**The GLMSELECT Procedure**

| Data Set | WORK.TRAIN_FINAL |
|---|---|
| Dependent Variable | Price_Log |
| Selection Method | LASSO |
| Stop at Specified Number of Effects | 3 |
| Choose Criterion | SBC |
| Effect Hierarchy Enforced | None |

| Number of Observations Read | 38527 |
|---|---|
| Number of Observations Used | 19137 |

| Dimensions | |
|---|---|
| Number of Effects | 43 |
| Number of Parameters | 43 |

**The GLMSELECT Procedure**

| | | LASSO Selection Summary | | |
|---|---|---|---|---|
| Step | Effect Entered | Effect Removed | Number Effects In | SBC |
| 0 | Intercept | | 1 | -18284.282 |
| 1 | r_entire | | 2 | -22993.005 |
| 2 | log_accom | | 3 | -28494.815* |
| | * Optimal Value of Criterion | | | |

Selection stopped at the first model containing the specified number of effects (3).



Coefficient Progression for Price_Log

## Fit Criteria for Price_Log



## Progression of Average Squared Errors for Price_Log



**The GLMSELECT Procedure**
**Selected Model**
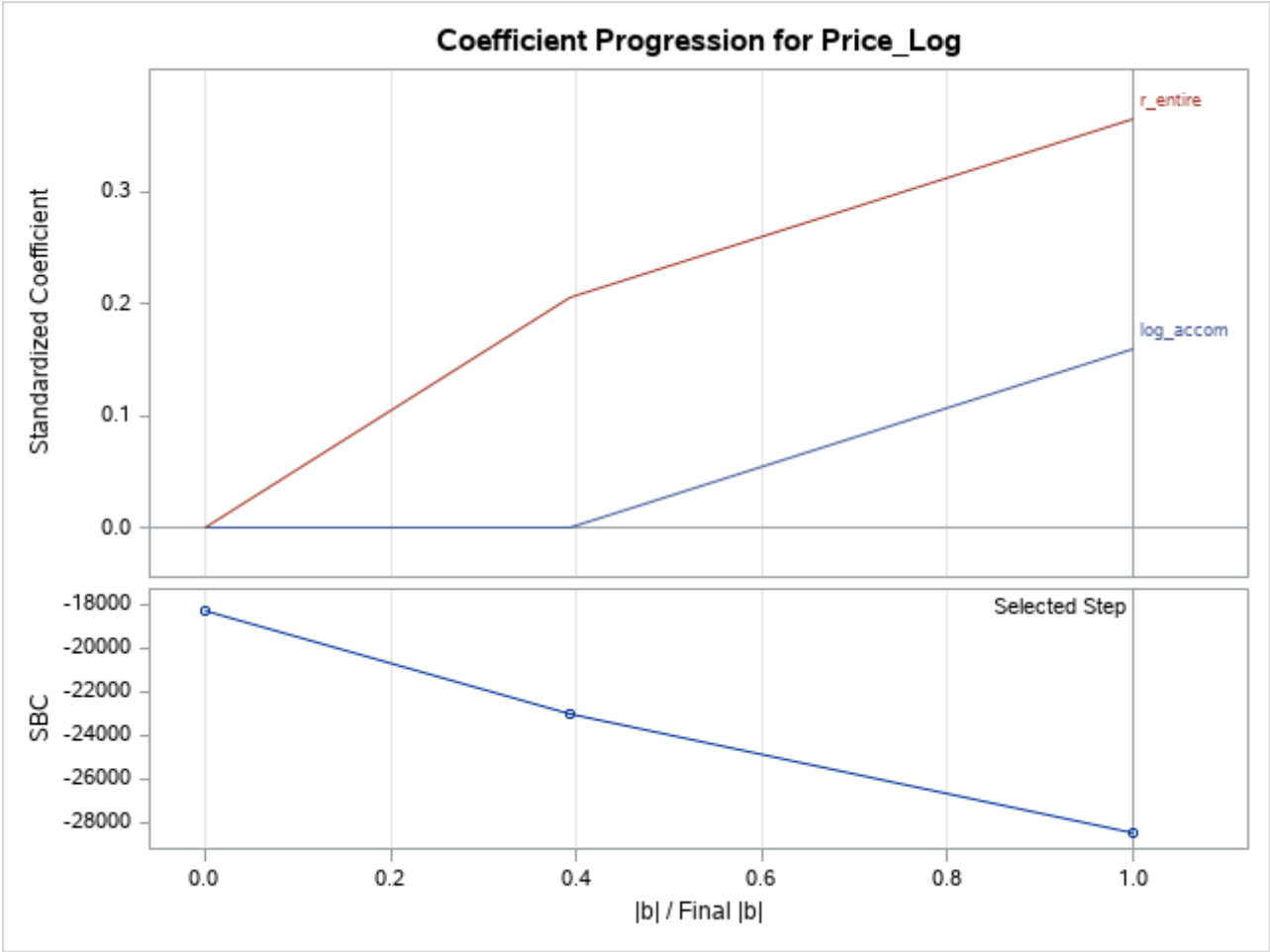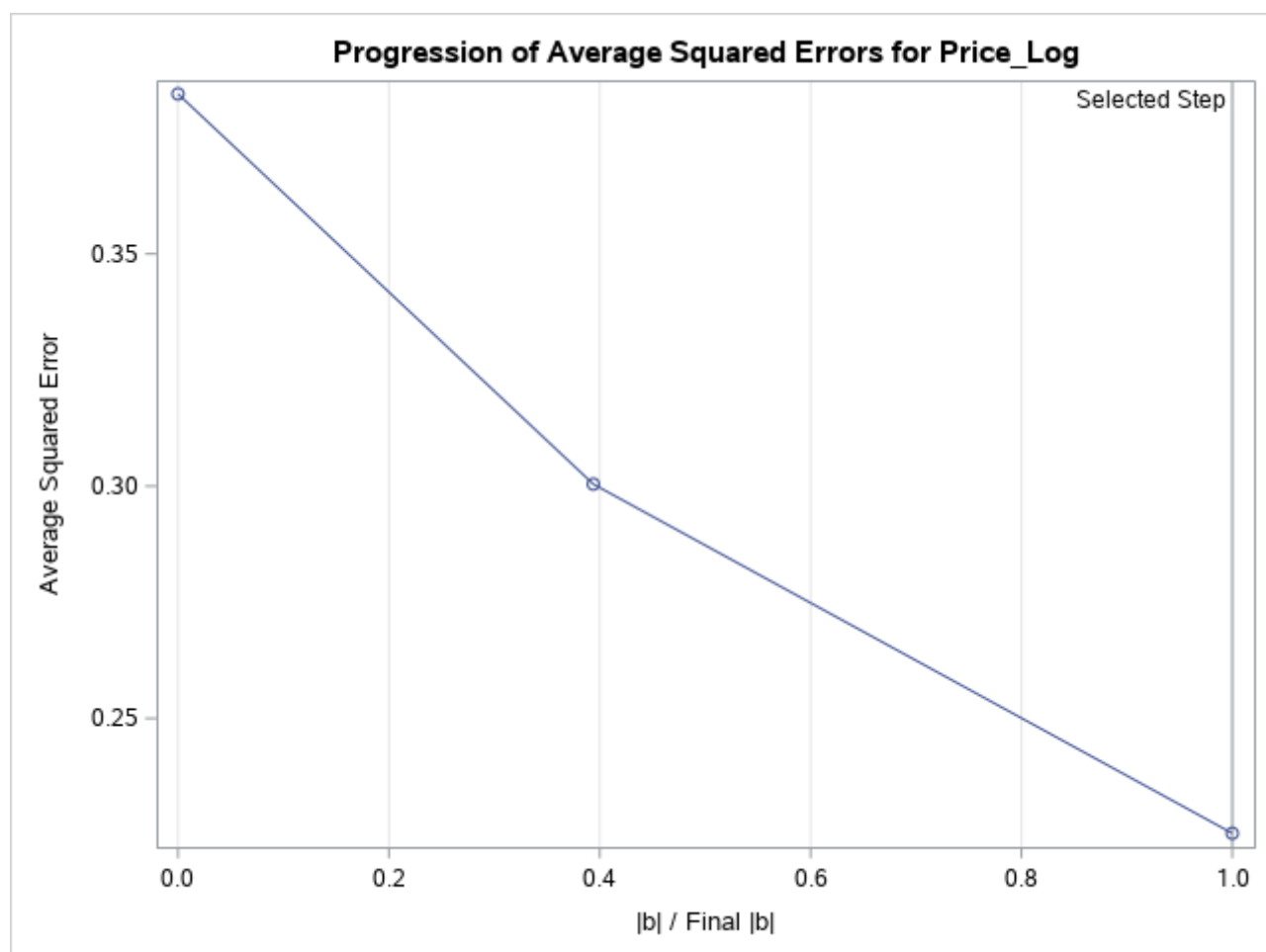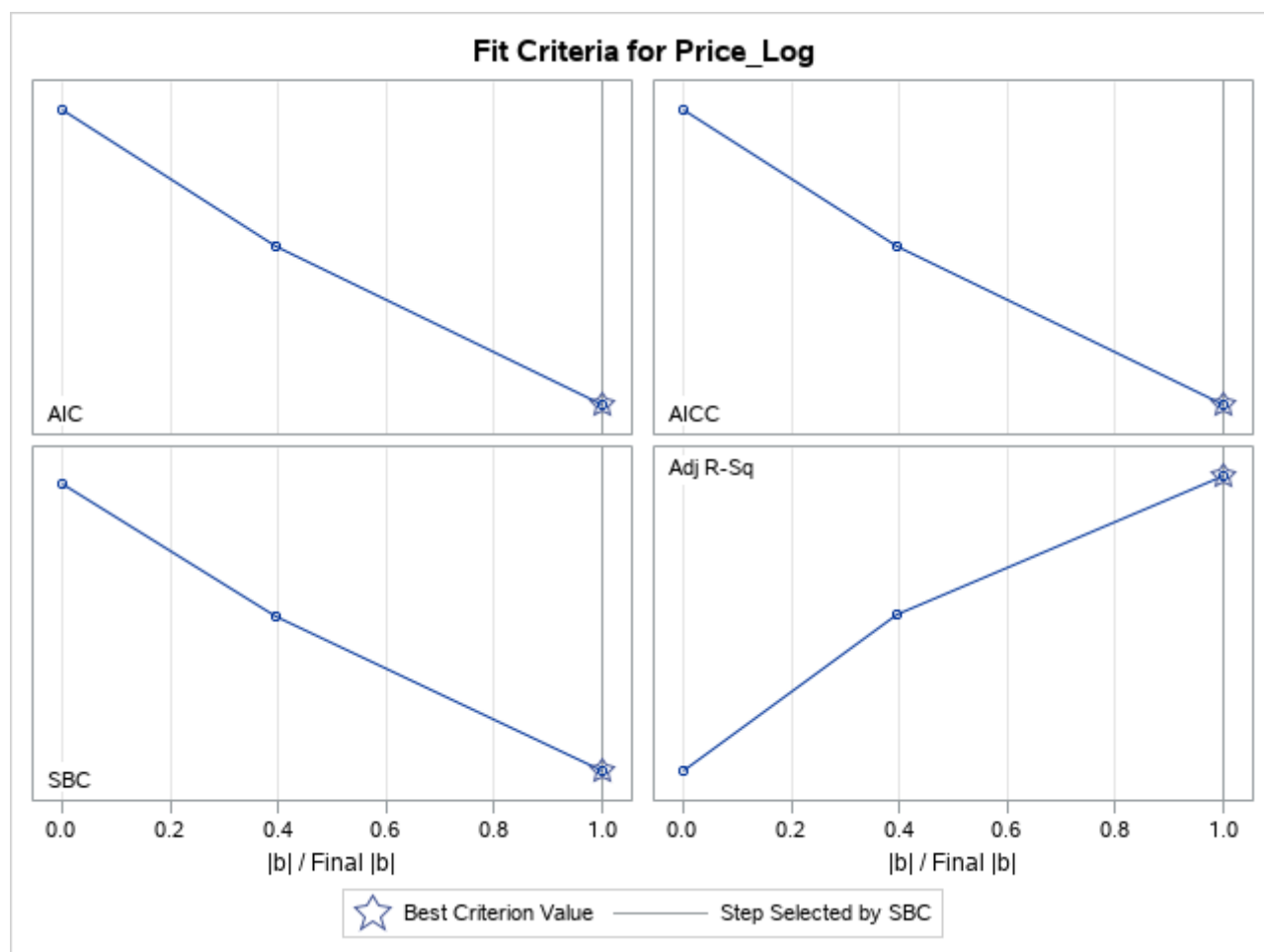
**The selected model, based on SBC, is the model at Step 2.**

