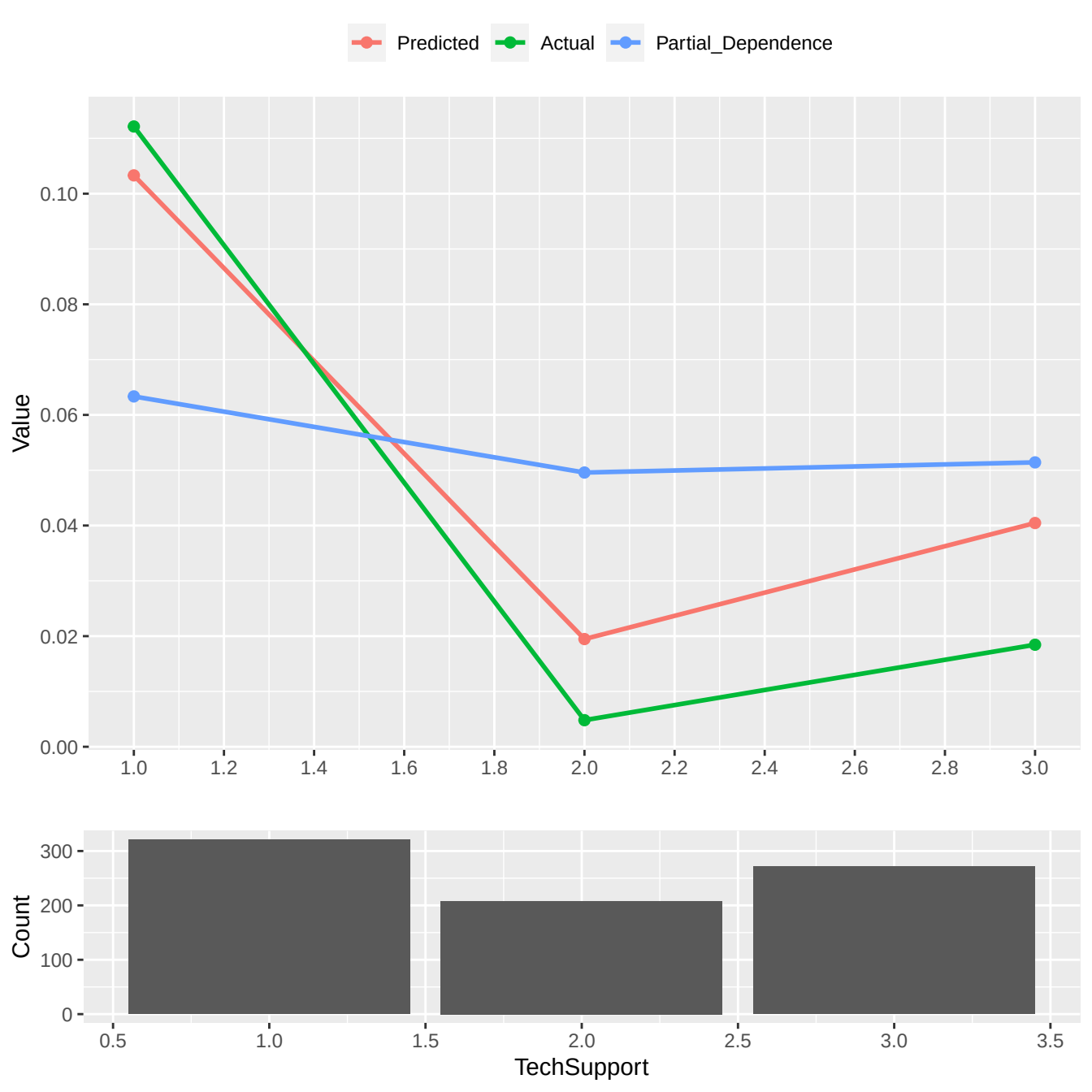

Data Drift
Significance
***
**
*
1.0
0.8
0.6
0.4
0.2
0.0
Frequency
y
residuals
pr
y

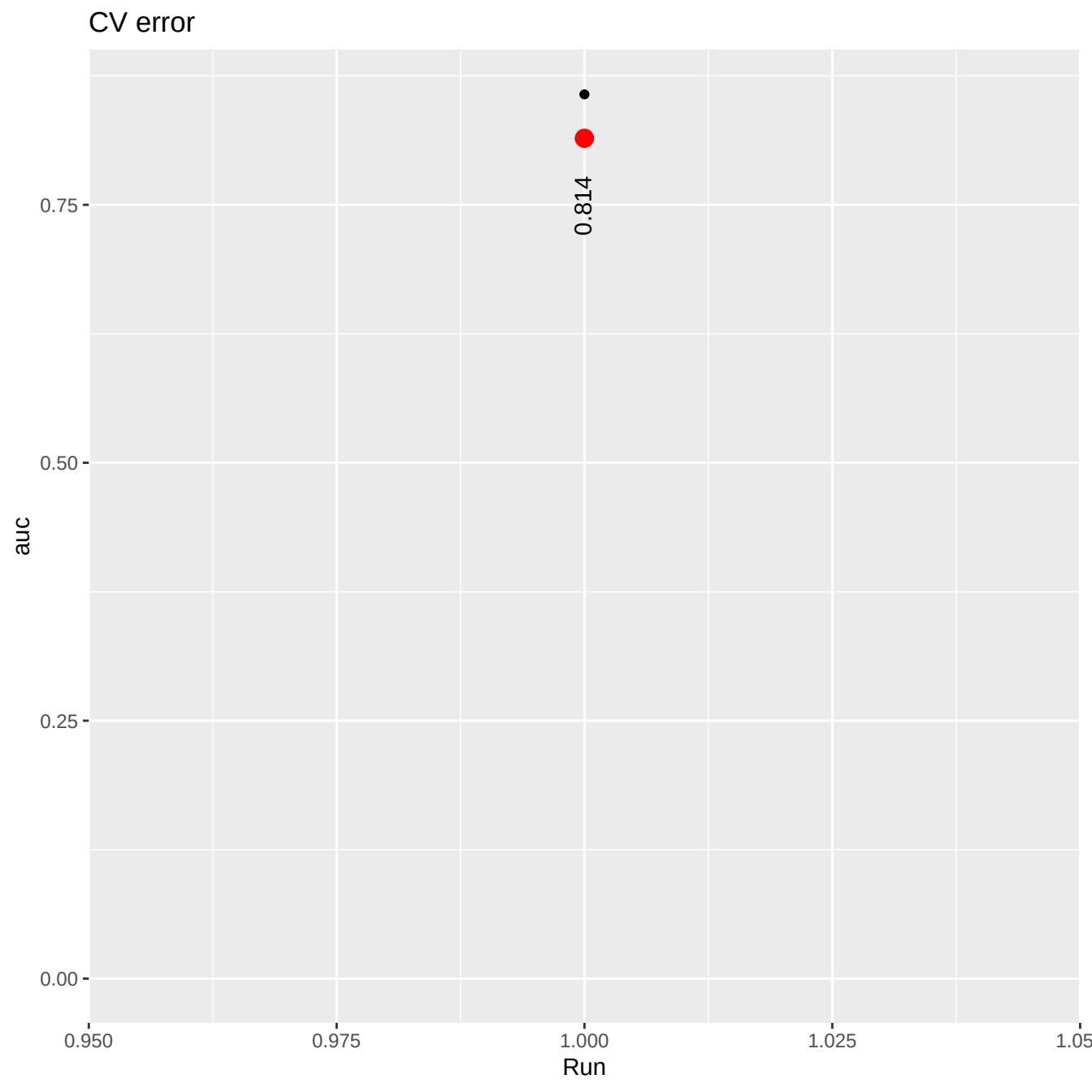

Predicted
Actual
Partial_Dependence
Value
0.10
0.08
0.06
0.04
0.02
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
300
200
100
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Contract

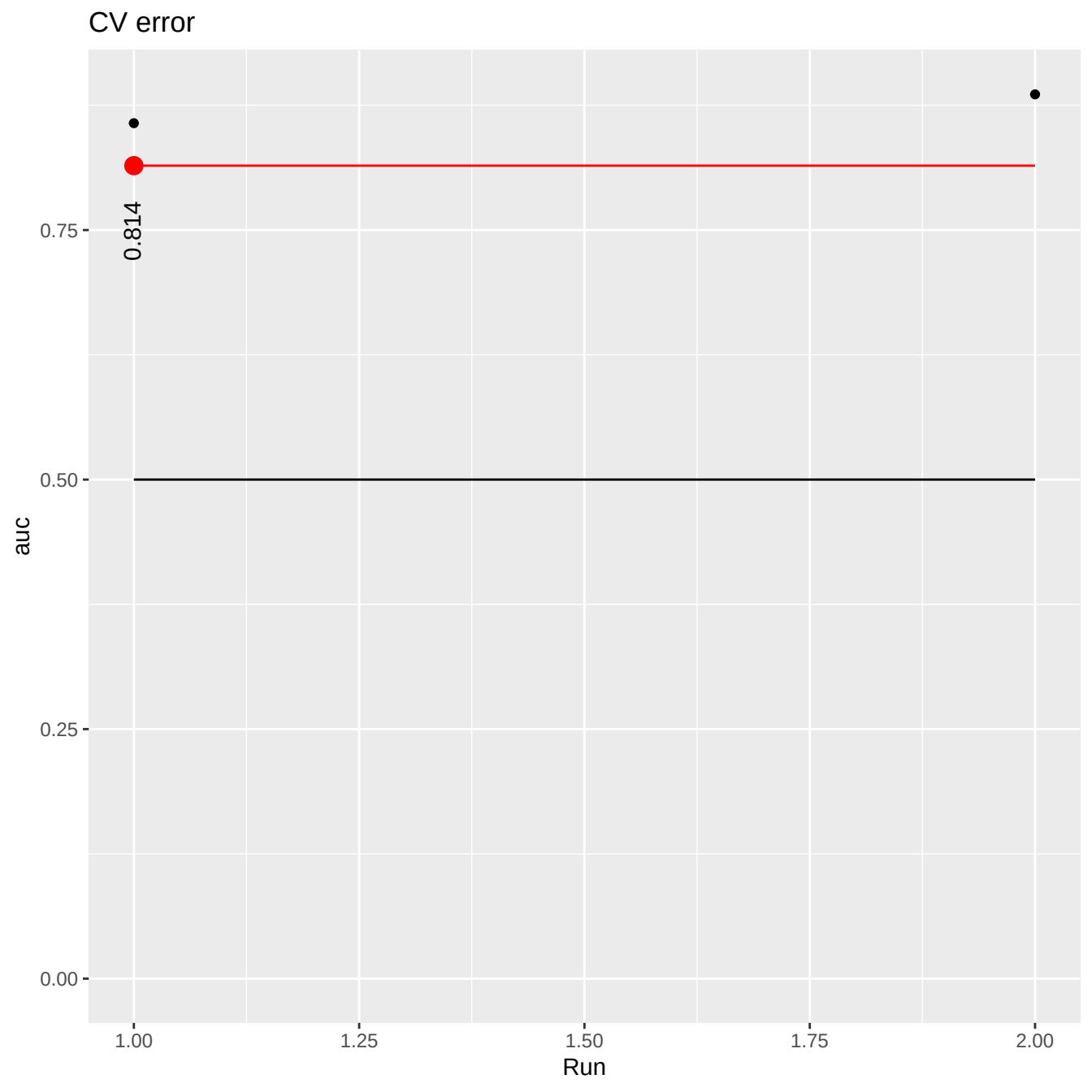

Predicted
Actual
Partial_Dependence
Value
0.30
0.25
0.20
0.15
0.10
0.05
0.00
20
30
40
50
60
70
80
90
100
110
120
Count
80
60
40
20
0
MonthlyCharges

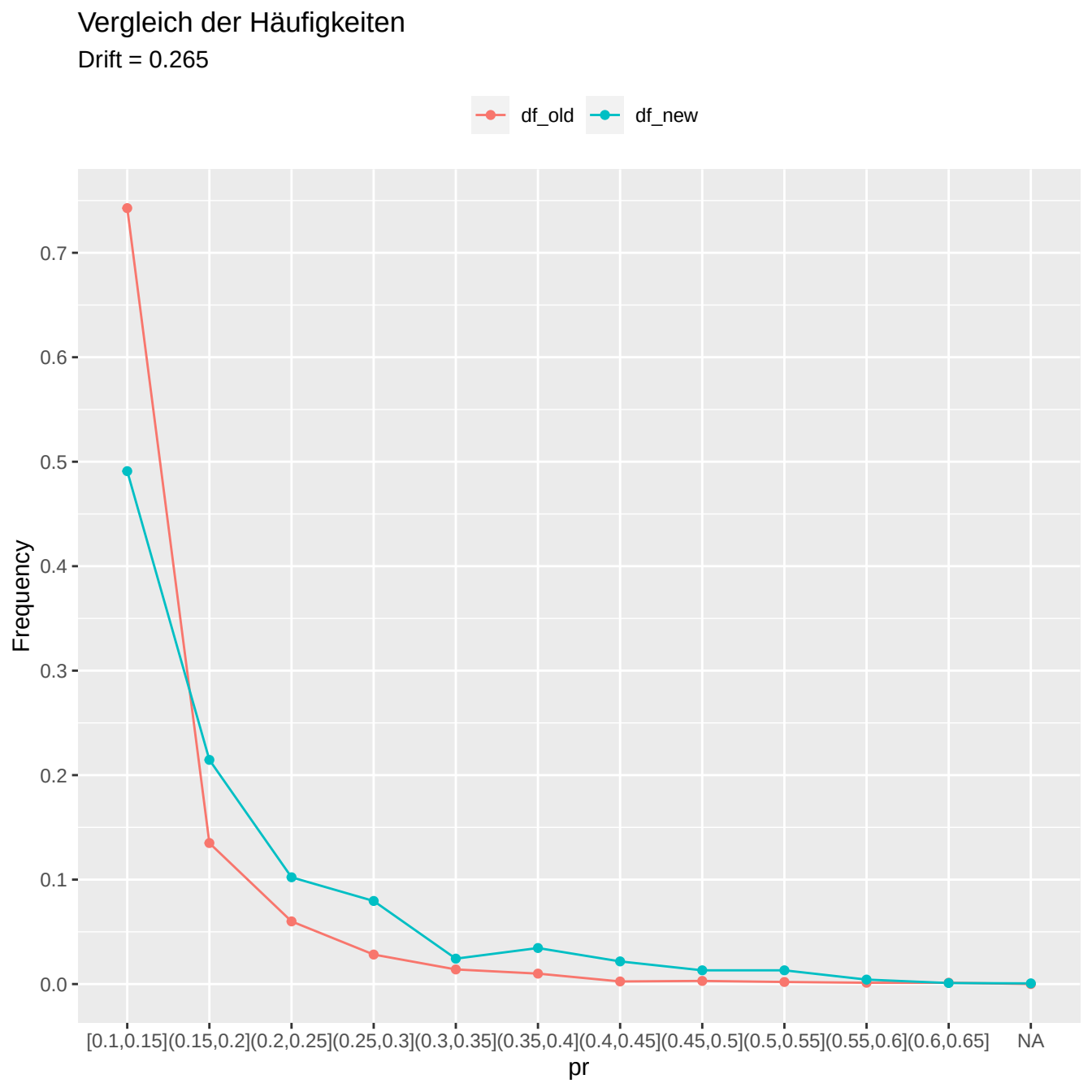

Predicted
Actual
Partial_Dependence
Value
0.15
0.10
0.05
0.00
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
Count
90
60
30
0
TotalCharges

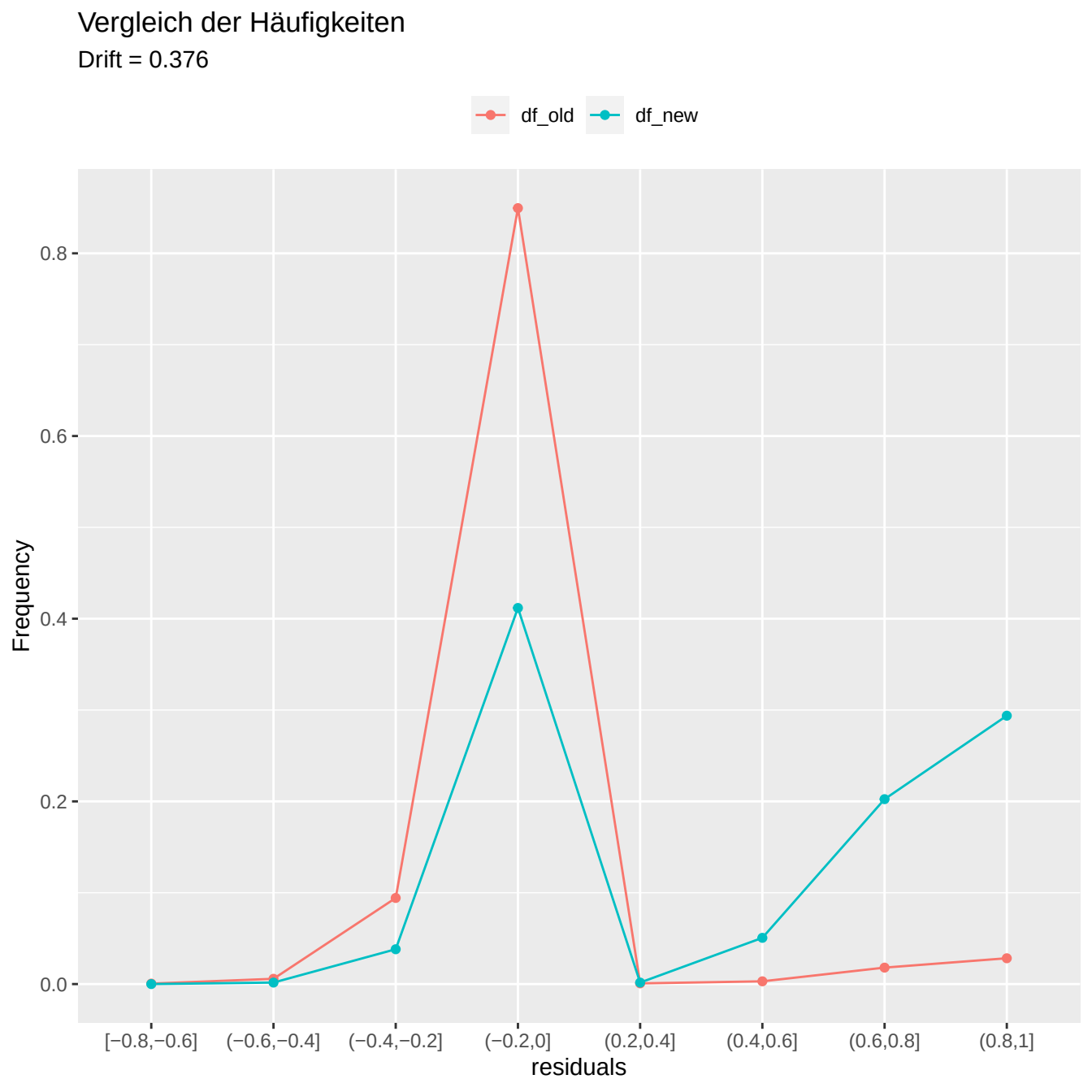

Predicted
Actual
Partial_Dependence
Value
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
50
60
70
Count
80
60
40
20
0
tenure