| Effects: | Intercept log_accom r_entire |
|---|---|

| Analysis of Variance | | | | |
|---|---|---|---|---|
| Source | DF | Sum of Squares | Mean Square | F Value |
| Model | 2 | 3046.46495 | 1523.23248 | 6761.29 |
| Error | 19134 | 4310.64406 | 0.22529 | |
| Corrected Total | 19136 | 7357.10901 | | |

| | |
|---|---|
| Root MSE | 0.47464 |
| Dependent Mean | 4.87761 |
| R-Square | 0.4141 |
| Adj R-Sq | 0.4140 |
| AIC | -9379.39330 |
| AICC | -9379.39121 |
| SBC | -28495 |

| Parameter Estimates | | |
|---|---|---|
| Parameter | DF | Estimate |
| Intercept | 1 | 4.575847 |
| log_accom | 1 | 0.144656 |
| r_entire | 1 | 0.473156 |

| Score Information | |
|---|---|
| Input Data Set | WORK.TEST_FINAL |
| Output Data Set | WORK.TEST_SCORE |
| | |
| Number of Observations Read | 9617 |
| Number of Observations Scored | 0 |
| Number of Residuals Computed | 0 |
| | |
| Residual Sum of Squares | . |

**The GLMSELECT Procedure**

| | |
|---|---|
| Data Set | WORK.TRAIN_FINAL |
| Dependent Variable | Price_Log |
| Selection Method | LASSO |
| Stop at Specified Number of Effects | 4 |
| Choose Criterion | SBC |
| Effect Hierarchy Enforced | None |

| | |
|---|---|
| Number of Observations Read | 38527 |
| Number of Observations Used | 19137 |

| Dimensions | |
|---|---|
| Number of Effects | 43 |
| Number of Parameters | 43 |

**The GLMSELECT Procedure**

| LASSO Selection Summary | | | | |
|---|---|---|---|---|
| Step | Effect Entered | Effect Removed | Number Effects In | SBC |
| 0 | Intercept | | 1 | -18284.282 |
| 1 | r_entire | | 2 | -22993.005 |
| 2 | log_accom | | 3 | -28494.815 |
| 3 | poly_accom | | 4 | -29416.248* |
| * Optimal Value of Criterion | | | | |

Selection stopped at the first model containing the specified number of effects (4).