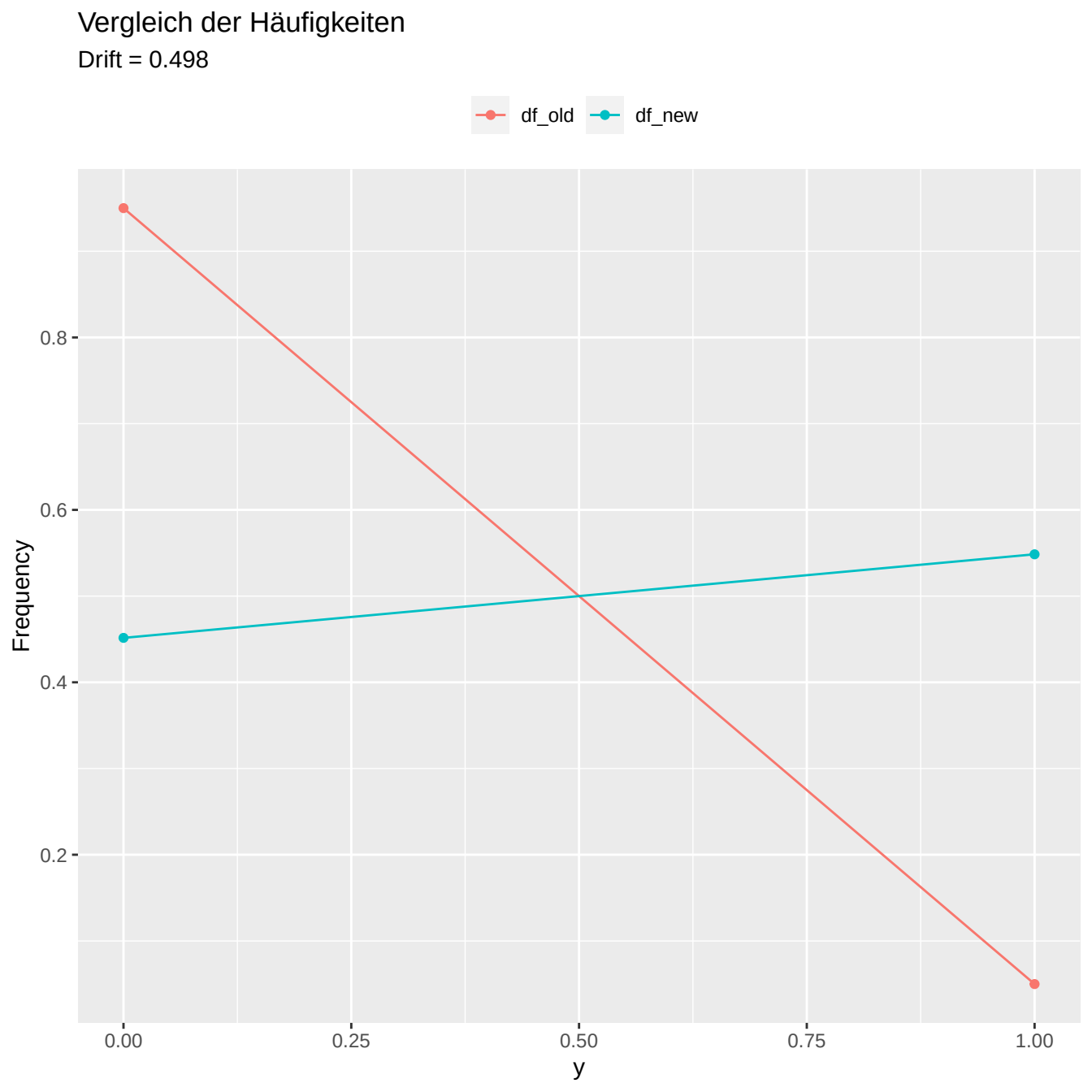

Predicted
Actual
Partial_Dependence
Value
0.10
0.08
0.06
0.04
0.02
0.00
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
300
200
100
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
TechSupport

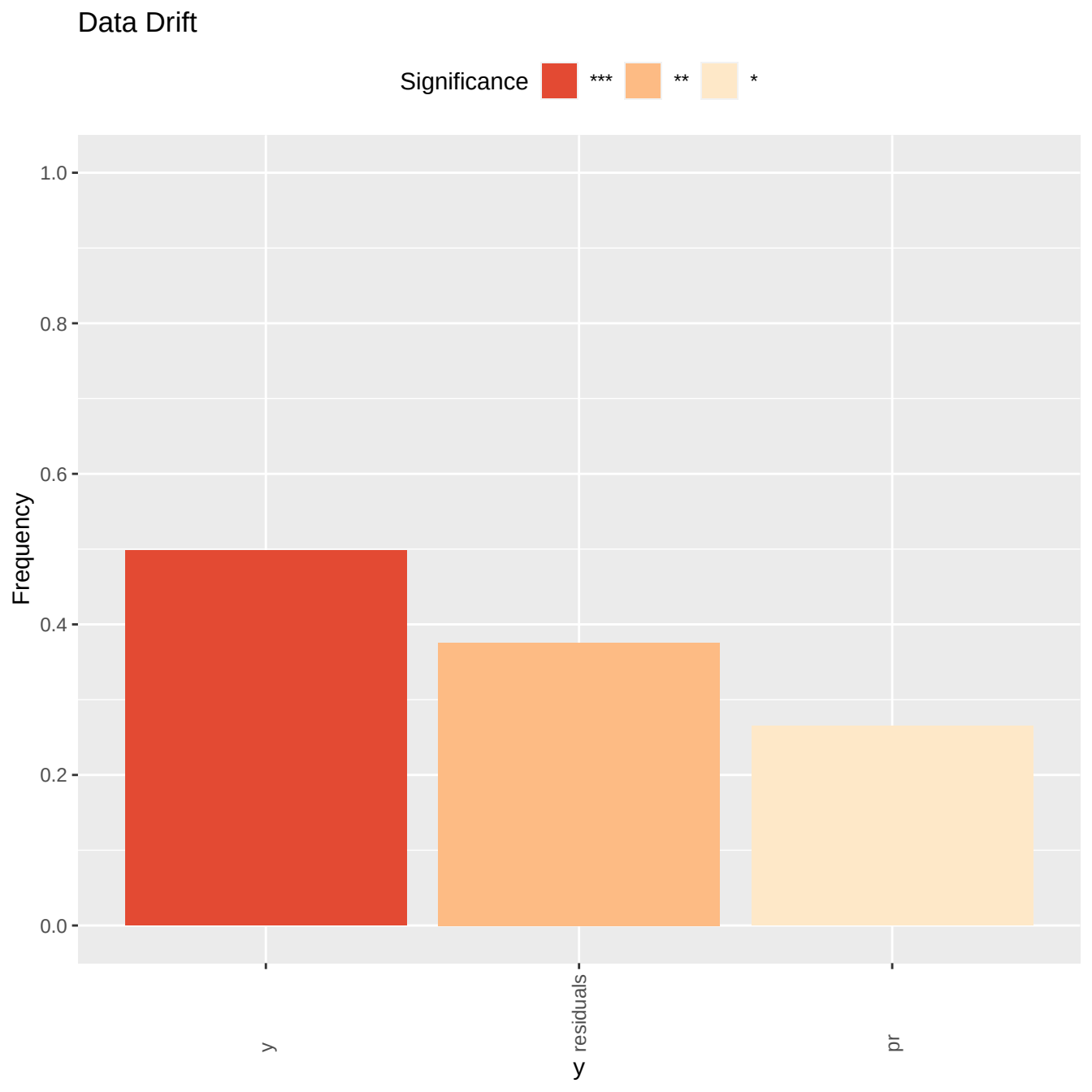

CV error
0.814
0.75
0.50
0.25
0.00
auc
0.950
0.975
1.000
1.025
1.05
Run

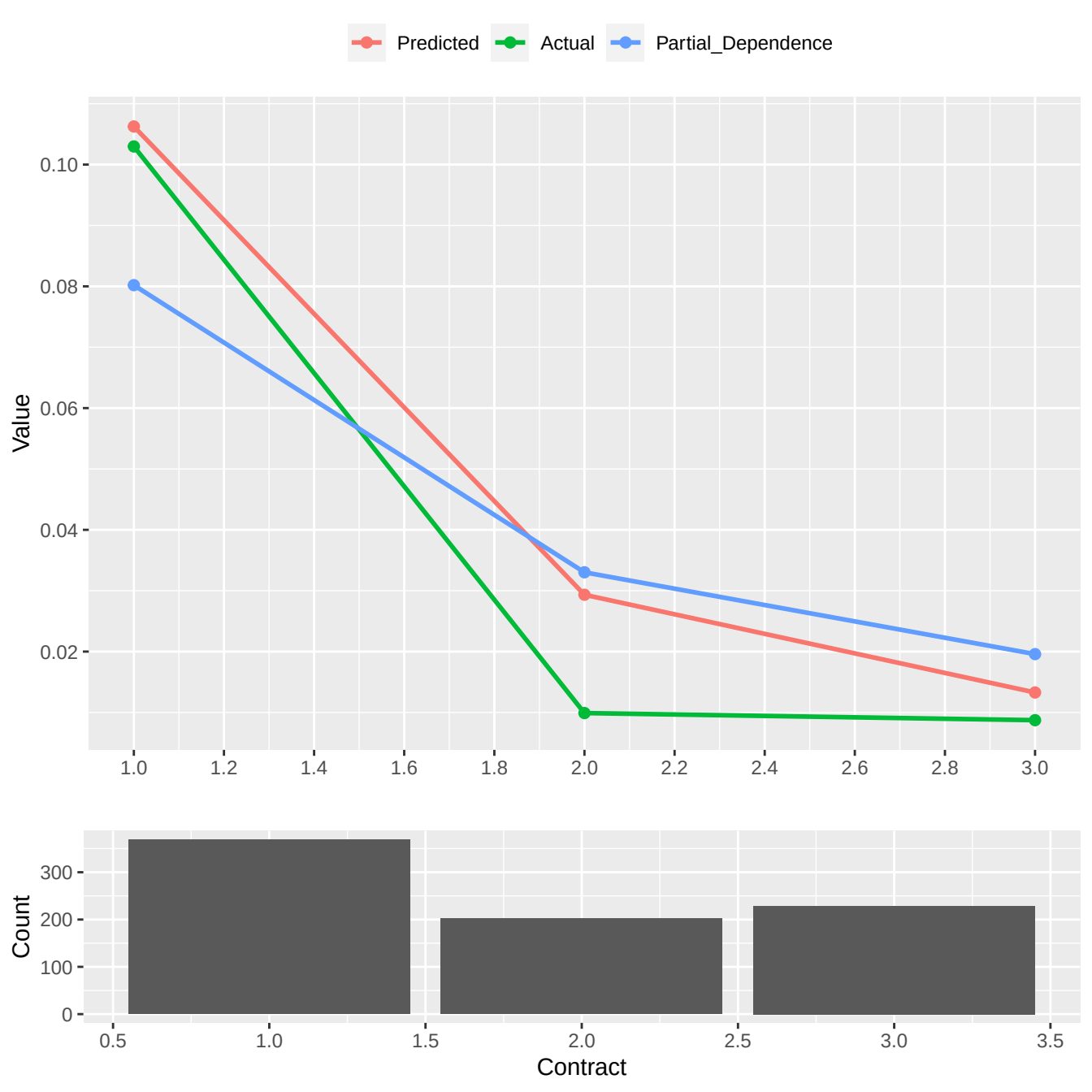

CV error
0.814
0.75
0.50
0.25
0.00
auc
1.00
1.25
1.50
1.75
2.00
Run

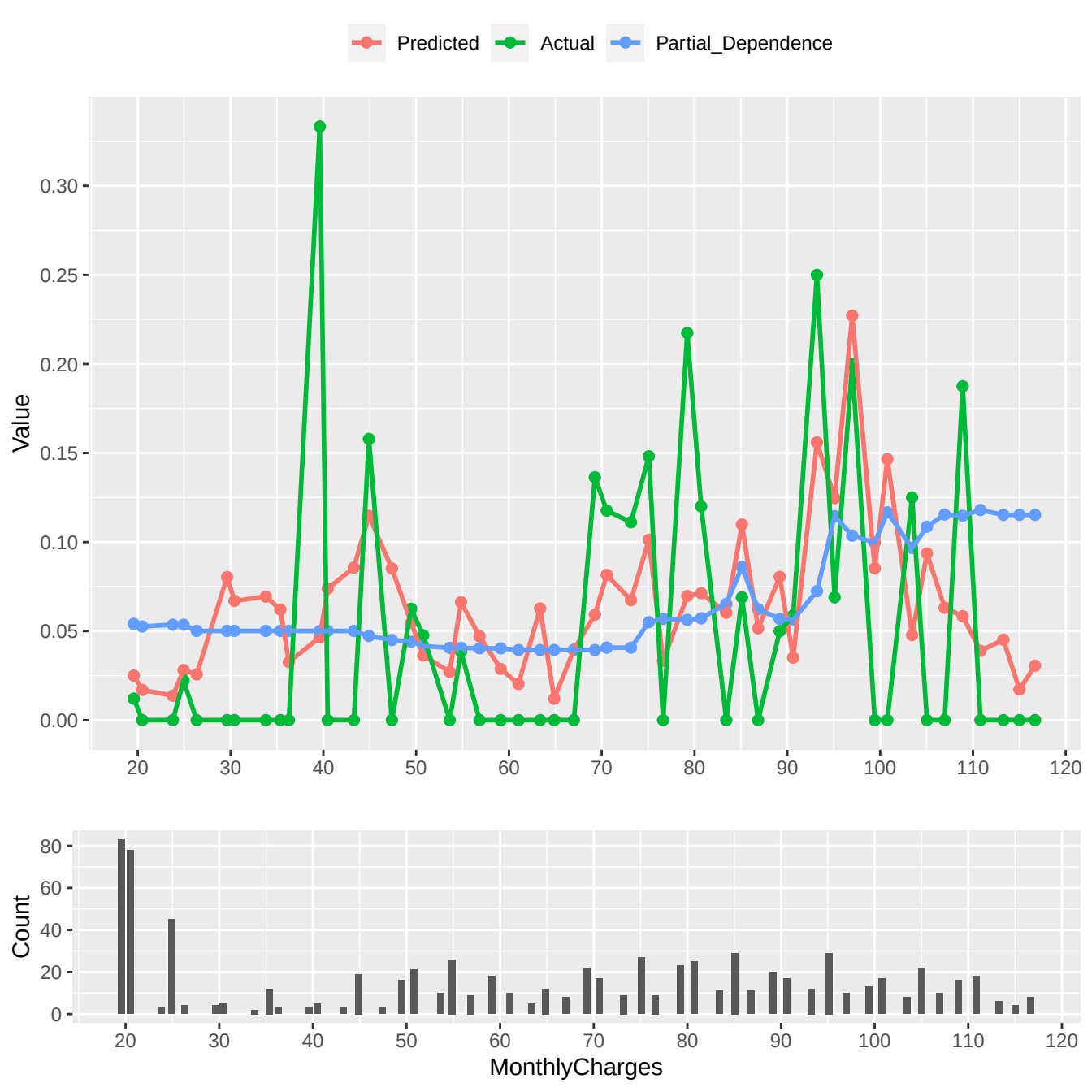

Vergleich der Häufigkeiten
Drift = 0.265
df_old
df_new
Frequency
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
[0.1,0.15]
(0.15,0.2]
(0.2,0.25]
(0.25,0.3]
(0.3,0.35]
(0.35,0.4]
(0.4,0.45]
(0.45,0.5]
(0.5,0.55]
(0.55,0.6]
(0.6,0.65]
NA
pr