## Fit Criteria for Price_Log



## Progression of Average Squared Errors for Price_Log



**The GLMSELECT Procedure**
**Selected Model**
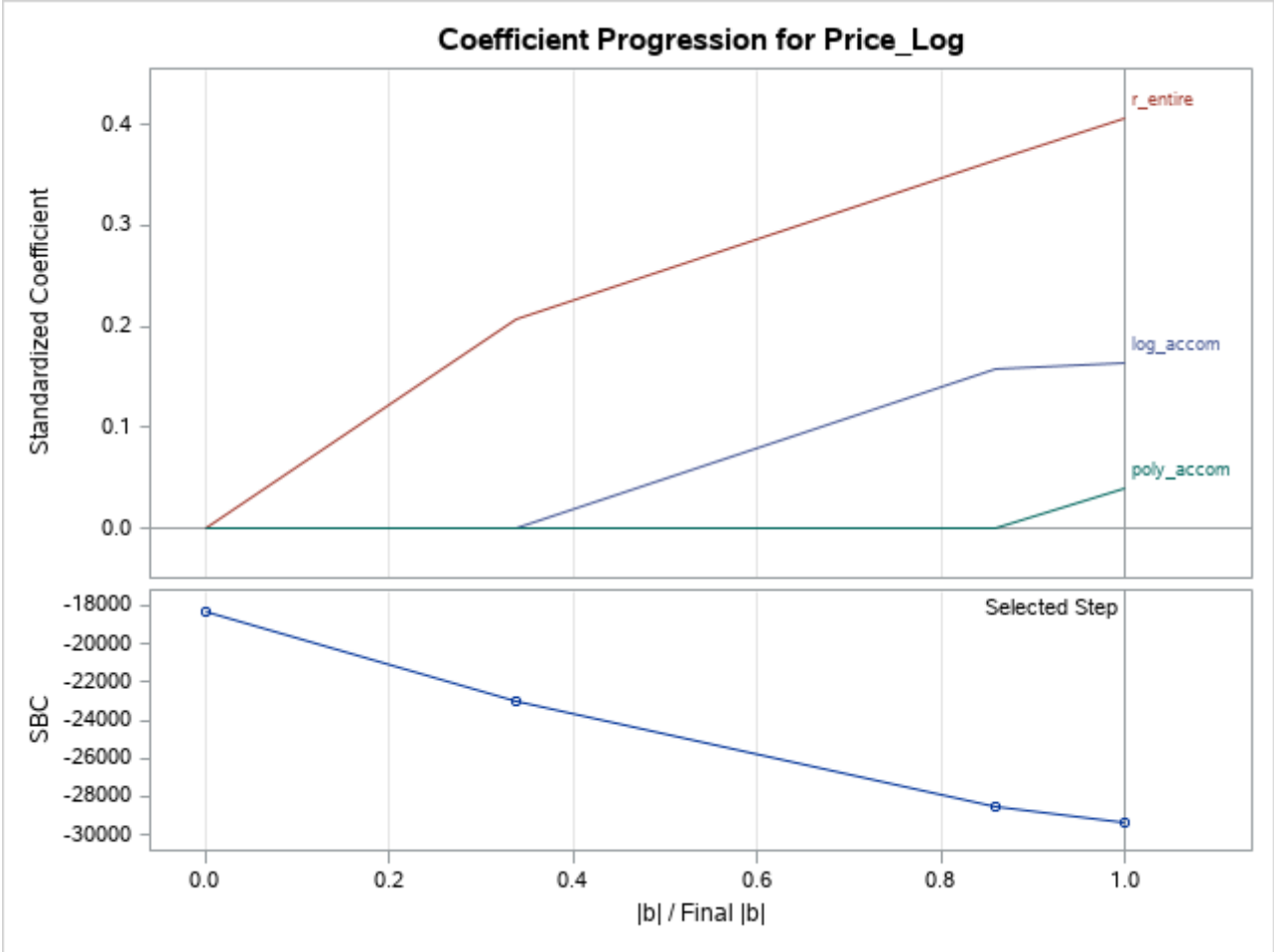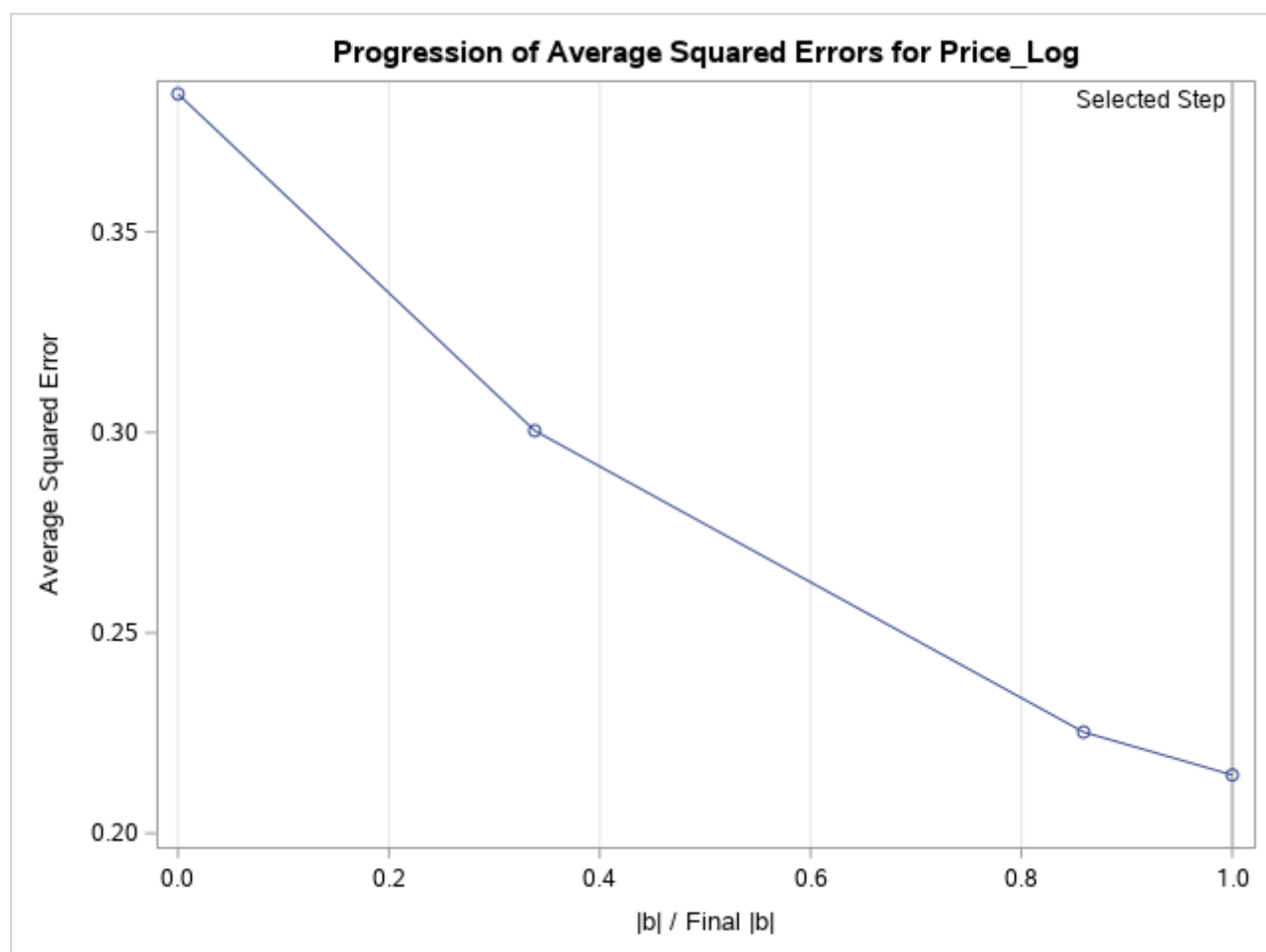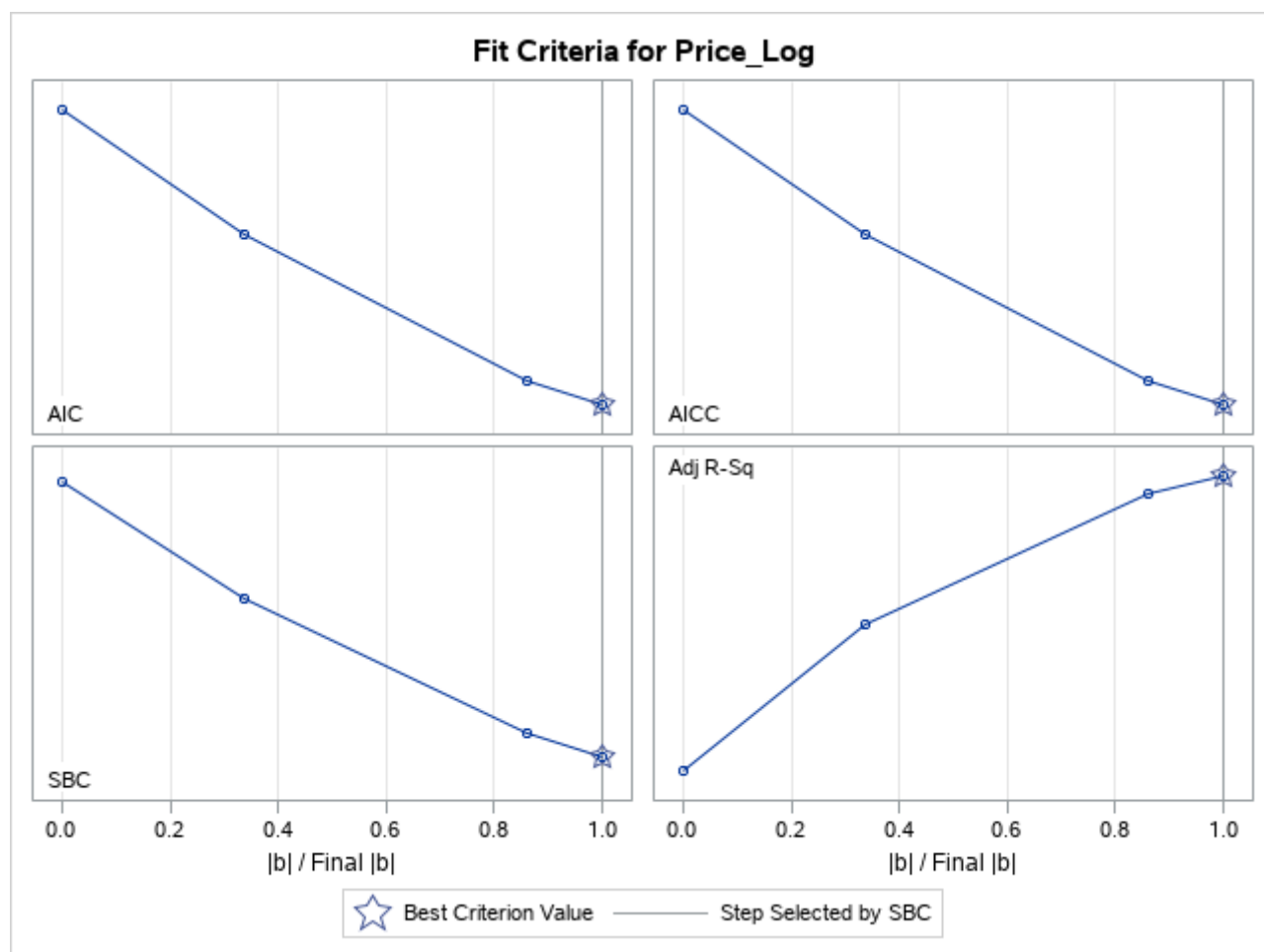
**The selected model, based on SBC, is the model at Step 3.**

| Effects: | Intercept log_accom r_entire poly_accom |
|---|---|

| Analysis of Variance | | | | |
|---|---|---|---|---|
| Source | DF | Sum of Squares | Mean Square | F Value |
| Model | 3 | 3251.21771 | 1083.73924 | 5050.11 |
| Error | 19133 | 4105.89130 | 0.21460 | |
| Corrected Total | 19136 | 7357.10901 | | |

| | |
|---|---|
| Root MSE | 0.46325 |
| Dependent Mean | 4.87761 |
| R-Square | 0.4419 |
| Adj R-Sq | 0.4418 |
| AIC | -10309 |
| AICC | -10309 |
| SBC | -29416 |

| Parameter Estimates | | |
|---|---|---|
| Parameter | DF | Estimate |
| Intercept | 1 | 4.537954 |
| log_accom | 1 | 0.149530 |
| r_entire | 1 | 0.526708 |
| poly_accom | 1 | 0.021951 |

| Score Information | |
|---|---|
| Input Data Set | WORK.TEST_FINAL |
| Output Data Set | WORK.TEST_SCORE |
| | |
| Number of Observations Read | 9617 |
| Number of Observations Scored | 0 |
| Number of Residuals Computed | 0 |
| | |
| Residual Sum of Squares | . |

**The GLMSELECT Procedure**

| | |
|---|---|
| Data Set | WORK.TRAIN_FINAL |
| Dependent Variable | Price_Log |
| Selection Method | LASSO |
| Stop at Specified Number of Effects | 5 |
| Choose Criterion | SBC |
| Effect Hierarchy Enforced | None |

| | |
|---|---|
| Number of Observations Read | 38527 |
| Number of Observations Used | 19137 |

| Dimensions | |
|---|---|
| Number of Effects | 43 |
| Number of Parameters | 43 |

**The GLMSELECT Procedure**

| | | | LASSO Selection Summary | |
|---|---|---|---|---|
| **Step** | **Effect Entered** | **Effect Removed** | **Number Effects In** | **SBC** |
| 0 | Intercept | | 1 | -18284.282 |
| 1 | r_entire | | 2 | -22993.005 |
| 2 | log_accom | | 3 | -28494.815 |
| 3 | poly_accom | | 4 | -29416.248 |
| 4 | log_bath | | 5 | -30105.923* |
| **\* Optimal Value of Criterion** | | | | |

Selection stopped at the first model containing the specified number of effects (5).