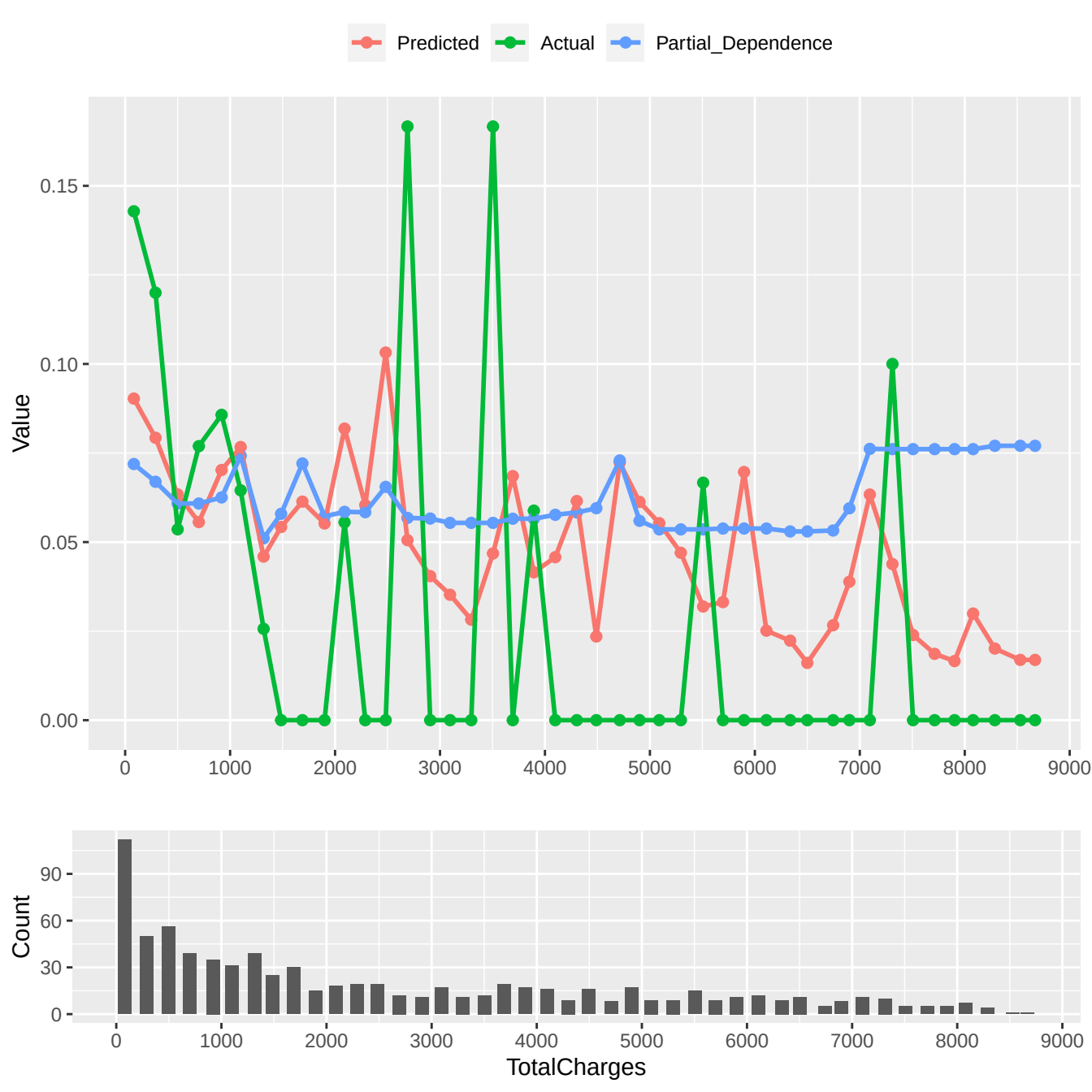

Vergleich der Häufigkeiten
Drift = 0.376
df_old
df_new
0.8
0.6
0.4
0.2
0.0
Frequency
[−0.8,−0.6]
(−0.6,−0.4]
(−0.4,−0.2]
(−0.2,0]
(0.2,0.4]
(0.4,0.6]
(0.6,0.8]
(0.8,1]
residuals


Vergleich der Häufigkeiten
Drift = 0.498
df_old
df_new
Frequency
0.8
0.6
0.4
0.2
0.00
0.25
0.50
0.75
1.00
y

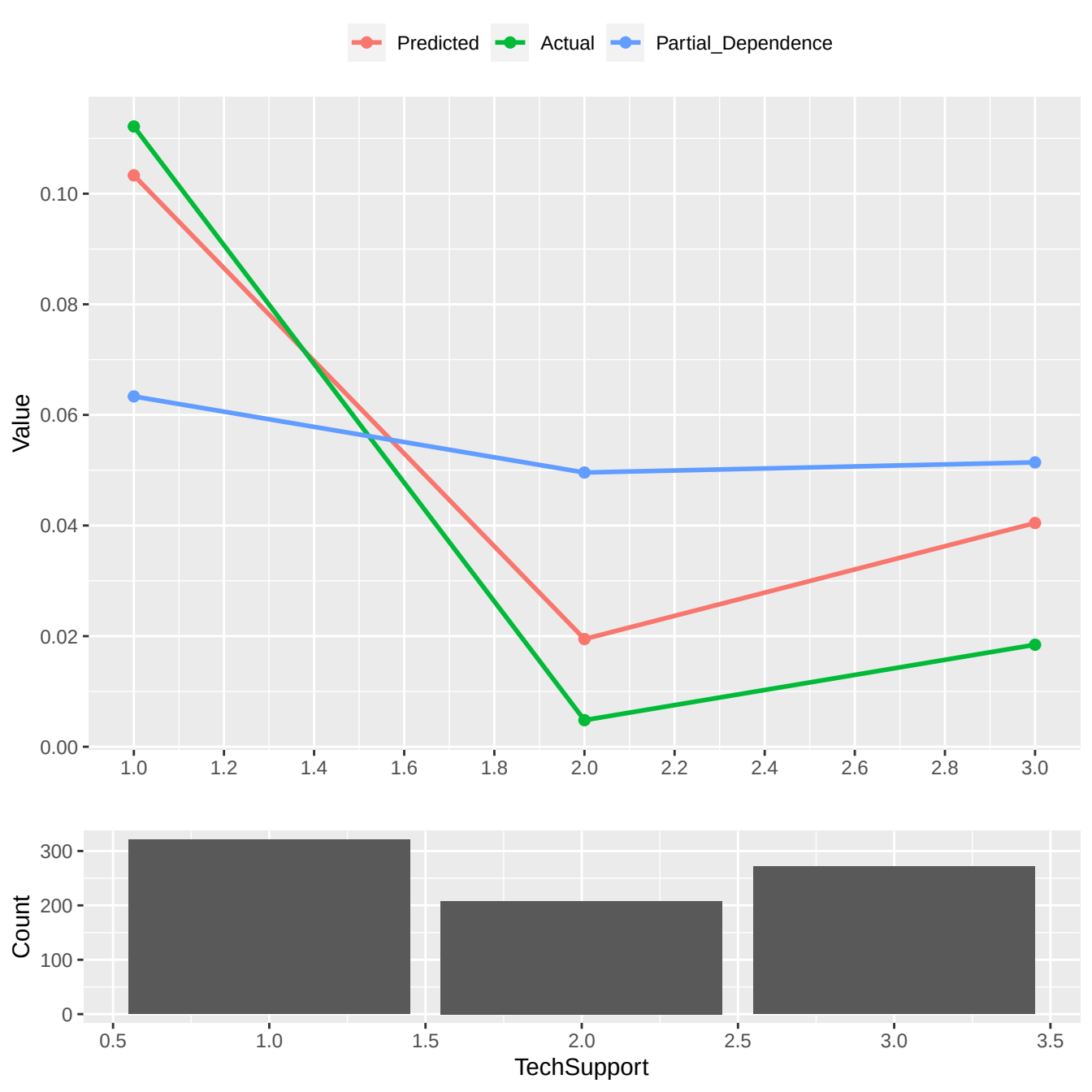

Data Drift
Significance
***
**
*
1.0
0.8
0.6
0.4
0.2
0.0
Frequency
y
residuals
pr
y

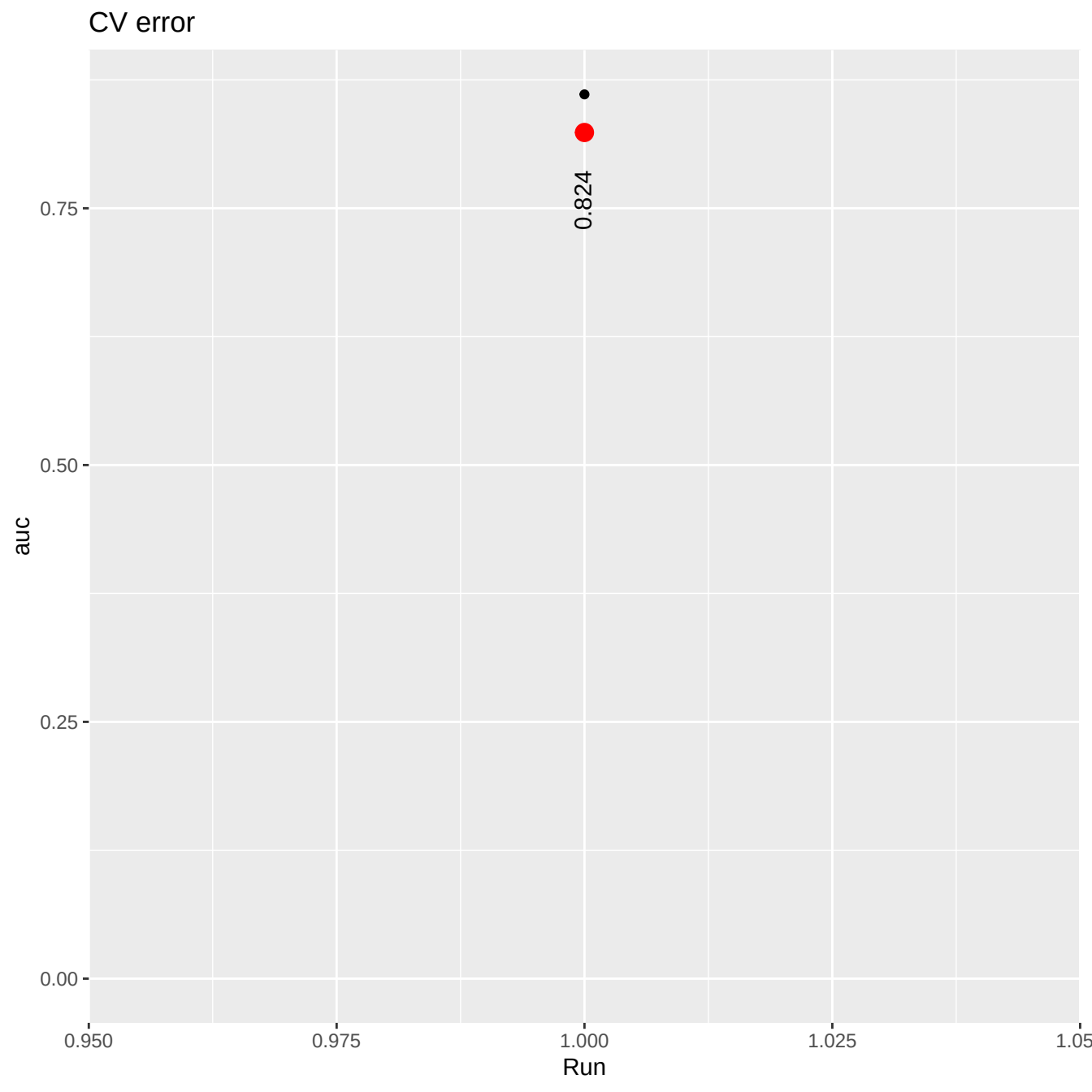

Predicted
Actual
Partial_Dependence
Value
0.10
0.08
0.06
0.04
0.02
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
300
200
100
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Contract

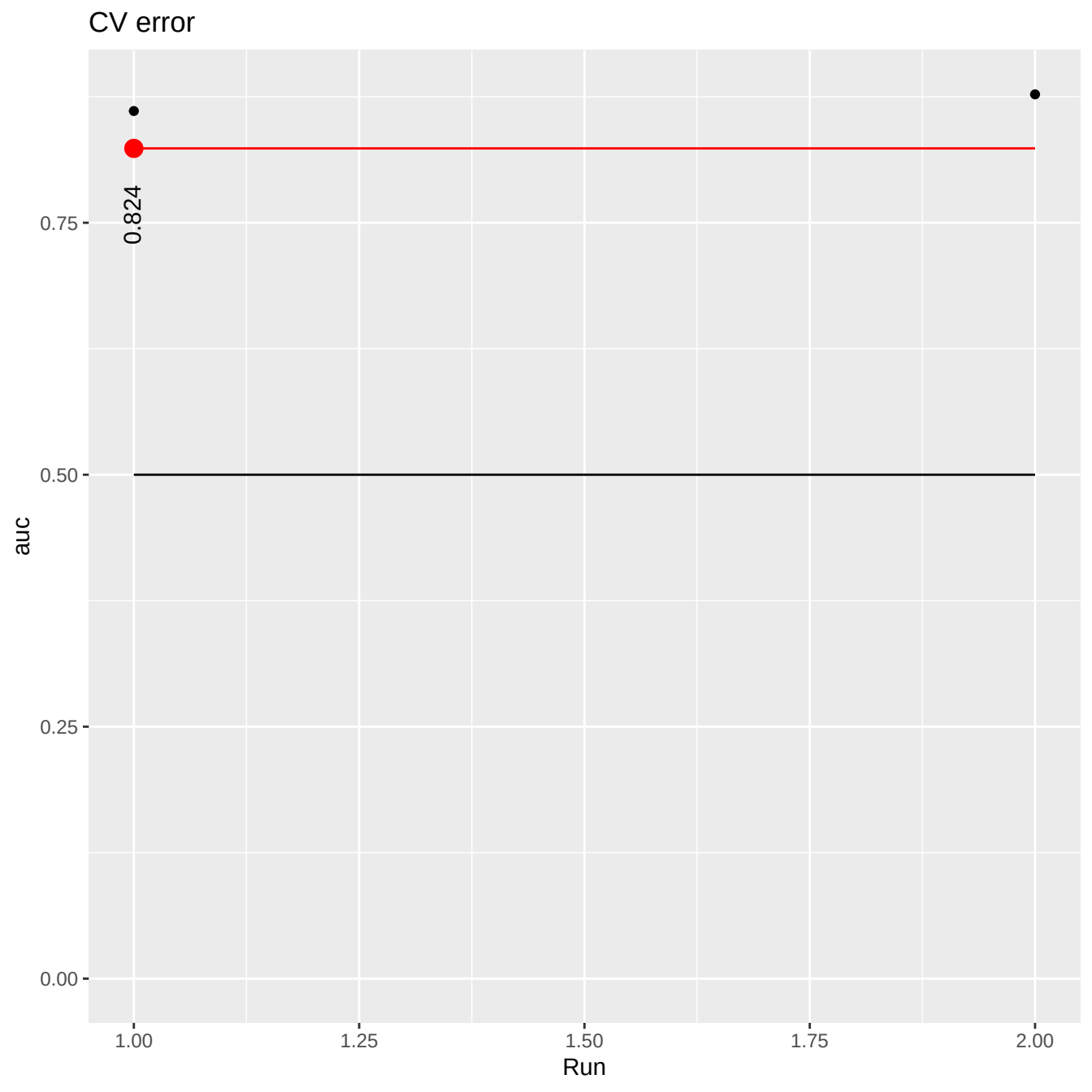

Predicted
Actual
Partial_Dependence
Value
0.30
0.25
0.20
0.15
0.10
0.05
0.00
20
30
40
50
60
70
80
90
100
110
120
Count
80
60
40
20
0
MonthlyCharges

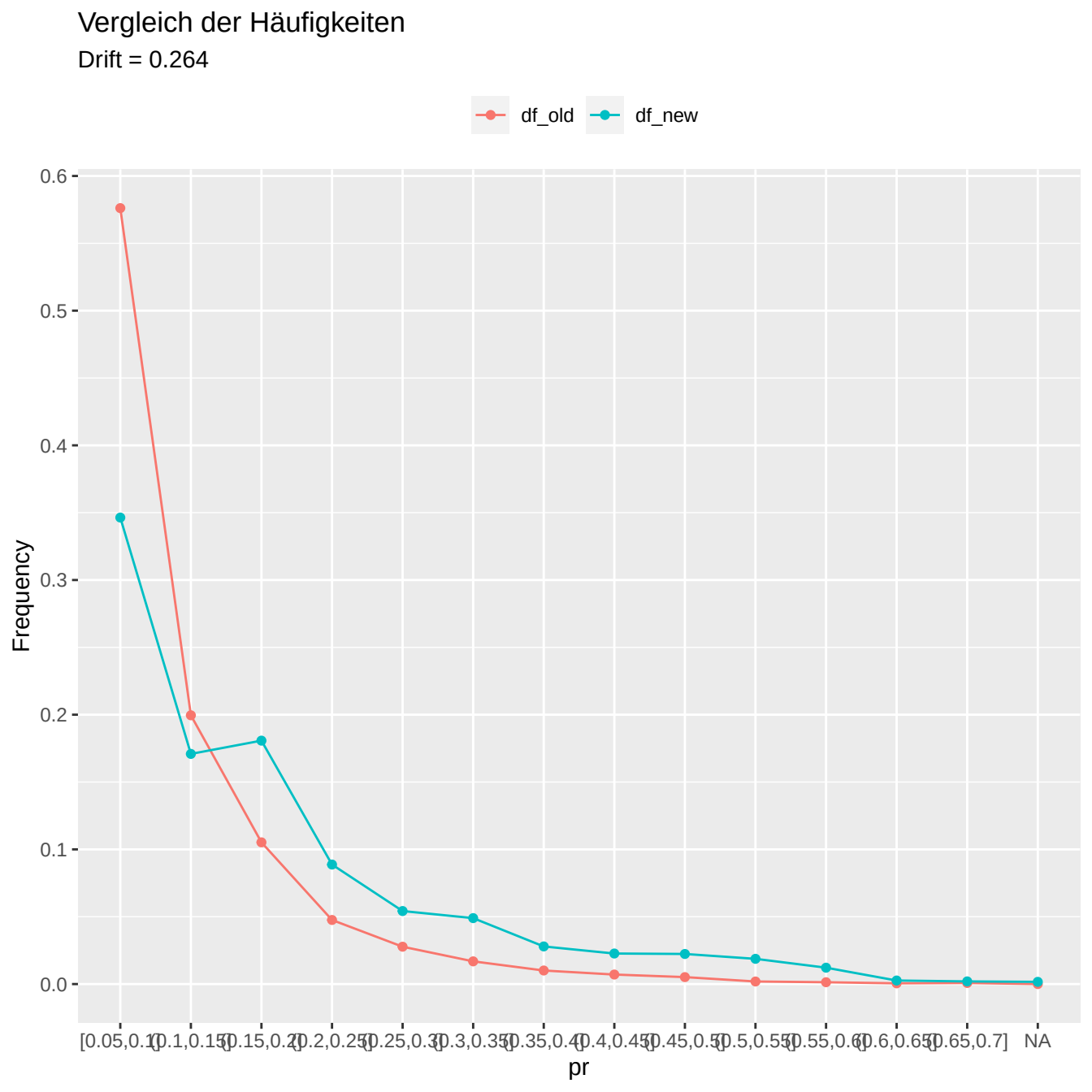

Predicted
Actual
Partial_Dependence
Value
0.15
0.10
0.05
0.00
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
Count
90
60
30
0
TotalCharges

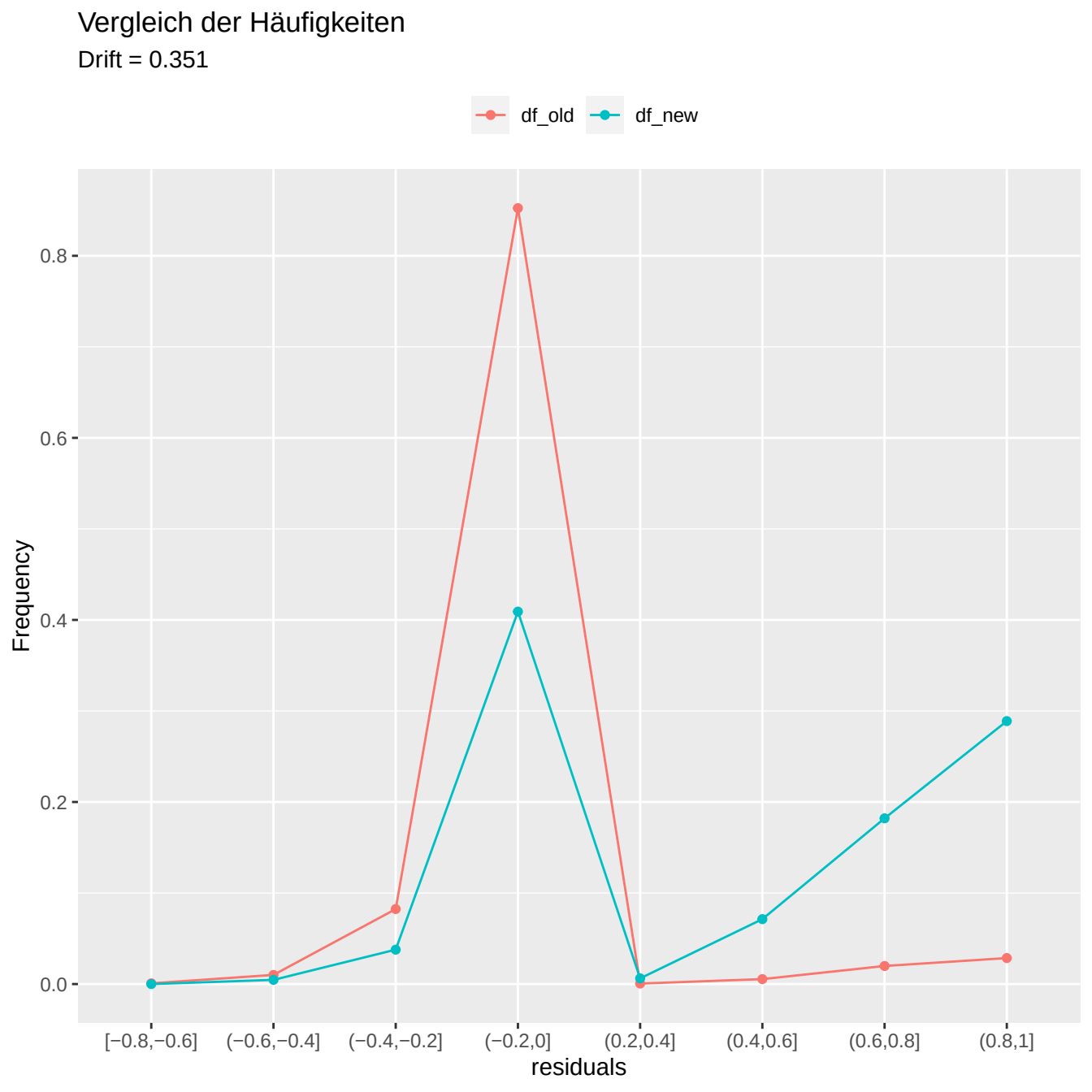

Predicted
Actual
Partial_Dependence
Value
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
50
60
70
Count
80
60
40
20
0
tenure

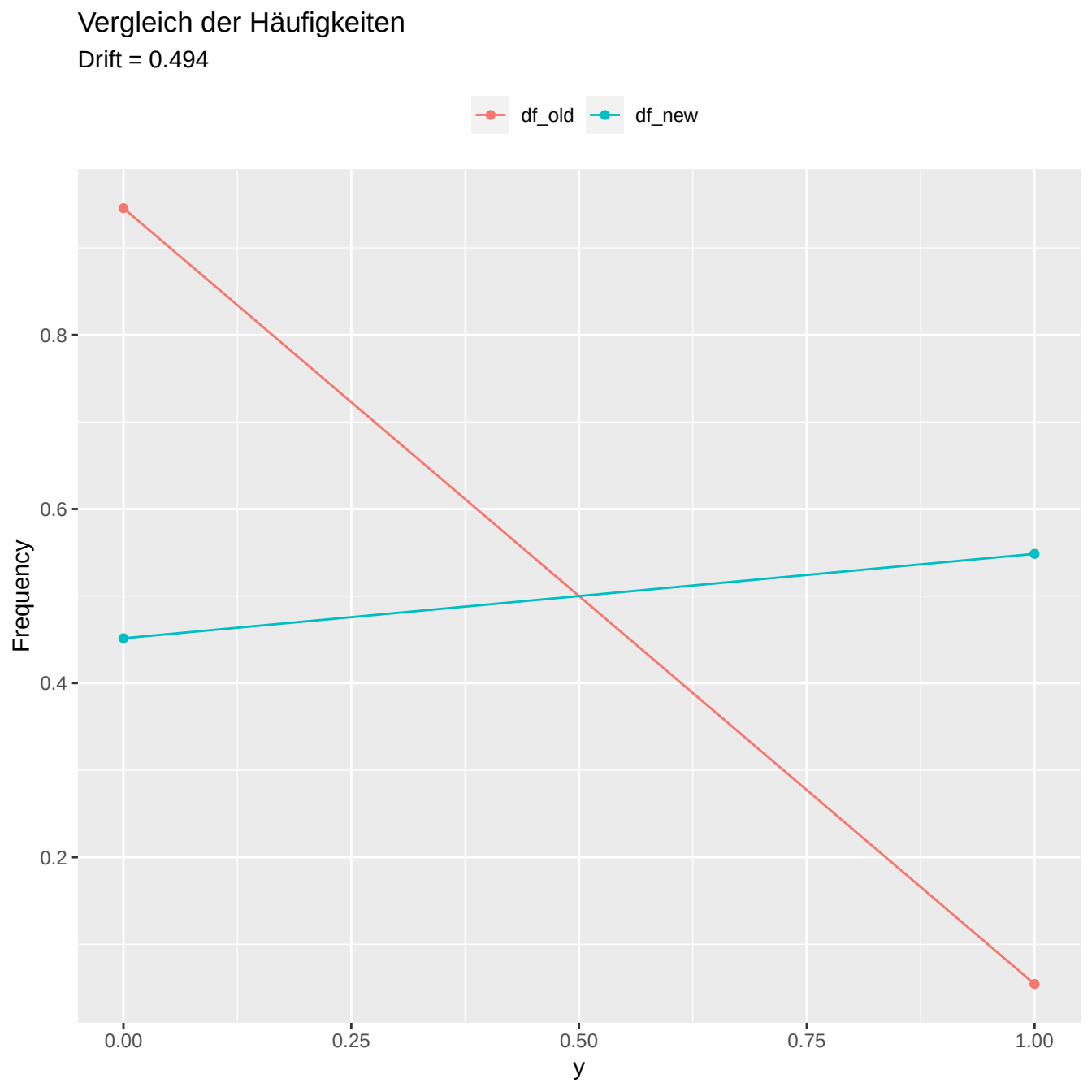

Predicted
Actual
Partial_Dependence
Value
0.10
0.08
0.06
0.04
0.02
0.00
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
300
200
100
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
TechSupport

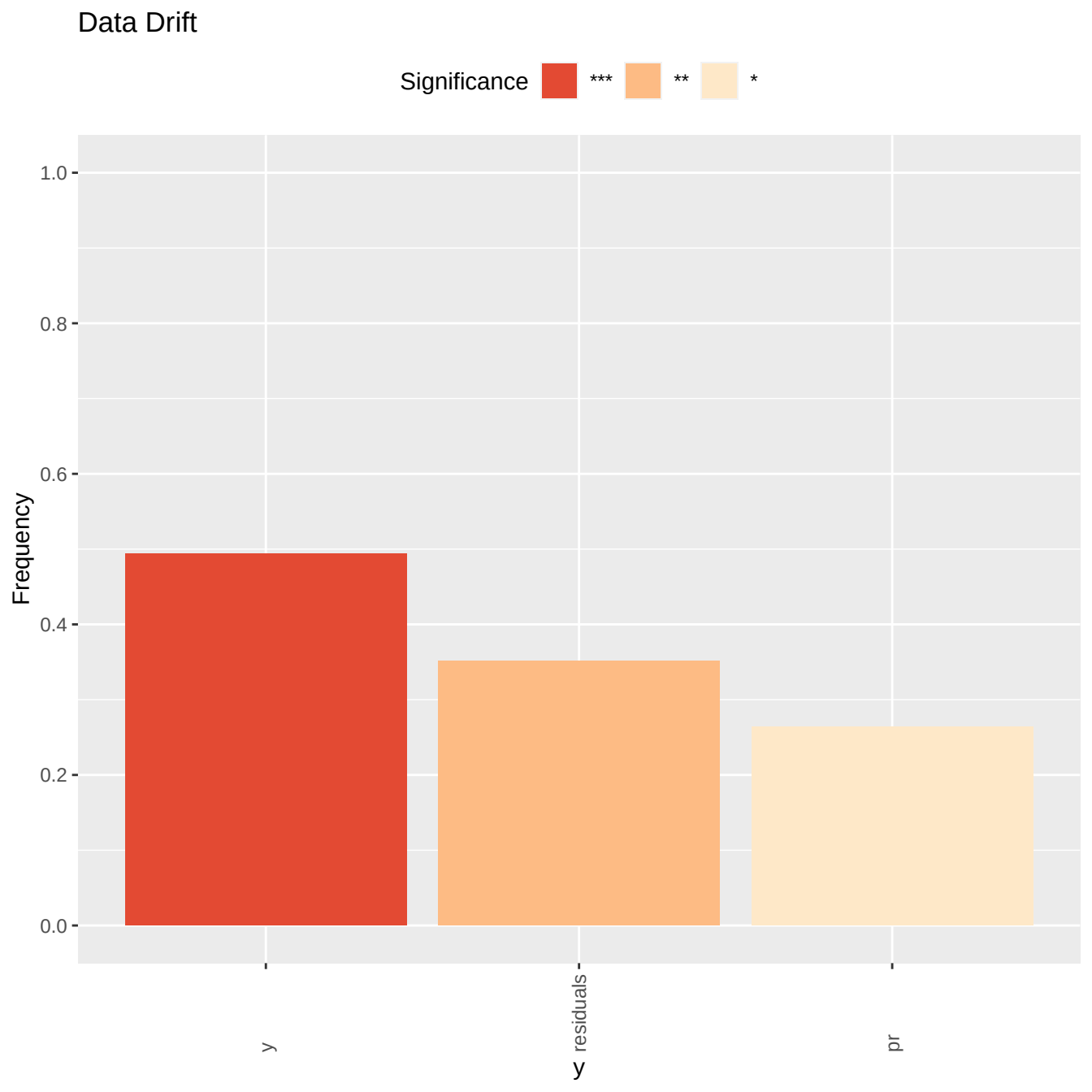

CV error
0.824
0.75
0.50
0.25
0.00
auc
0.950
0.975
1.000
1.025
1.05
Run

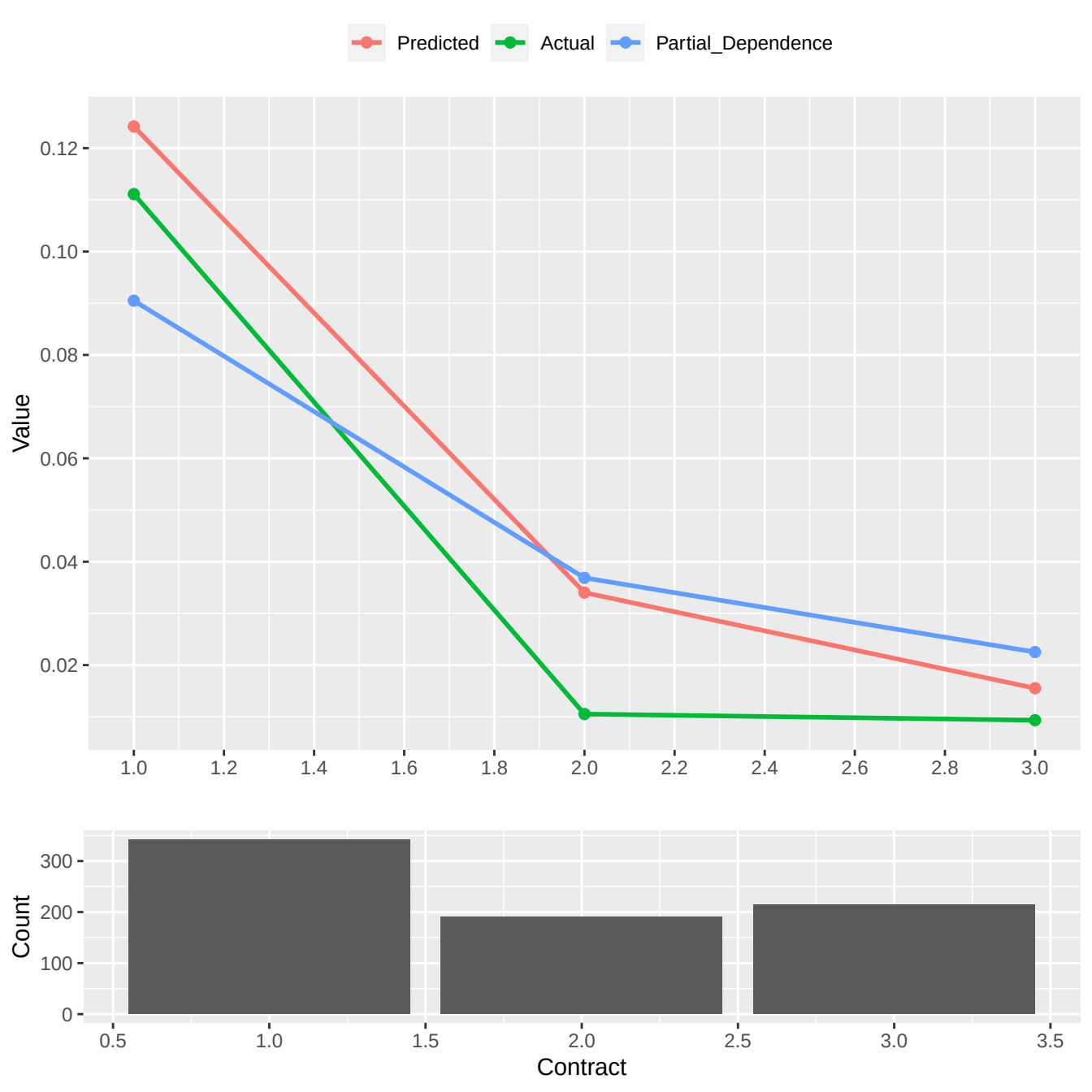

CV error
0.824
0.75
0.50
0.25
0.00
auc
1.00
1.25
1.50
1.75
2.00
Run


Vergleich der Häufigkeiten
Drift = 0.264
df_old
df_new
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Frequency
[0.05,0.1]
(0.1,0.15]
(0.15,0.2]
(0.2,0.25]
(0.25,0.3]
(0.3,0.35]
(0.35,0.4]
(0.4,0.45]
(0.45,0.5]
(0.5,0.55]
(0.55,0.6]
(0.6,0.65]
(0.65,0.7]
NA
pr


Vergleich der Häufigkeiten
Drift = 0.351
df_old
df_new
0.8
0.6
0.4
0.2
0.0
Frequency
[−0.8,−0.6]
(−0.6,−0.4]
(−0.4,−0.2]
(−0.2,0]
(0.2,0.4]
(0.4,0.6]
(0.6,0.8]
(0.8,1]
residuals


Vergleich der Häufigkeiten
Drift = 0.494
df_old
df_new
Frequency
0.8
0.6
0.4
0.2
0.00
0.25
0.50
0.75
1.00
y

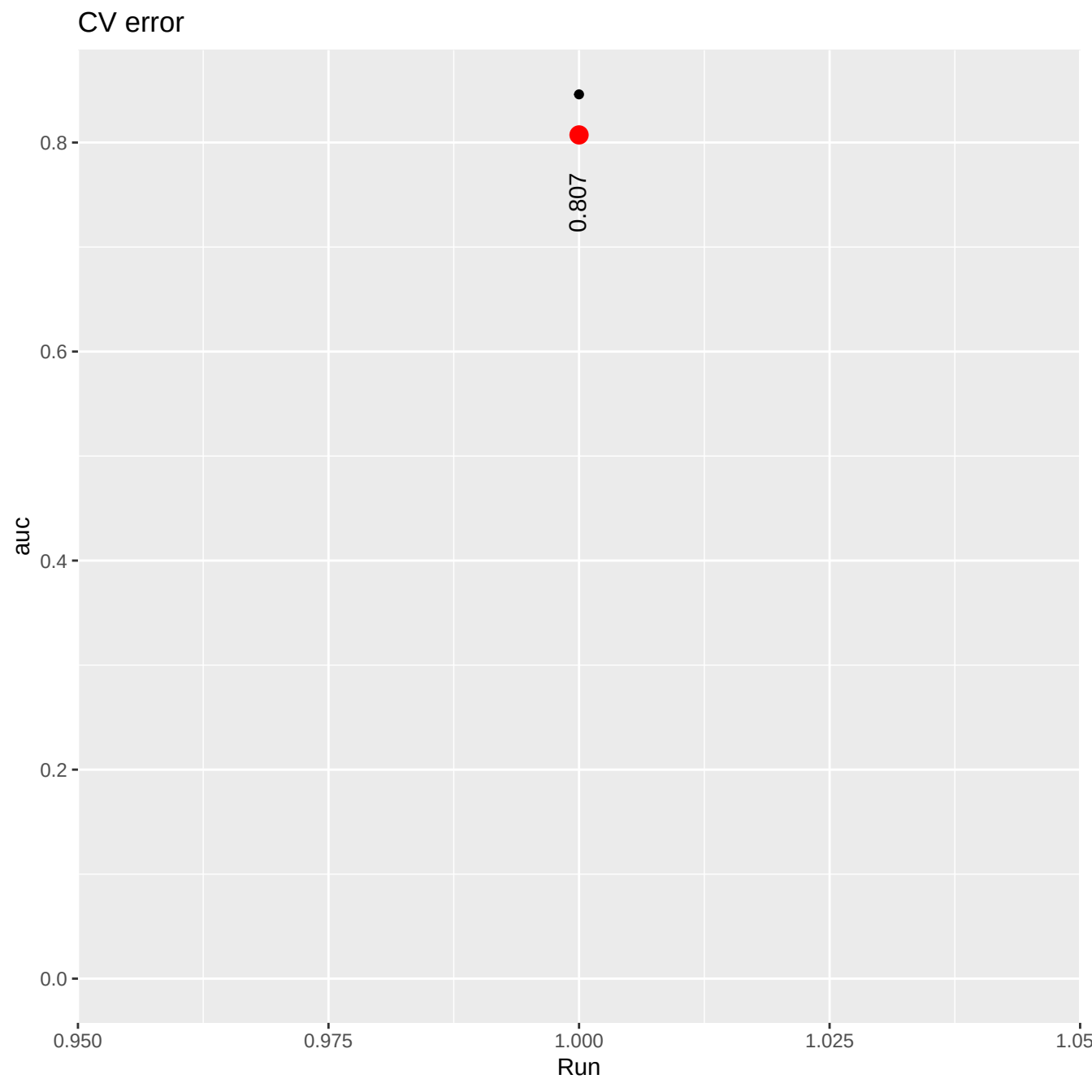

Data Drift
Significance
***
**
*
1.0
0.8
0.6
0.4
0.2
0.0
Frequency
y
residuals
pr
y