Coefficient Progression for Price_Log

## Fit Criteria for Price_Log



AIC

AICC

SBC

Adj R-Sq

|b| / Final |b|

|b| / Final |b|

☆ Best Criterion Value ——— Step Selected by SBC

## Progression of Average Squared Errors for Price_Log



Selected Step

Average Squared Error

|b| / Final |b|

**The GLMSELECT Procedure**
**Selected Model**
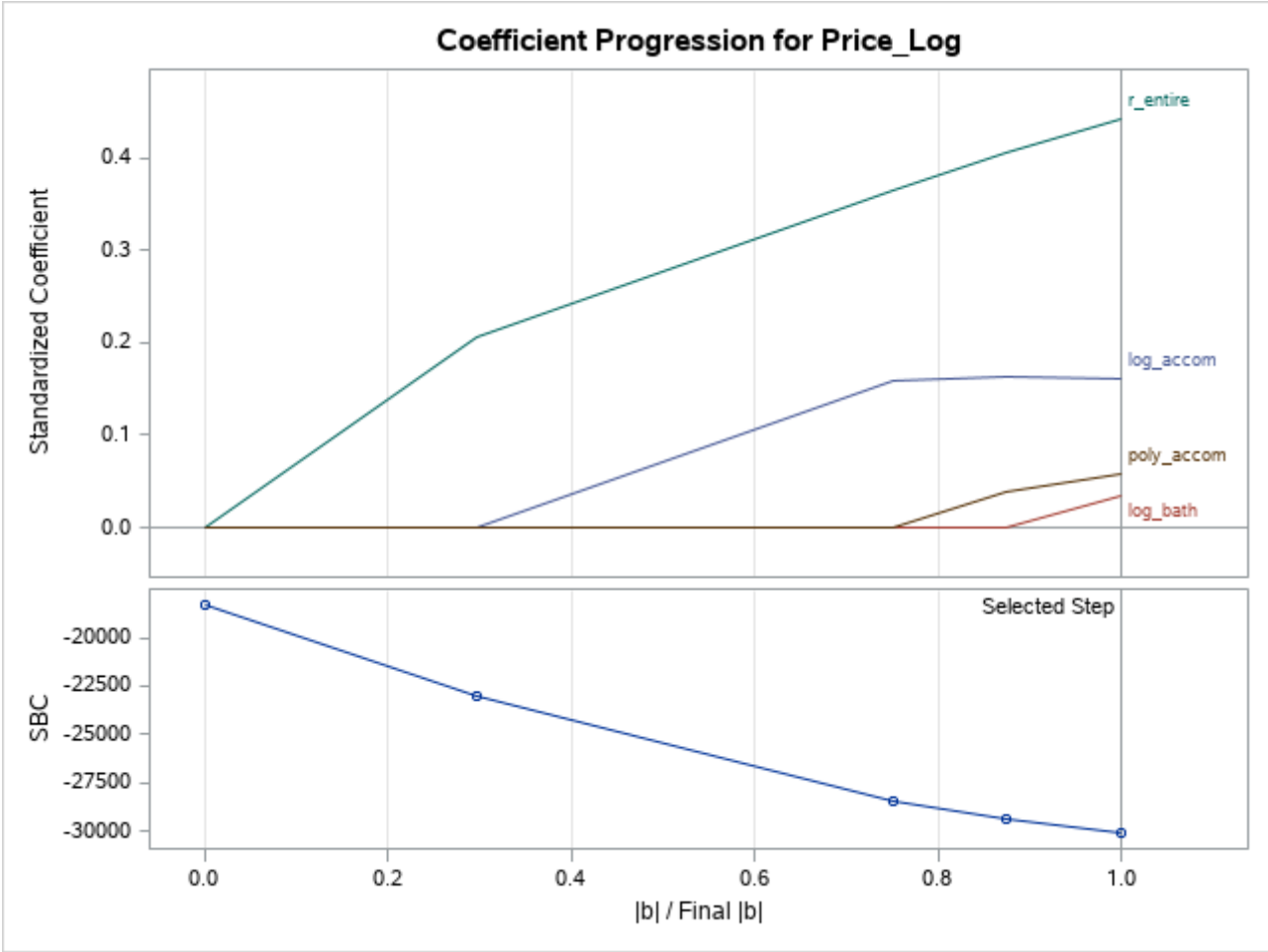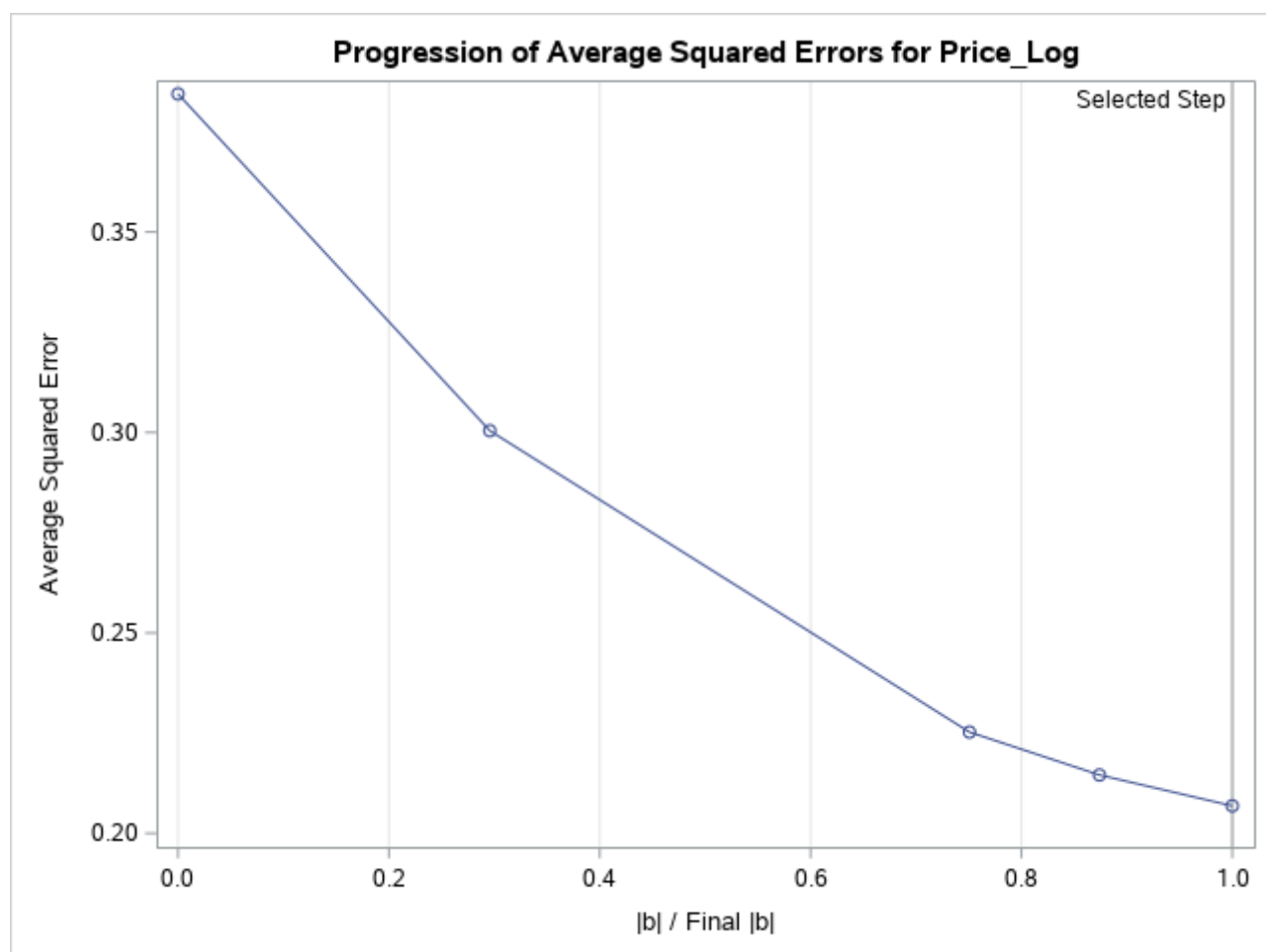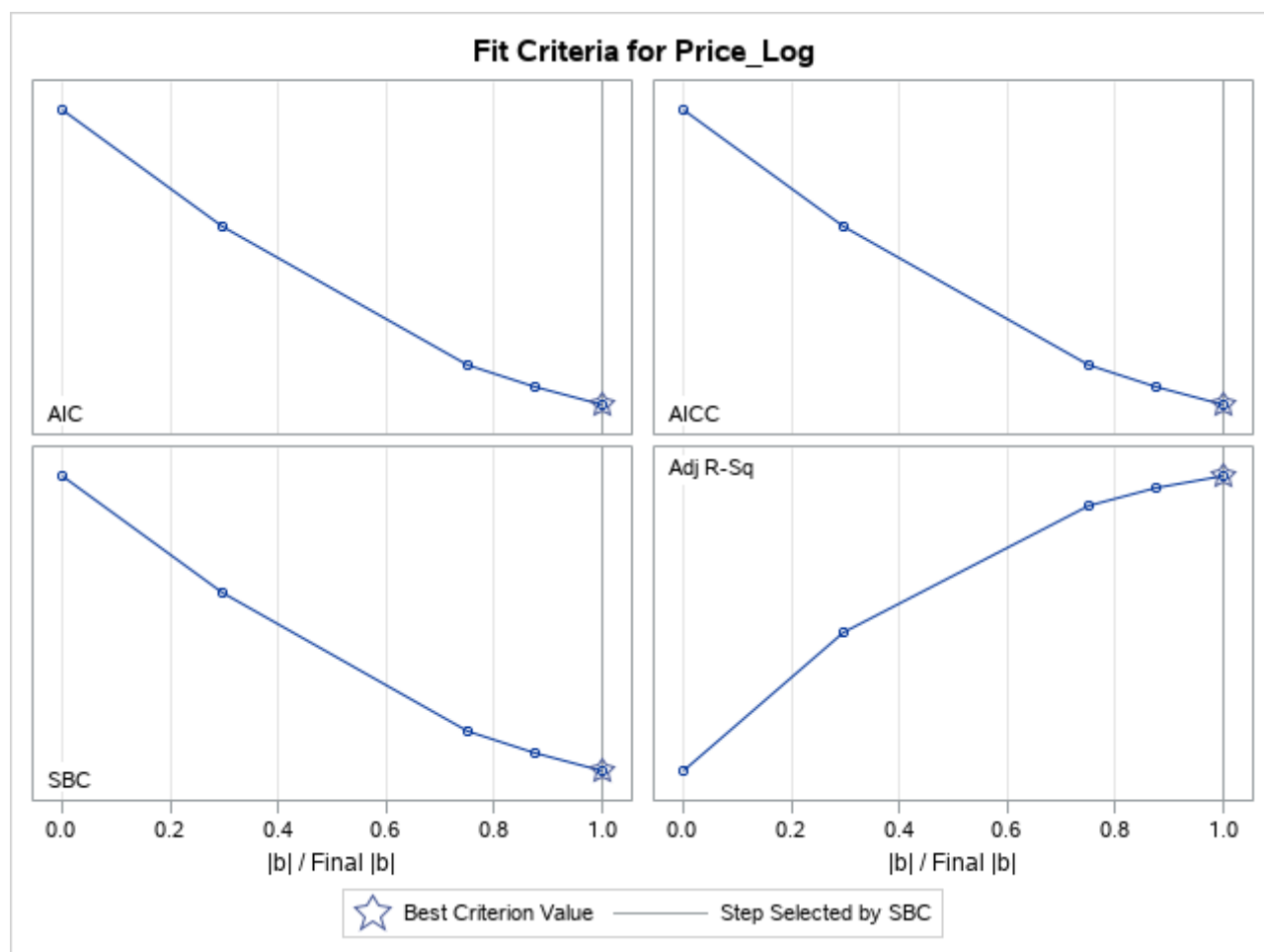
**The selected model, based on SBC, is the model at Step 4.**