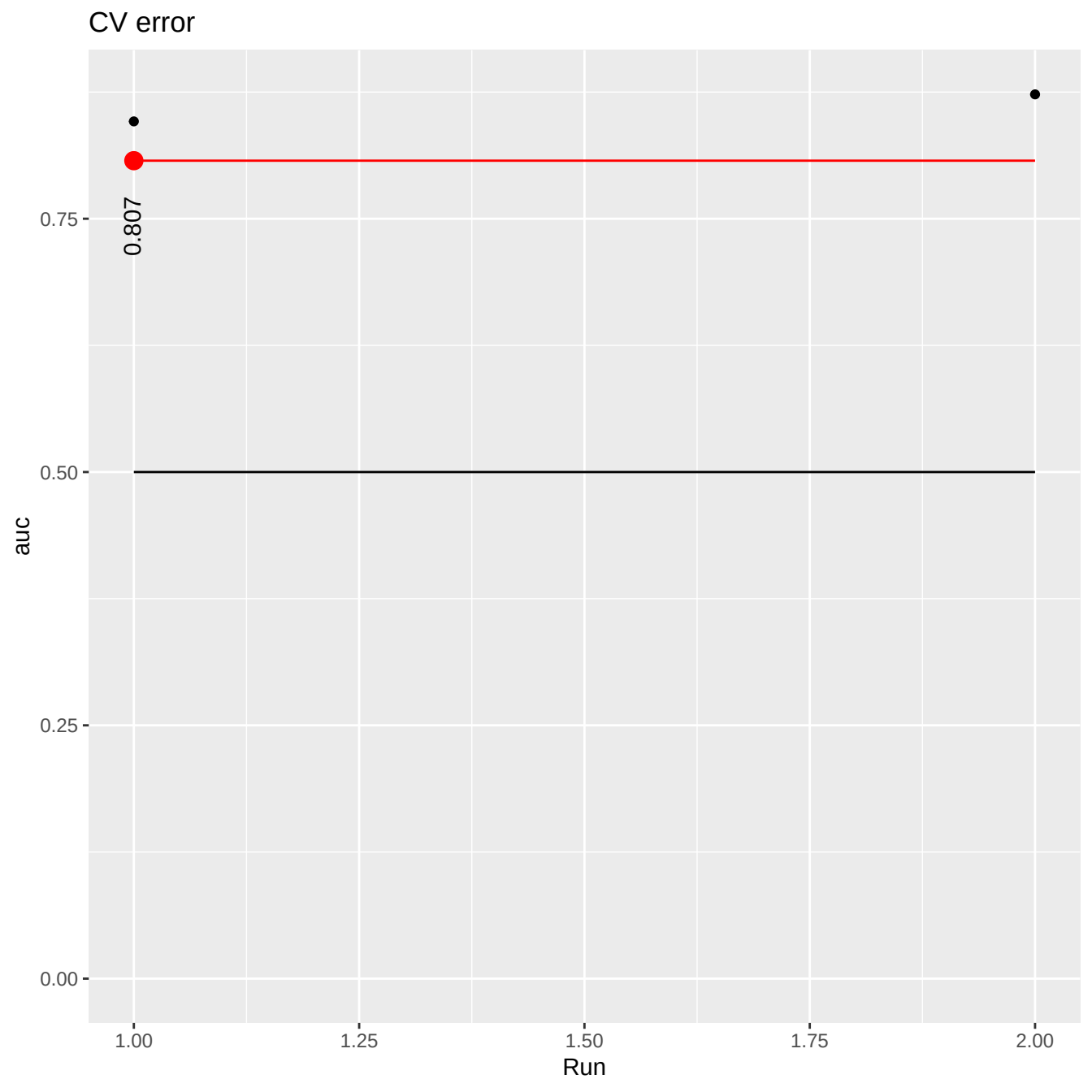

Predicted
Actual
Partial_Dependence
Value
0.12
0.10
0.08
0.06
0.04
0.02
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
300
200
100
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Contract

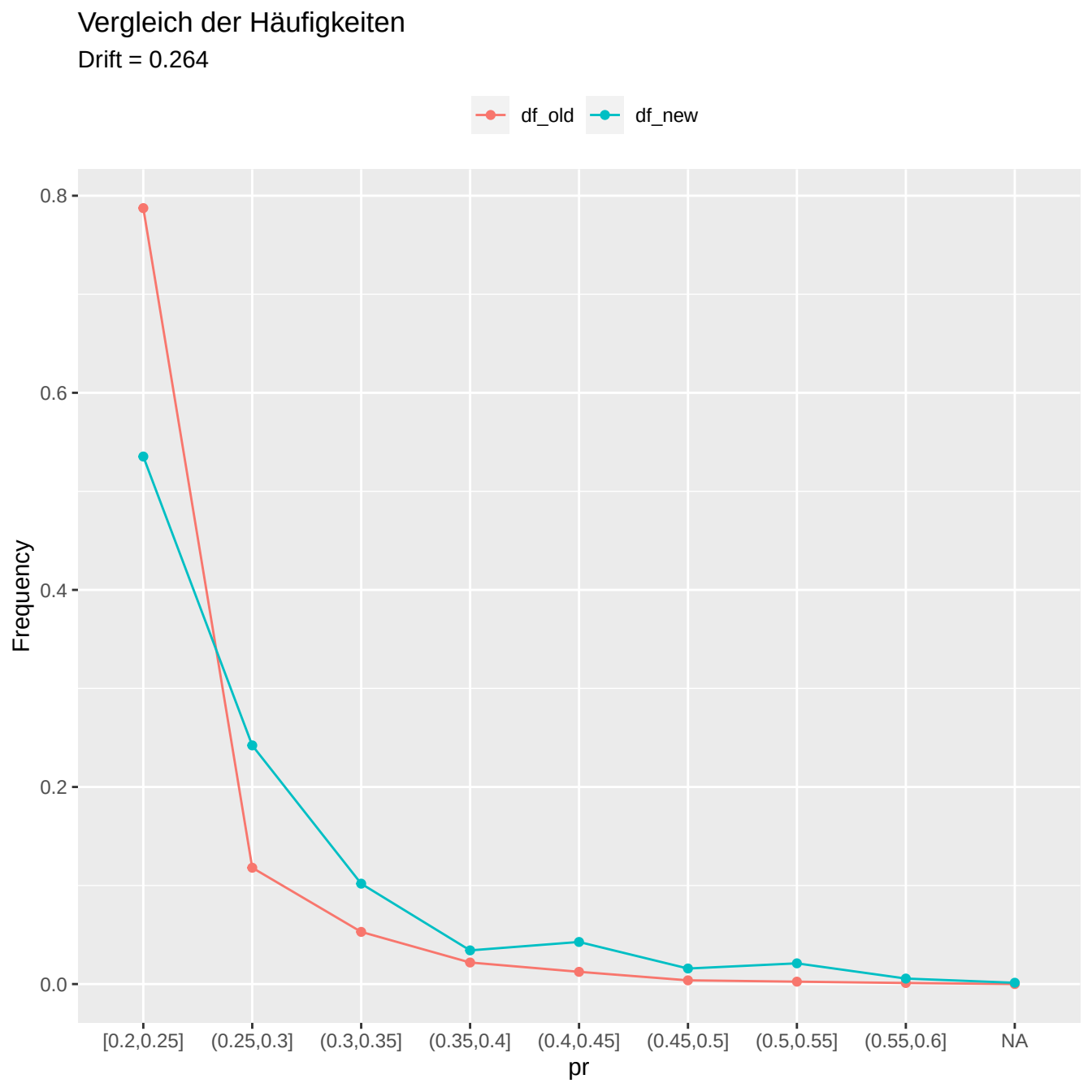

Predicted
Actual
Partial_Dependence
Value
0.30
0.25
0.20
0.15
0.10
0.05
0.00
20
30
40
50
60
70
80
90
100
110
120
Count
60
40
20
0
MonthlyCharges

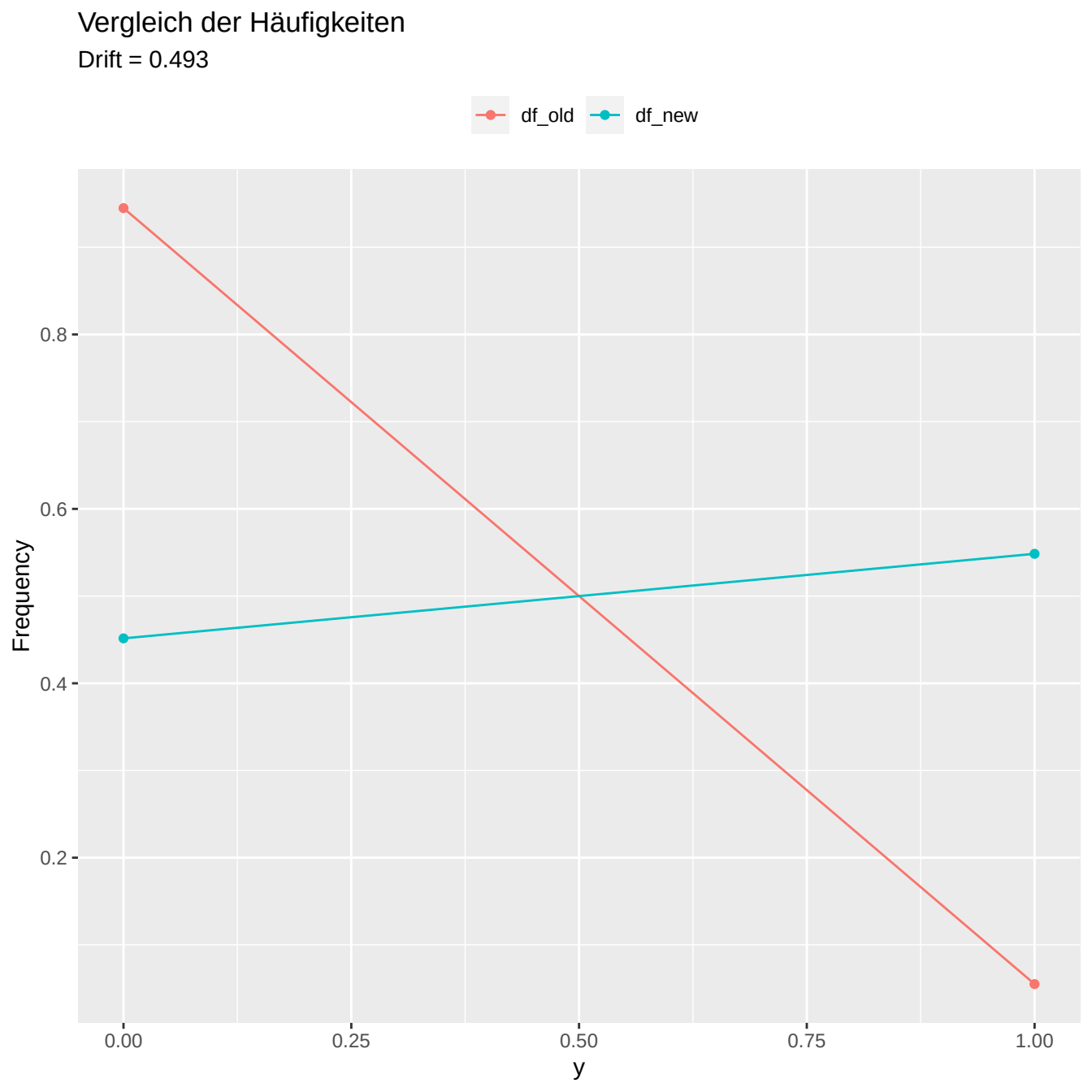

Predicted
Actual
Partial_Dependence
Value
0.20
0.15
0.10
0.05
0.00
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
Count
100
75
50
25
0
TotalCharges

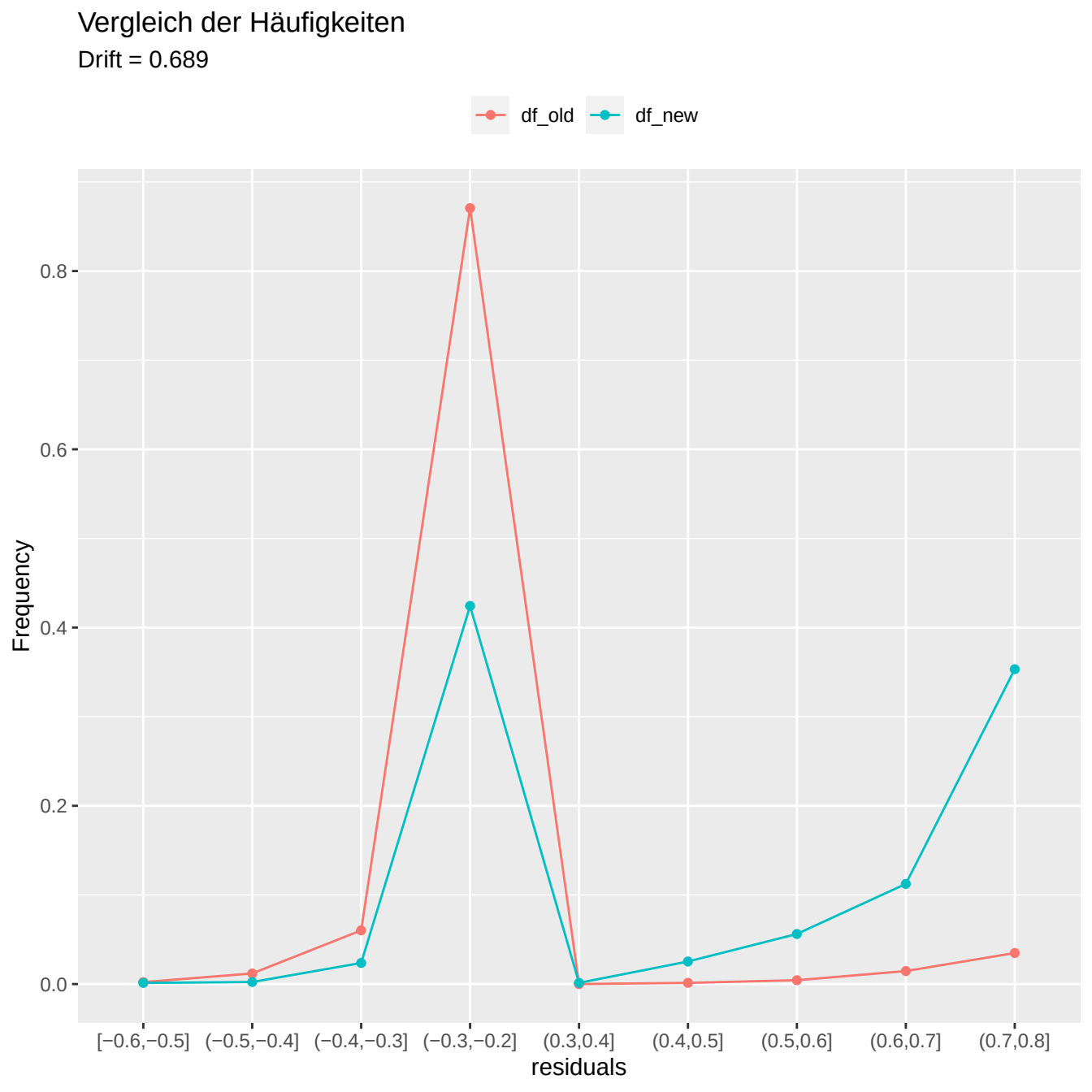

Predicted
Actual
Partial_Dependence
Value
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
50
60
70
Count
80
60
40
20
0
tenure

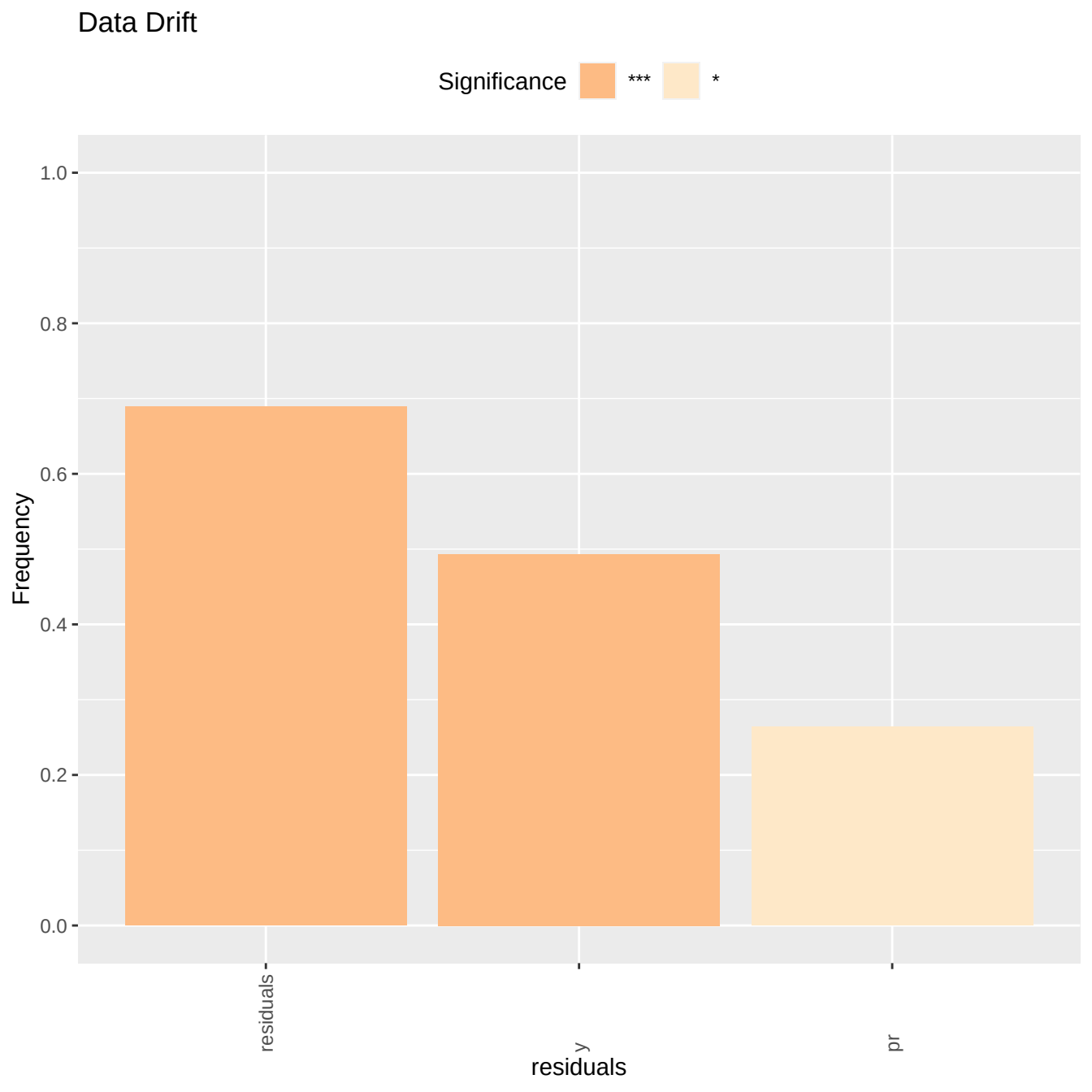

CV error
0.807
auc
0.0
0.2
0.4
0.6
0.8
0.950
0.975
1.000
1.025
1.05
Run

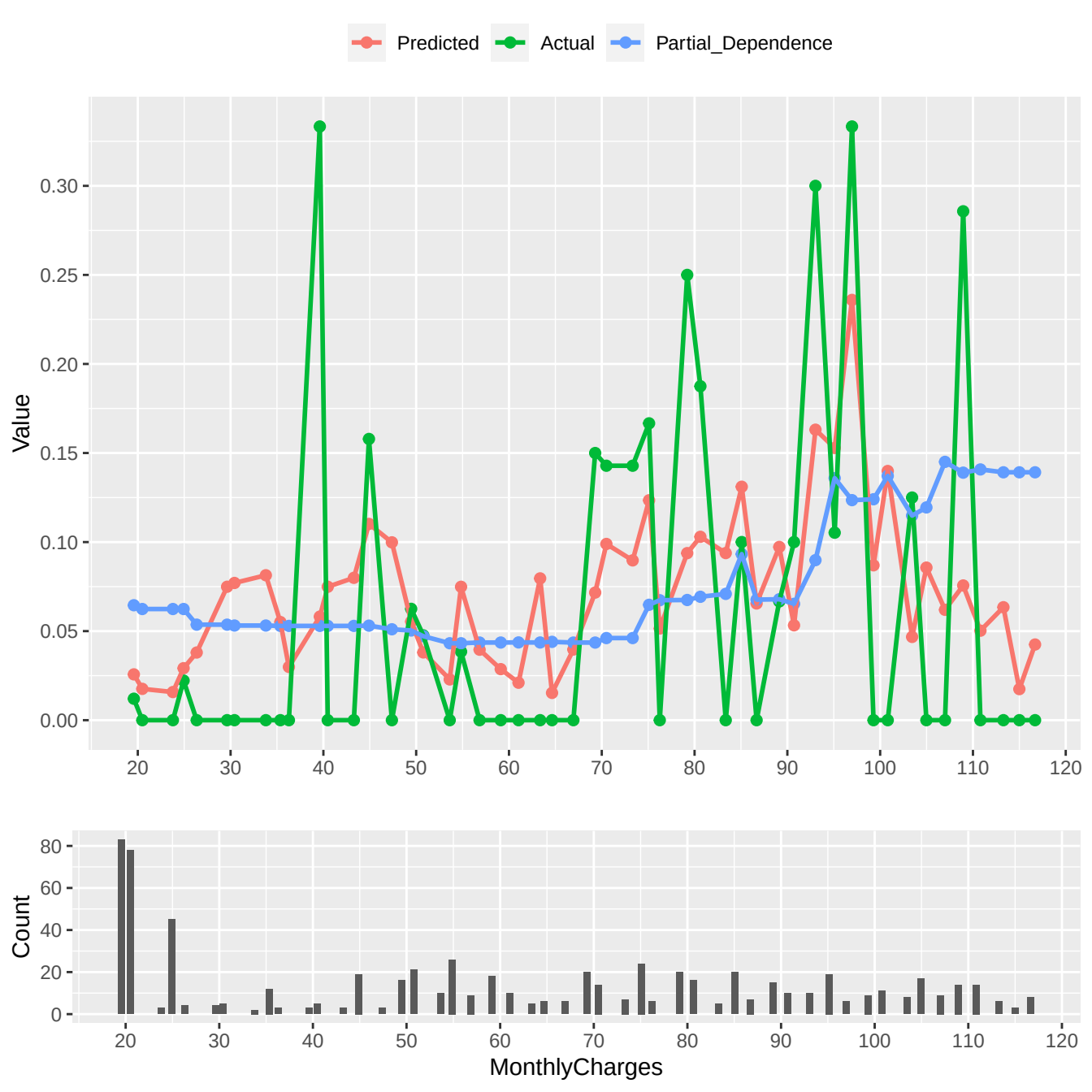

CV error
0.807
0.75
0.50
0.25
0.00
auc
1.00
1.25
1.50
1.75
2.00
Run

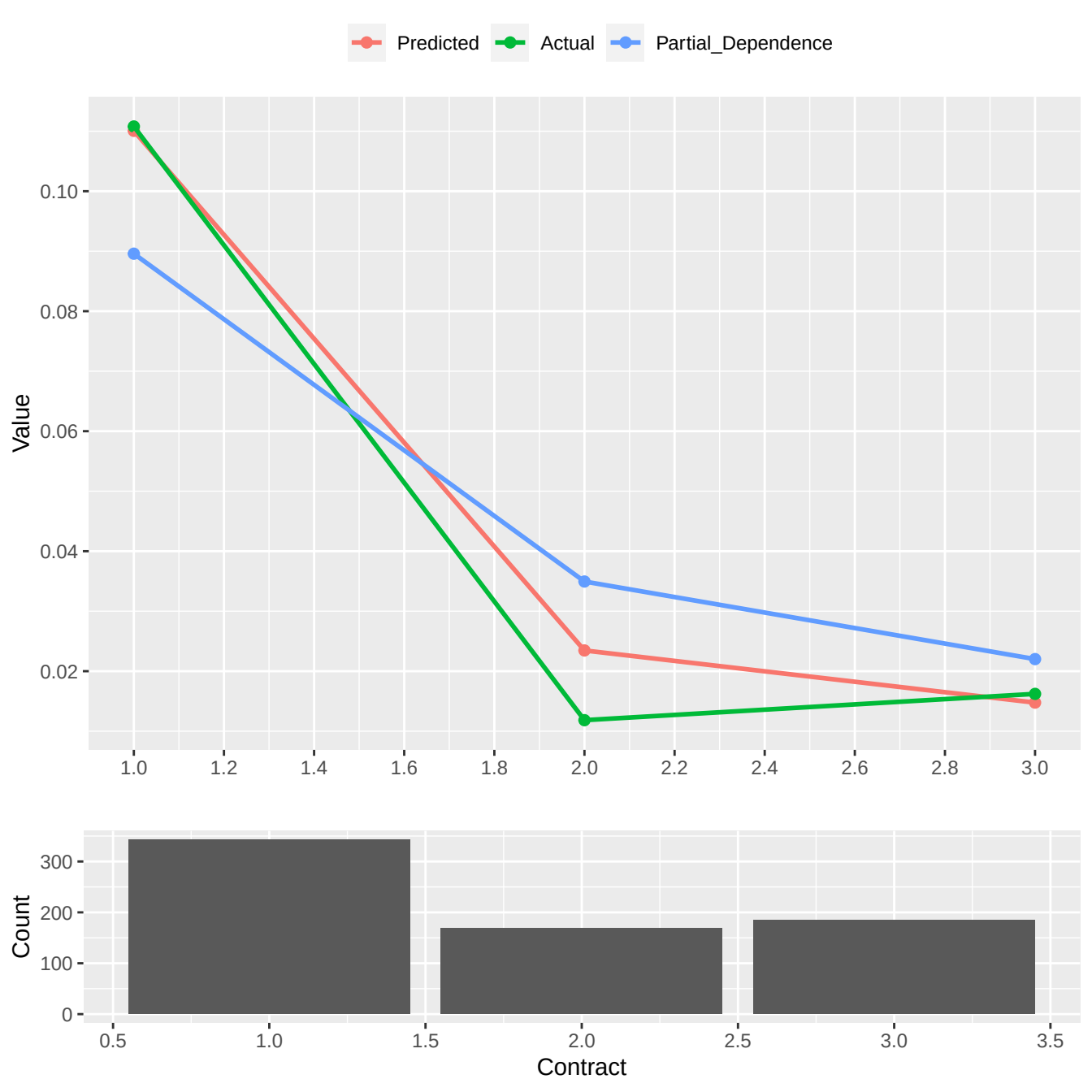

Vergleich der Häufigkeiten
Drift = 0.264
df_old
df_new
0.8
0.6
0.4
0.2
0.0
Frequency
[0.2,0.25]
(0.25,0.3]
(0.3,0.35]
(0.35,0.4]
(0.4,0.45]
(0.45,0.5]
(0.5,0.55]
(0.55,0.6]
NA
pr

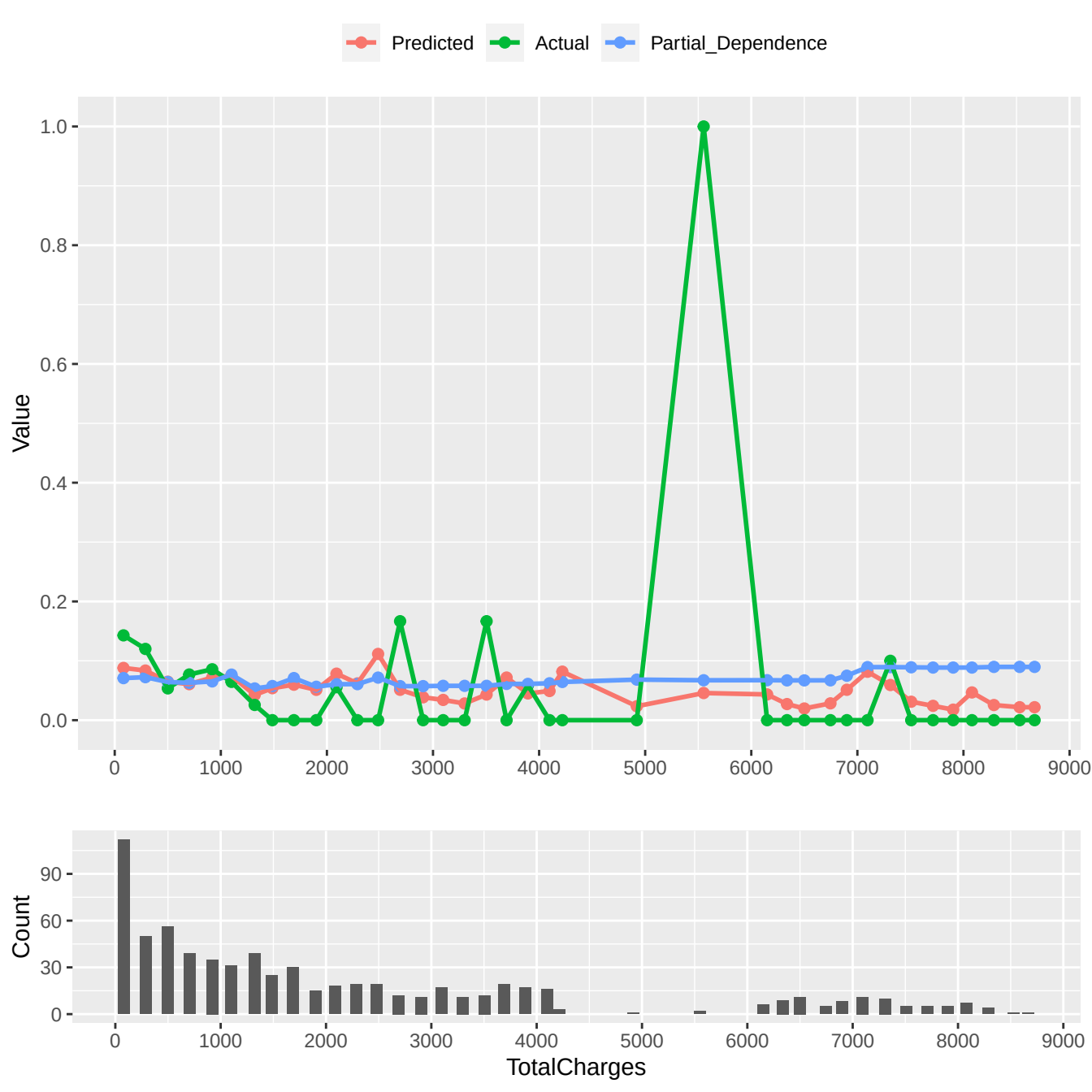

Vergleich der Häufigkeiten
Drift = 0.493
df_old
df_new
Frequency
0.8
0.6
0.4
0.2
0.00
0.25
0.50
0.75
1.00
y


Vergleich der Häufigkeiten
Drift = 0.689
df_old
df_new
0.8
0.6
0.4
0.2
0.0
Frequency
[−0.6,−0.5]
(−0.5,−0.4]
(−0.4,−0.3]
(−0.3,−0.2]
(0.3,0.4]
(0.4,0.5]
(0.5,0.6]
(0.6,0.7]
(0.7,0.8]
residuals

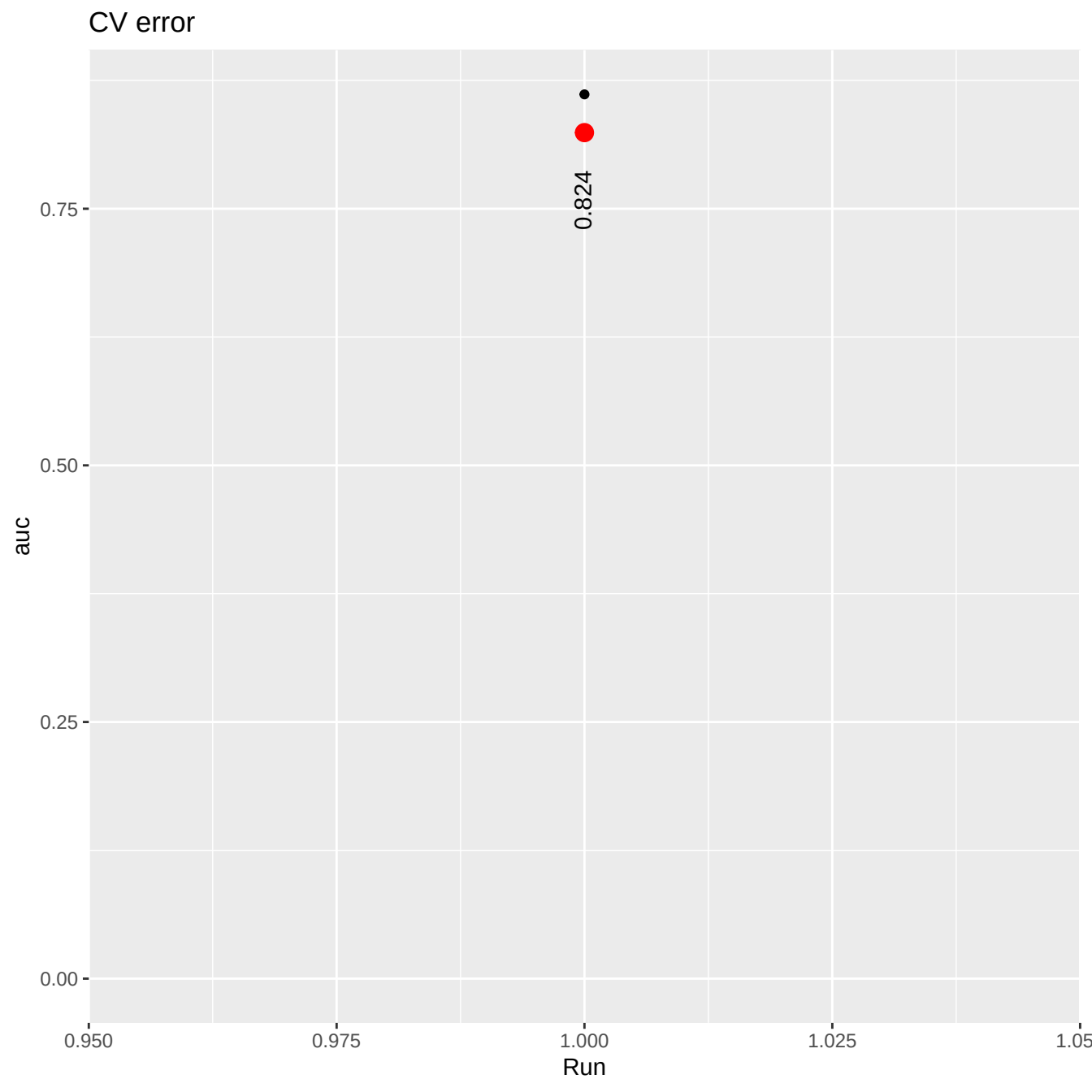

Data Drift
Significance
***
*
1.0
0.8
0.6
0.4
0.2
0.0
Frequency
residuals
y
pr
residuals