| Effects: | Intercept log_accom log_bath r_entire poly_accom |
|---|---|

### Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Value |
|---|---|---|---|---|
| Model | 4 | 3398.59463 | 849.64866 | 4106.46 |
| Error | 19132 | 3958.51438 | 0.20691 | |
| Corrected Total | 19136 | 7357.10901 | | |

| | |
|---|---|
| Root MSE | 0.45487 |
| Dependent Mean | 4.87761 |
| R-Square | 0.4619 |
| Adj R-Sq | 0.4618 |
| AIC | -11006 |
| AICC | -11006 |
| SBC | -30106 |

### Parameter Estimates

| Parameter | DF | Estimate |
|---|---|---|
| Intercept | 1 | 4.513781 |
| log_accom | 1 | 0.147572 |
| log_bath | 1 | 0.036476 |
| r_entire | 1 | 0.573572 |
| poly_accom | 1 | 0.032723 |

### Score Information

| Input Data Set | WORK.TEST_FINAL |
|---|---|
| Output Data Set | WORK.TEST_SCORE |
| | |
| Number of Observations Read | 9617 |
| Number of Observations Scored | 0 |
| Number of Residuals Computed | 0 |
| | |
| Residual Sum of Squares | . |

### The GLMSELECT Procedure

| Data Set | WORK.TRAIN_FINAL |
|---|---|
| Dependent Variable | Price_Log |
| Selection Method | LASSO |
| Stop Criterion | None |
| Choose Criterion | SBC |
| Effect Hierarchy Enforced | None |

| Number of Observations Read | 38527 |
|---|---|
| Number of Observations Used | 19137 |

### Dimensions

| Number of Effects | 43 |
|---|---|
| Number of Parameters | 43 |

**The GLMSELECT Procedure**

| | LASSO Selection Summary | | | |
|---|---|---|---|---|
| Step | Effect Entered | Effect Removed | Number Effects In | SBC |
| 0 | Intercept | | 1 | -18284.282 |
| 1 | r_entire | | 2 | -22993.005 |
| 2 | log_accom | | 3 | -28494.815 |
| 3 | poly_accom | | 4 | -29416.248 |
| 4 | log_bath | | 5 | -30105.923 |
| 5 | p_group1 | | 6 | -30306.602 |
| 6 | poly_avail | | 7 | -30304.380 |
| 7 | p_house | | 8 | -30528.795 |
| 8 | hcount_level3 | | 9 | -30546.804 |
| 9 | n_bronx | | 10 | -30913.406 |
| 10 | p_loft | | 11 | -31238.696 |
| 11 | hcount_level2 | | 12 | -31445.517 |
| 12 | poly_guests | | 13 | -31507.131 |
| 13 | poly_max | | 14 | -31508.194 |
| 14 | poly_min | | 15 | -31537.415 |
| 15 | p_townhouse | | 16 | -31536.427 |
| 16 | p_group2 | | 17 | -31761.464 |
| 17 | b_real | | 18 | -31824.917 |
| 18 | log_bedsper | | 19 | -31872.706 |
| 19 | h_verified | | 20 | -31917.436 |
| 20 | b_futon | | 21 | -31949.831 |
| 21 | h_super | | 22 | -32044.155 |
| 22 | log_max | | 23 | -32067.523 |
| 23 | r_shared | | 24 | -32094.183 |
| 24 | p_apart | | 25 | -32161.444 |
| 25 | require_phone | | 26 | -32212.703 |
| 26 | h_profile | | 27 | -32224.502 |
| 27 | instant | | 28 | -32238.442* |
| 28 | log_guest | | 29 | -32231.327 |
| 29 | log_avil30 | | 30 | -32226.850 |
| 30 | b_air | | 31 | -32228.496 |
| 31 | log_min | | 32 | -32220.056 |
| 32 | require_pic | | 33 | -32210.503 |
| 33 | poly_bath | | 34 | -32203.648 |
| 34 | b_couch | | 35 | -32196.654 |
| | * Optimal Value of Criterion | | | |

Selection stopped because the change of the maximum absolute correction is tiny.
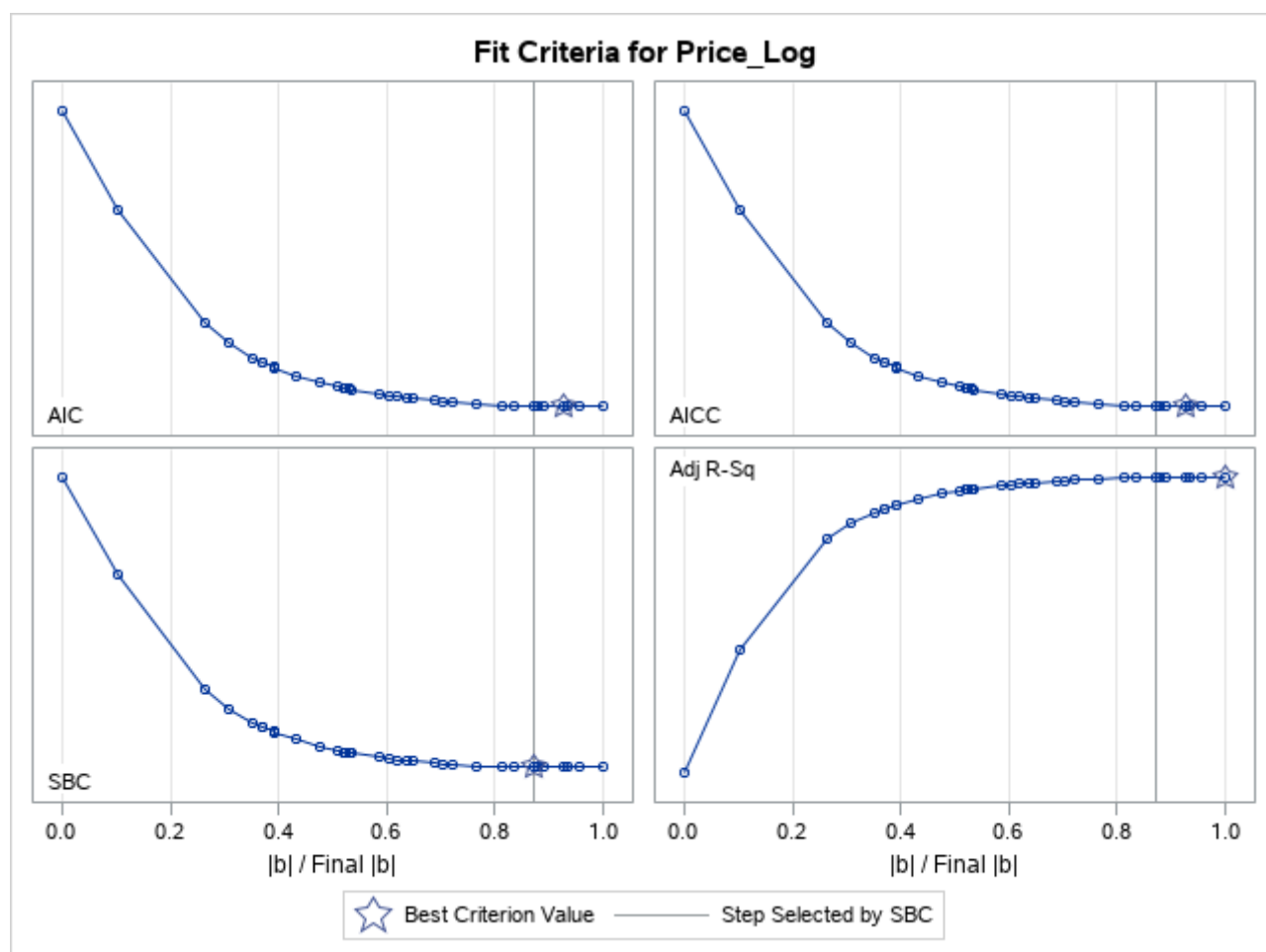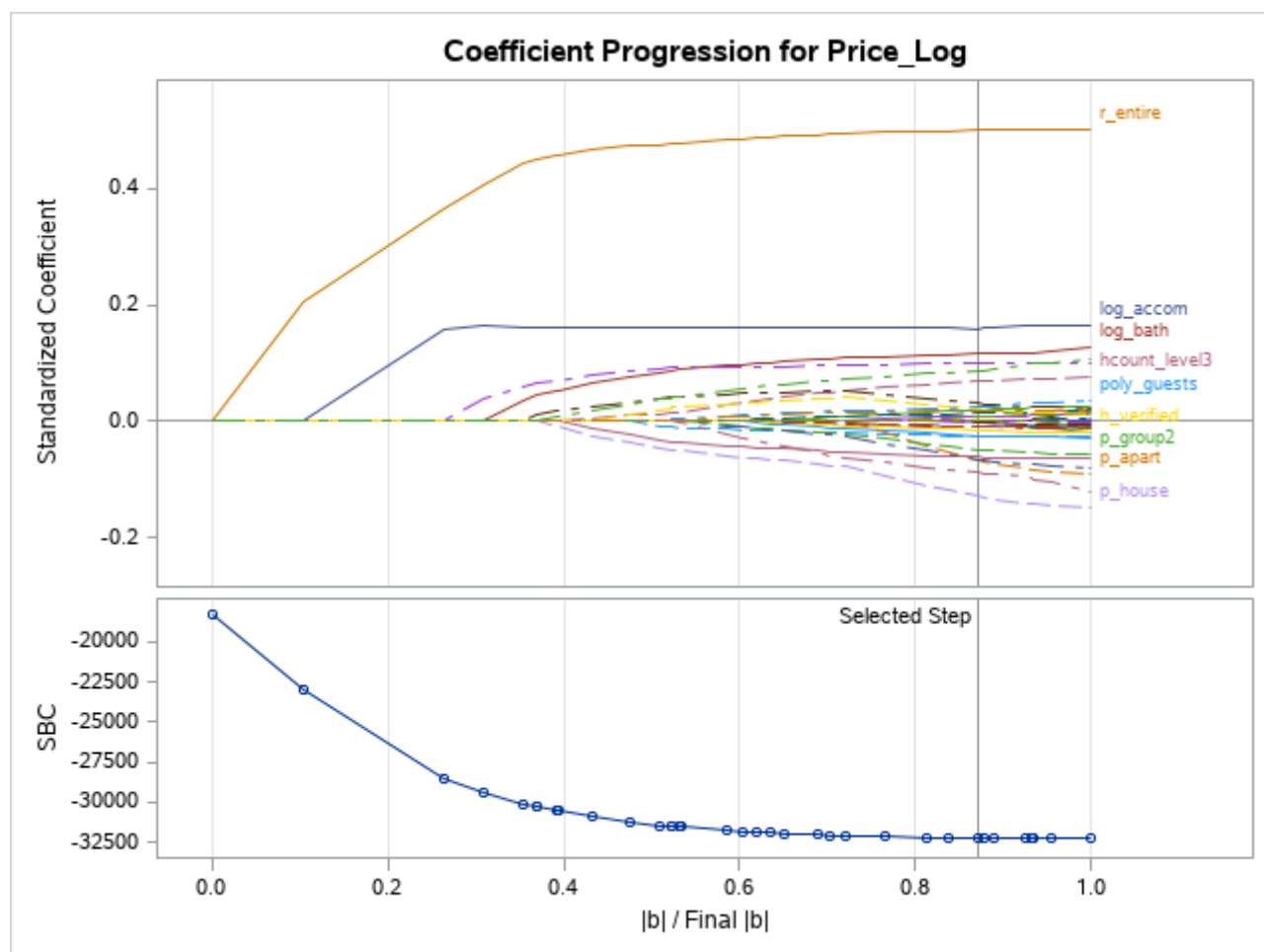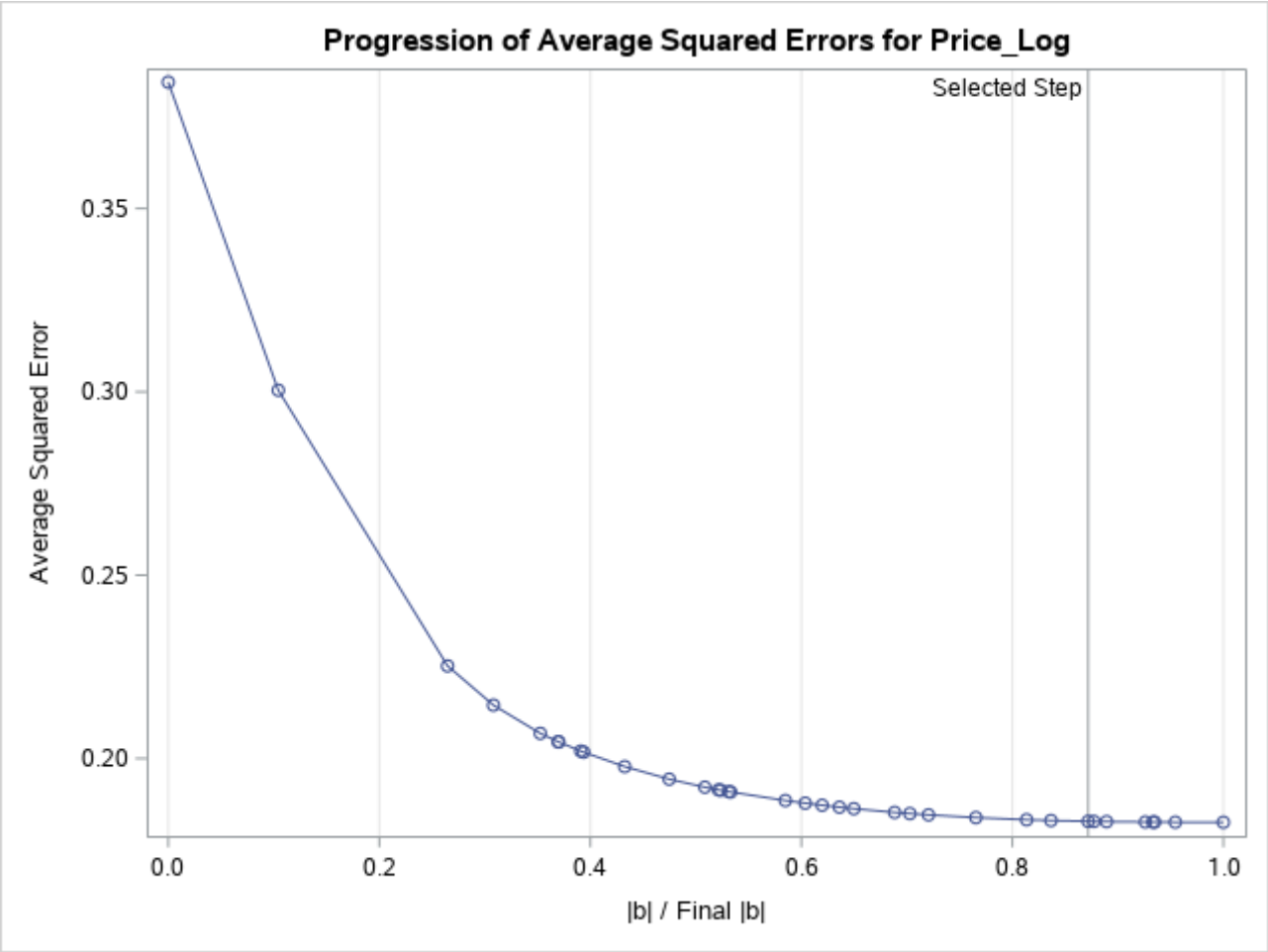
## Coefficient Progression for Price_Log