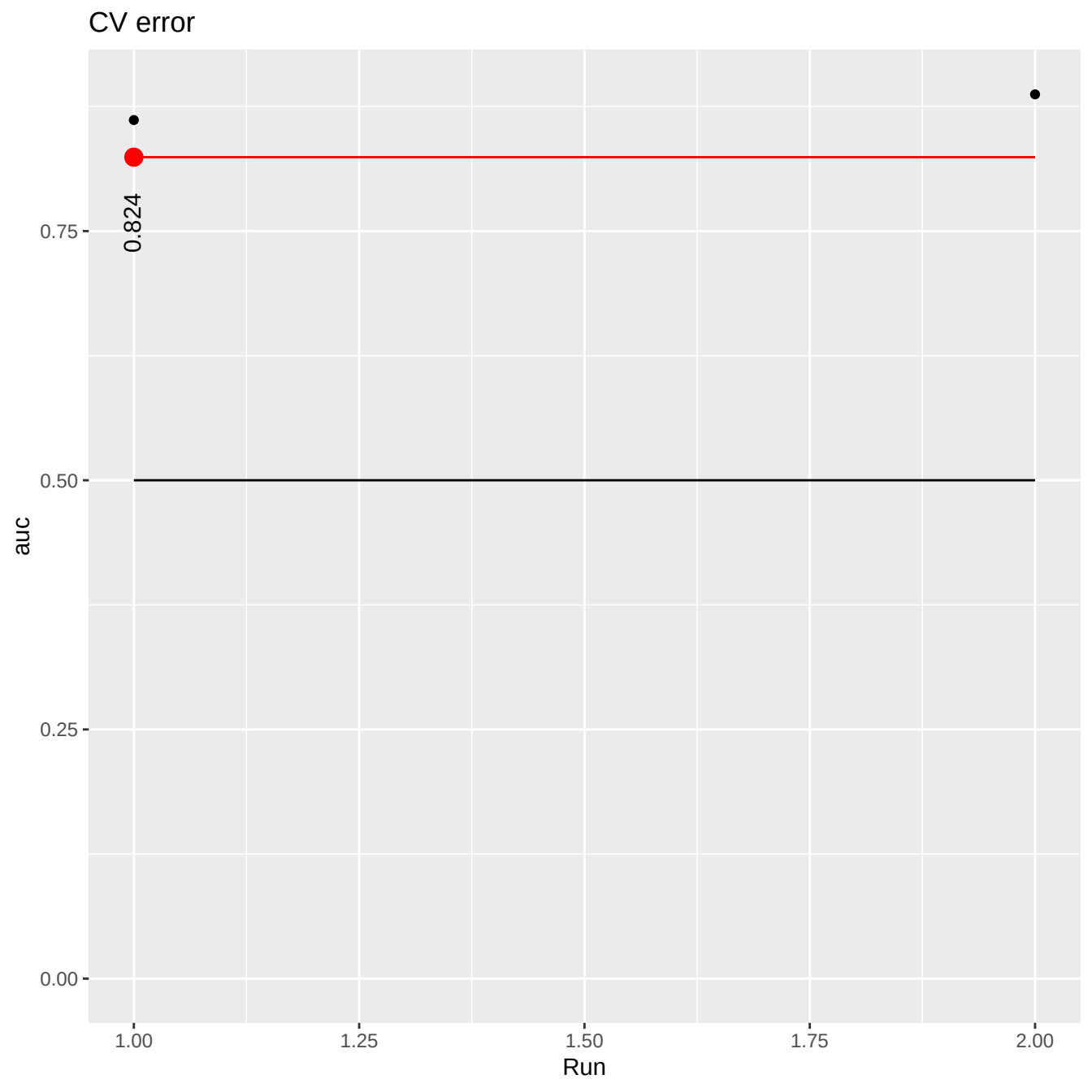

Predicted
Actual
Partial_Dependence
Value
0.30
0.25
0.20
0.15
0.10
0.05
0.00
20
30
40
50
60
70
80
90
100
110
120
Count
80
60
40
20
0
MonthlyCharges

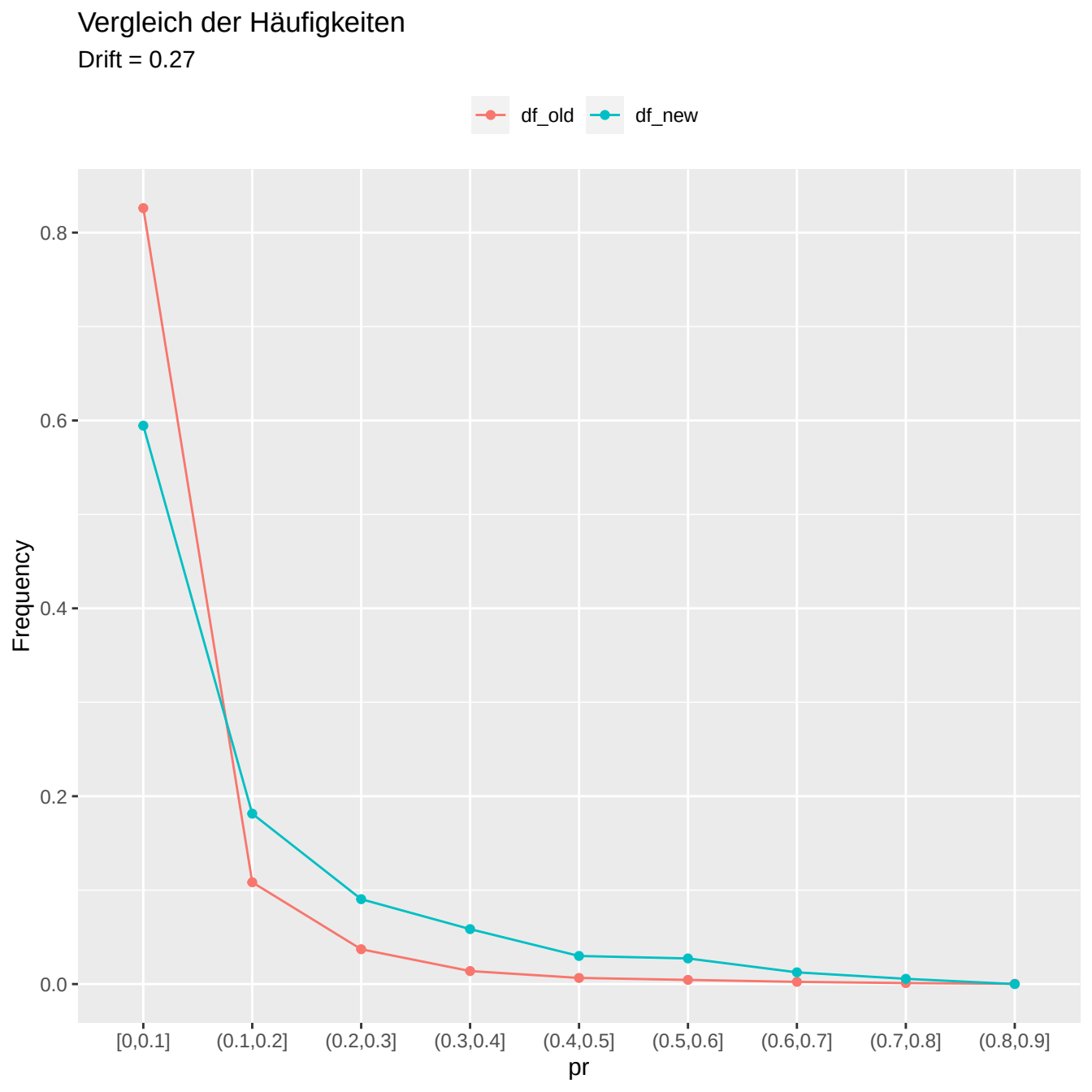

Predicted
Actual
Partial_Dependence
Value
0.10
0.08
0.06
0.04
0.02
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
300
200
100
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Contract

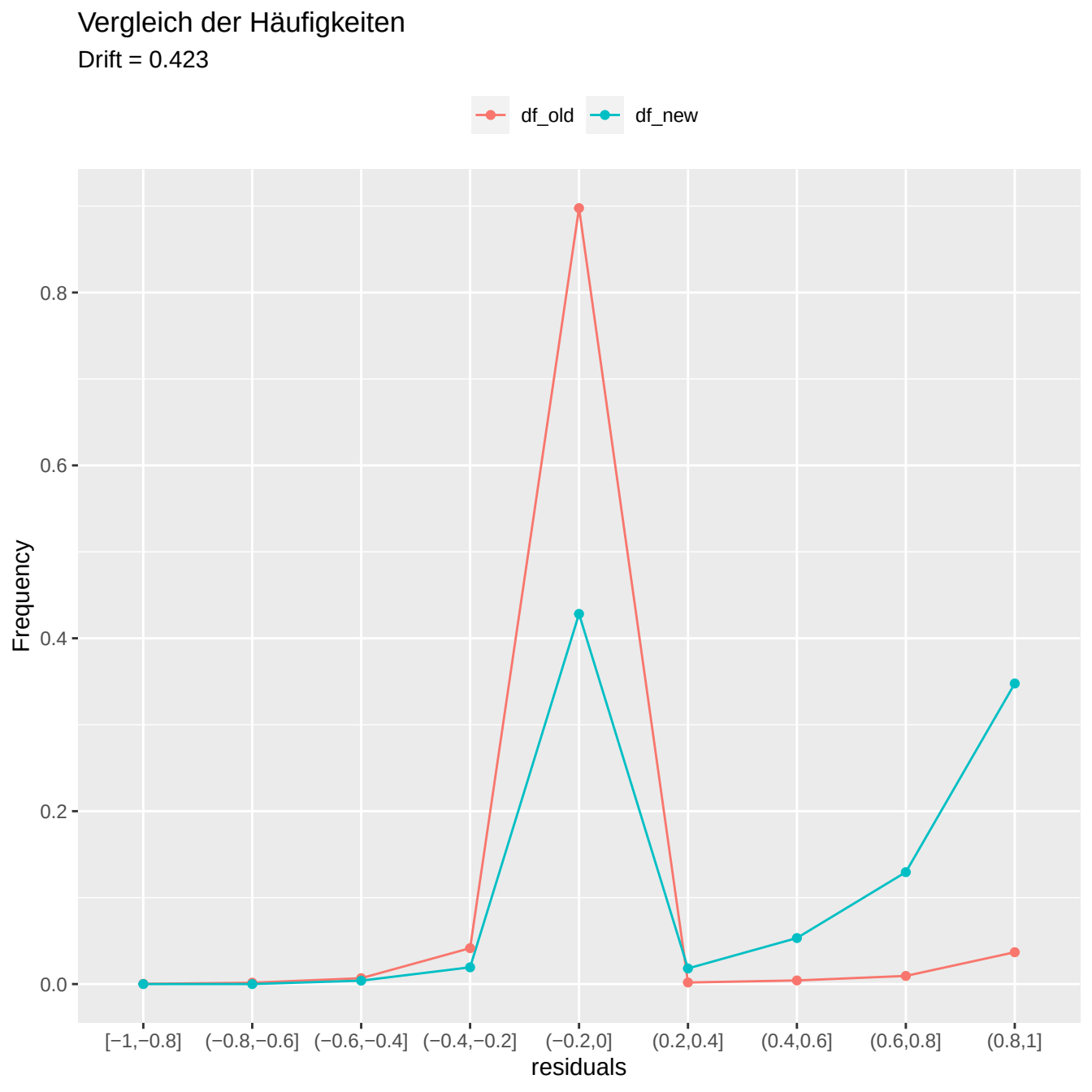

Predicted
Actual
Partial_Dependence
Value
1.0
0.8
0.6
0.4
0.2
0.0
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
Count
90
60
30
0
TotalCharges

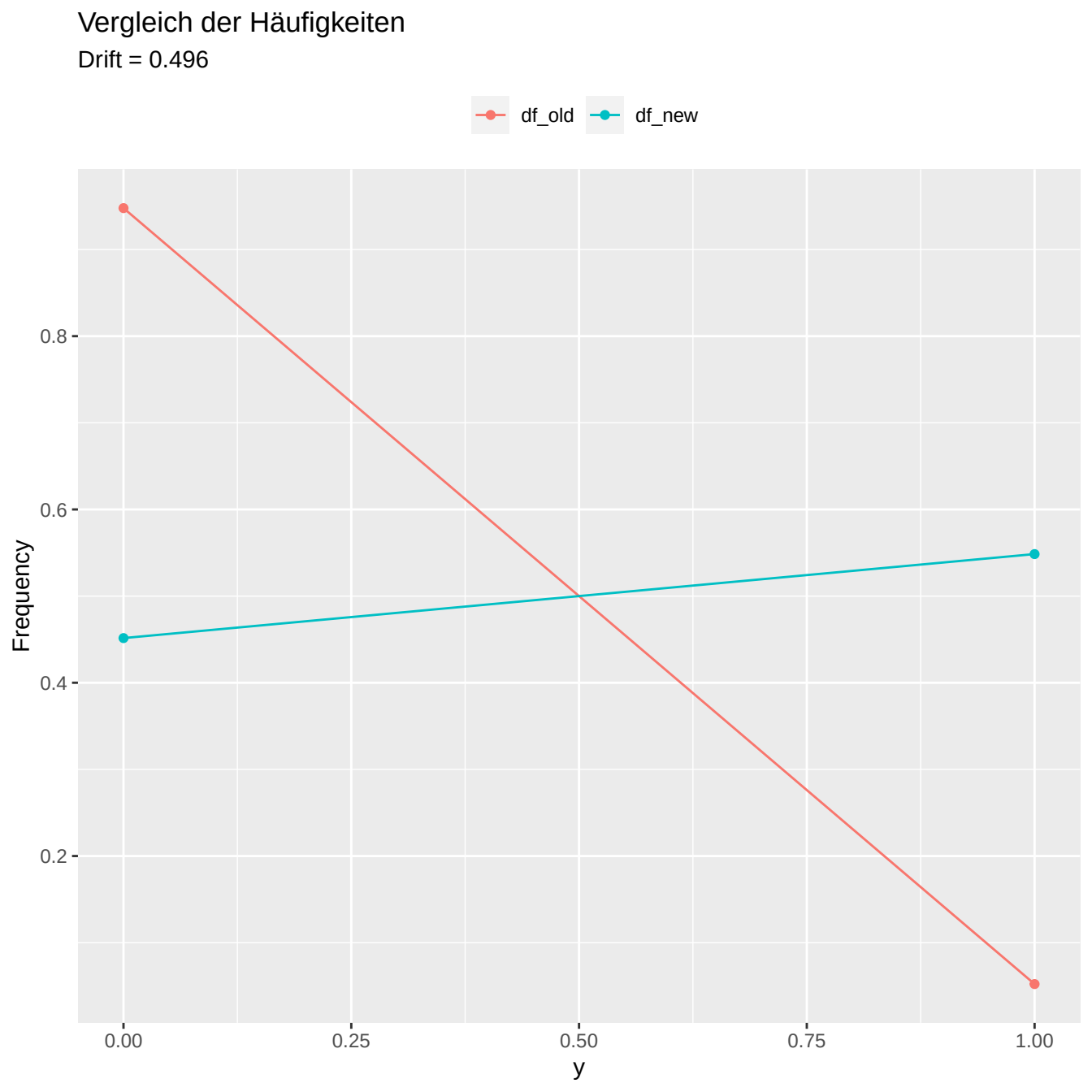

Predicted
Actual
Partial_Dependence
Value
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
50
60
70
Count
60
40
20
0
tenure

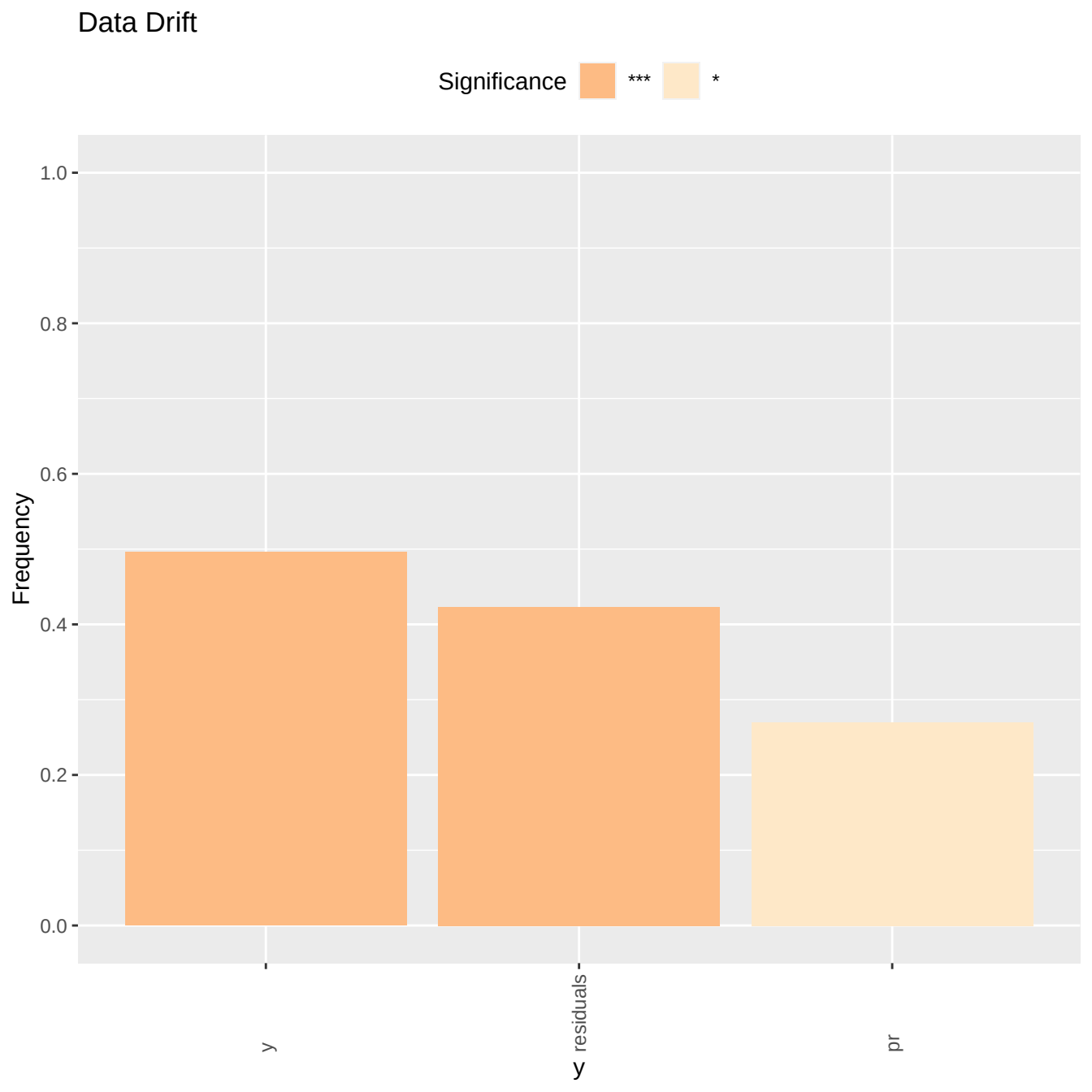

CV error
0.824
0.75
0.50
0.25
0.00
auc
0.950
0.975
1.000
1.025
1.05
Run

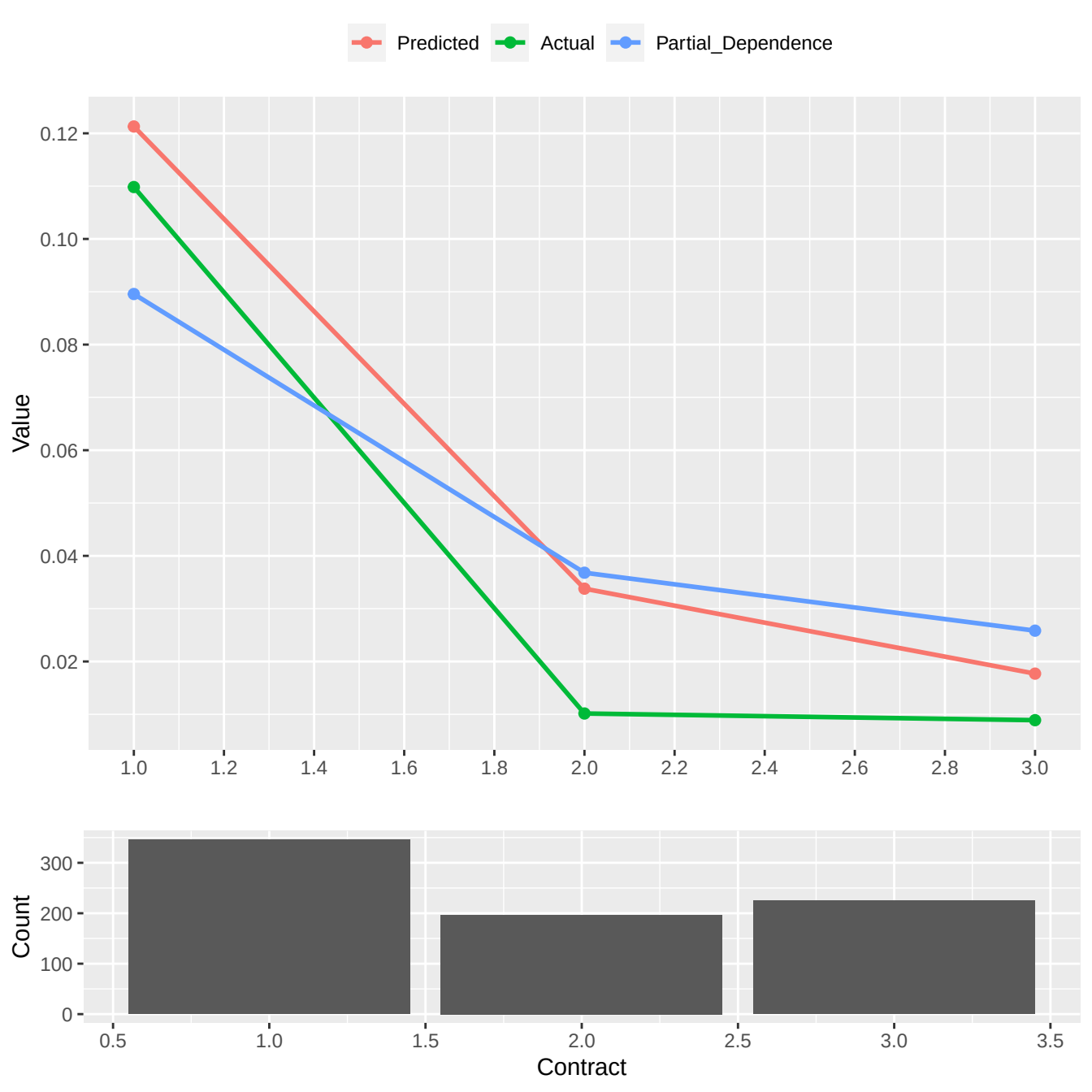

CV error
0.824
0.75
0.50
0.25
0.00
auc
1.00
1.25
1.50
1.75
2.00
Run

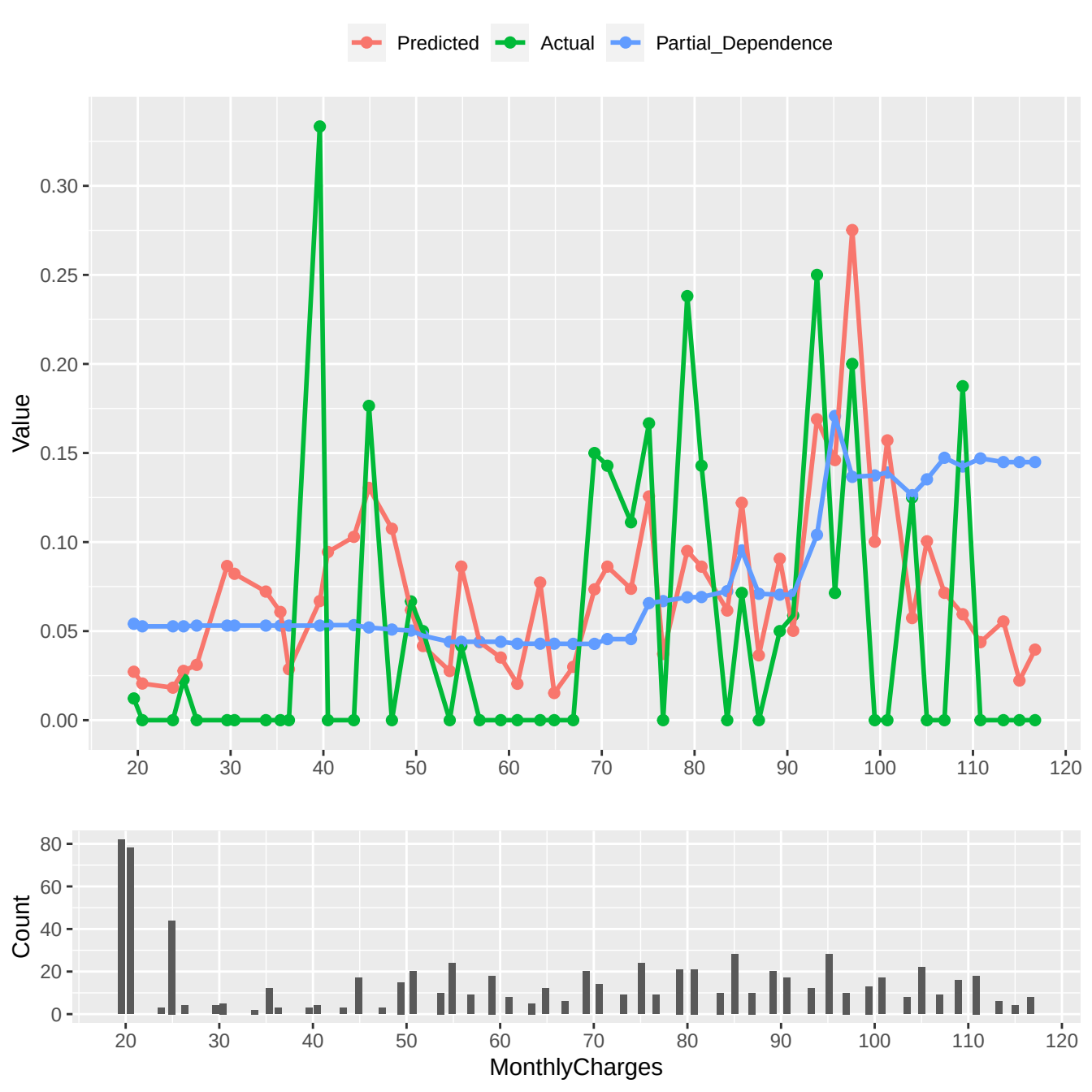

Vergleich der Häufigkeiten
Drift = 0.27
df_old
df_new
0.8
0.6
0.4
0.2
0.0
Frequency
[0,0.1]
(0.1,0.2]
(0.2,0.3]
(0.3,0.4]
(0.4,0.5]
(0.5,0.6]
(0.6,0.7]
(0.7,0.8]
(0.8,0.9]
pr


Vergleich der Häufigkeiten
Drift = 0.423
df_old
df_new
Frequency
0.0
0.2
0.4
0.6
0.8
[−1,−0.8]
(−0.8,−0.6]
(−0.6,−0.4]
(−0.4,−0.2]
(−0.2,0]
(0.2,0.4]
(0.4,0.6]
(0.6,0.8]
(0.8,1]
residuals


Vergleich der Häufigkeiten
Drift = 0.496
df_old
df_new
Frequency
0.8
0.6
0.4
0.2
0.00
0.25
0.50
0.75
1.00
y

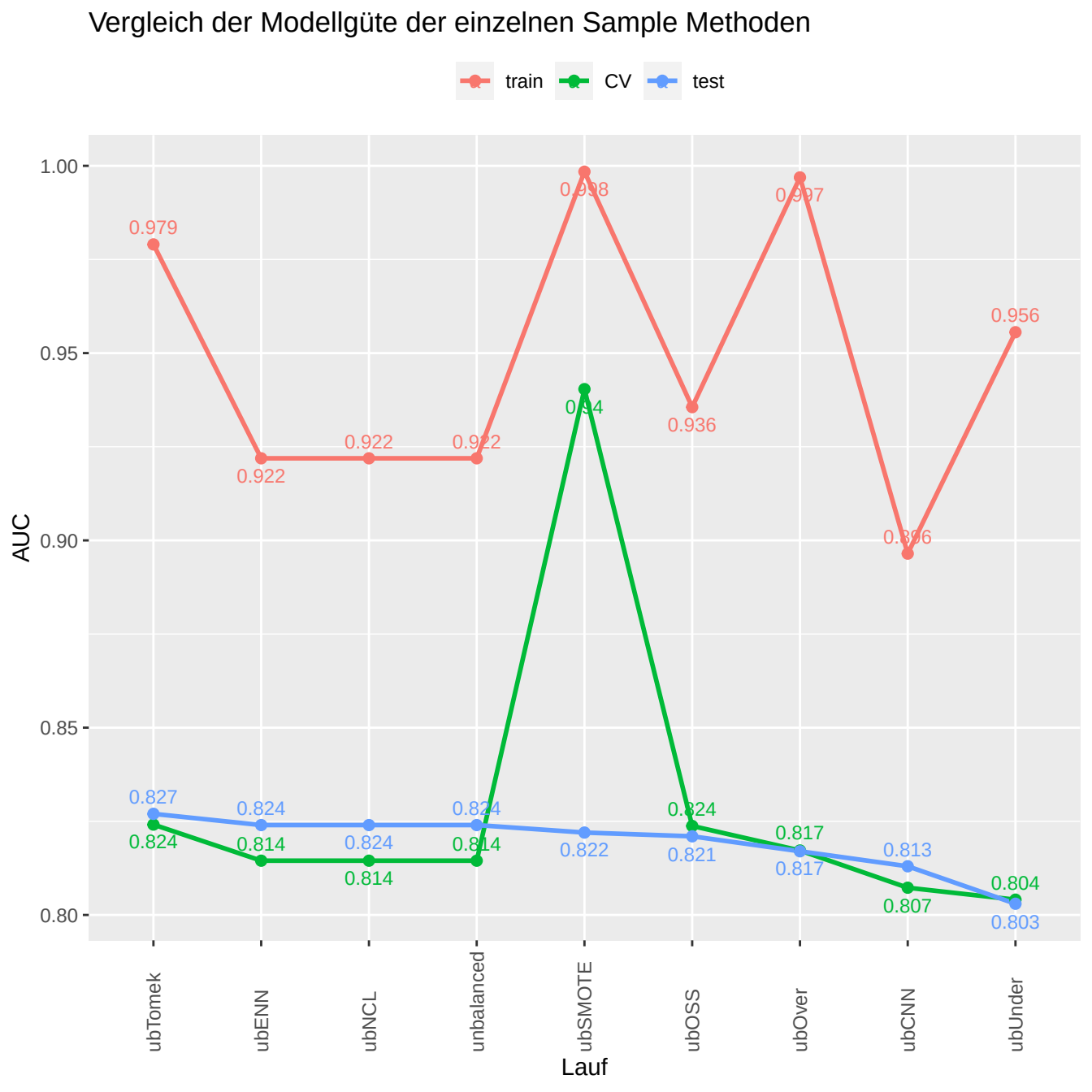

Data Drift
Significance
***
*
1.0
0.8
0.6
0.4
0.2
0.0
Frequency
y
residuals
pr
y

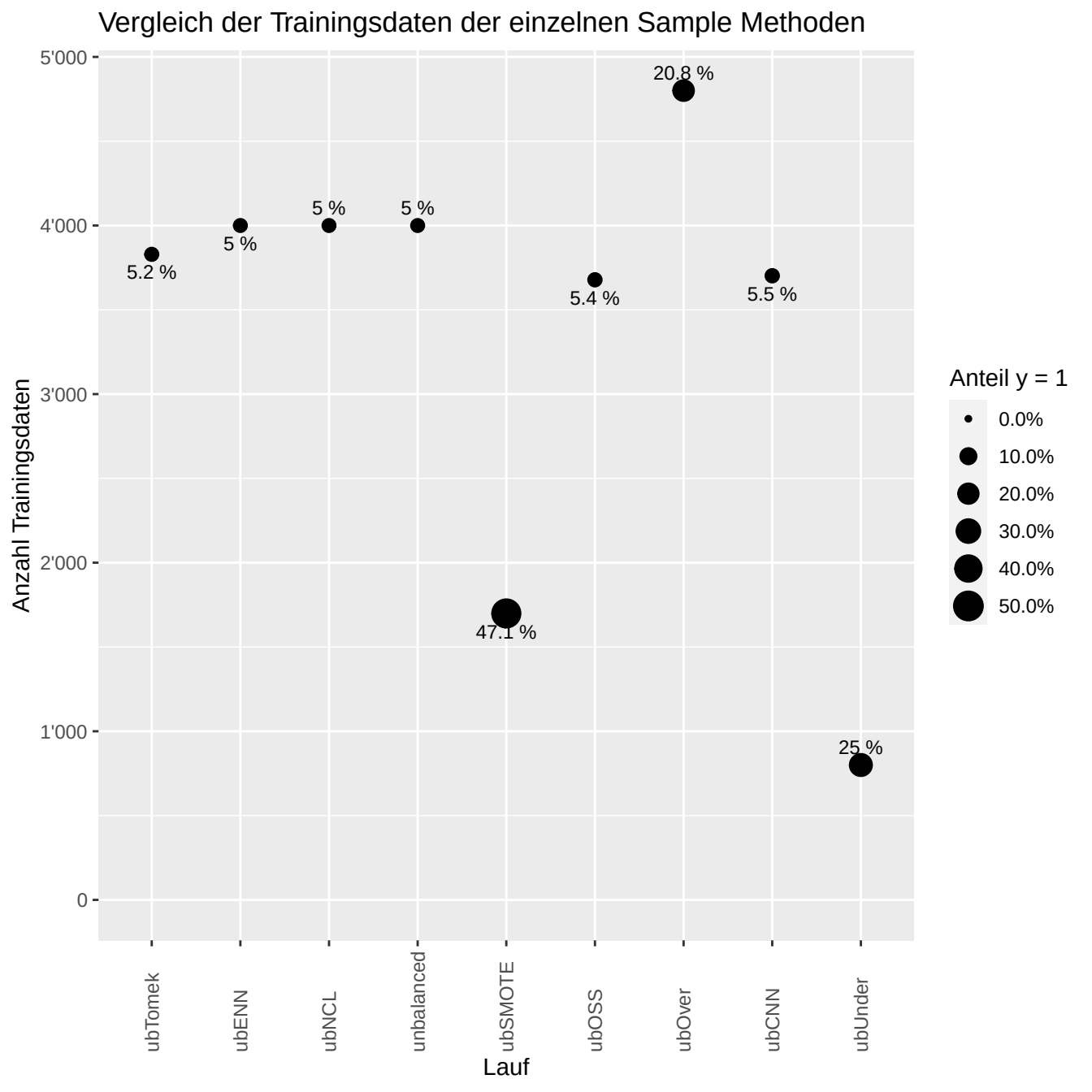

Predicted
Actual
Partial_Dependence
Value
0.12
0.10
0.08
0.06
0.04
0.02
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Count
300
200
100
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Contract


Predicted
Actual
Partial_Dependence
Value
0.30
0.25
0.20
0.15
0.10
0.05
0.00
20
30
40
50
60
70
80
90
100
110
120
Count
80
60
40
20
0
MonthlyCharges


Predicted
Actual
Partial_Dependence
Value
0.20
0.15
0.10
0.05
0.00
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
Count
100
75
50
25
0
TotalCharges


Predicted
Actual
Partial_Dependence
Value
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
50
60
70
Count
80
60
40
20
0
tenure


Vergleich der Modellgüte der einzelnen Sample Methoden
train
CV
test
AUC
Lauf
1.00
0.95
0.90
0.85
0.80
ubTomek
ubENN
ubNCL
unbalanced
ubSMOTE
ubOSS
ubOver
ubCNN
ubUnder
0.979
0.922
0.922
0.922
0.998
0.936
0.997
0.896
0.956
0.824
0.814
0.814
0.814
0.934
0.824
0.817
0.807
0.804
0.827
0.824
0.824
0.824
0.822
0.821
0.817
0.813
0.803


Vergleich der Trainingsdaten der einzelnen Sample Methoden
Anzahl Trainingsdaten
5'000
4'000
3'000
2'000
1'000
0
5.2 %
5 %
5 %
5 %
47.1 %
5.4 %
20.8 %
5.5 %
25 %
ubTomek
ubENN
ubNCL
unbalanced
ubSMOTE
ubOSS
ubOver
ubCNN
ubUnder
Lauf
Anteil y = 1
0.0%
10.0%
20.0%
30.0%
40.0%
50.0%