## Fit Criteria for Price_Log

## Progression of Average Squared Errors for Price_Log



---

**The GLMSELECT Procedure**
**Selected Model**

**The selected model, based on SBC, is the model at Step 27.**

| Effects: | Intercept log_accom log_bath log_max log_bedsper n_bronx r_entire h_super h_profile h_verified b_futon b_real instant require_phone hcount_level2 hcount_level3 p_apart p_group2 p_house p_loft p_townhouse r_shared p_group1 poly_accom poly_guests poly_min poly_max poly_avail |
|---|---|

### Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Value |
|---|---|---|---|---|
| Model | 27 | 3857.73222 | 142.87897 | 780.22 |
| Error | 19109 | 3499.37679 | 0.18313 | |
| Corrected Total | 19136 | 7357.10901 | | |

| | |
|---|---|
| Root MSE | 0.42793 |
| Dependent Mean | 4.87761 |
| R-Square | 0.5244 |
| Adj R-Sq | 0.5237 |
| AIC | -13320 |
| AICC | -13319 |
| SBC | -32238 |

### Parameter Estimates

| Parameter | DF | Estimate |
|---|---|---|
| Intercept | 1 | 4.585200 |
| log_accom | 1 | 0.145127 |
| log_bath | 1 | 0.121667 |
| log_max | 1 | 0.004642 |

| Parameter Estimates | | |
|---|---|---|
| Parameter | DF | Estimate |
| log_bedsper | 1 | -0.029311 |
| n_bronx | 1 | -0.326507 |
| r_entire | 1 | 0.648439 |
| h_super | 1 | 0.032650 |
| h_profile | 1 | -0.058495 |
| h_verified | 1 | -0.020091 |
| b_futon | 1 | -0.087589 |
| b_real | 1 | 0.070705 |
| instant | 1 | -0.003420 |
| require_phone | 1 | 0.029076 |
| hcount_level2 | 1 | -0.034264 |
| hcount_level3 | 1 | 0.138494 |
| p_apart | 1 | -0.105948 |
| p_group2 | 1 | -0.328477 |
| p_house | 1 | -0.331994 |
| p_loft | 1 | 0.071987 |
| p_townhouse | 1 | -0.239362 |
| r_shared | 1 | -0.076879 |
| p_group1 | 1 | 0.115389 |
| poly_accom | 1 | 0.056143 |
| poly_guests | 1 | 0.013898 |
| poly_min | 1 | -0.047705 |
| poly_max | 1 | 0.015991 |
| poly_avail | 1 | 0.052545 |

| Score Information | |
|---|---|
| Input Data Set | WORK.TEST_FINAL |
| Output Data Set | WORK.TEST_SCORE |
| | |
| Number of Observations Read | 9617 |
| Number of Observations Scored | 0 |
| Number of Residuals Computed | 0 |
| | |
| Residual Sum of Squares | . |