

## Selection Method LASSO Using Cross Validation

### The GLMSELECT Procedure

Data Set	WORK.FORMODELANDPRED
Dependent Variable	price_doc
Selection Method	LASSO
Stop Criterion	Cross Validation
Choose Criterion	Cross Validation
Cross Validation Method	Random
Cross Validation Fold	5
Effect Hierarchy Enforced	None
Random Number Seed	1

Number of Observations Read	30446
Number of Observations Used	25452

Dimensions	
Number of Effects	38
Number of Parameters	38



## Selection Method LASSO Using Cross Validation

### The GLMSELECT Procedure

LASSO Selection Summary				
Step	Effect Entered	Effect Removed	Number Effects In	CV PRESS
0	Intercept		1	2.70406E17
1	full_sq		2	1.8253E17
2	kremlin_km		3	1.48771E17
3	swim_pool_km		4	1.44031E17
4	fitness_km		5	1.42929E17
5	timestamp		6	1.38614E17
6	raion_popul		7	1.37576E17
7	university_km		8	1.37472E17
8	kitch_sq		9	1.36704E17
9	big_road1_km		10	1.35648E17
10	X0_17_all		11	1.35623E17
11	floor		12	1.34608E17
12	hospital_beds_raion		13	1.34317E17
13	healthcare_centers_r		14	1.34153E17
14	public_transport_sta		15	1.34099E17
15	metro_km_walk		16	1.3409E17
16	big_market_km		17	1.3328E17
17	ice_rink_km		18	1.32011E17
18	build_count_wood		19	1.31699E17
19	university_top_20_ra		20	1.31628E17
20	big_road2_km		21	1.31606E17
21		X0_17_all	20	1.3159E17
22		public_transport_sta	19	1.31586E17
23	railroad_km		20	1.31351E17*
* Optimal Value of Criterion				

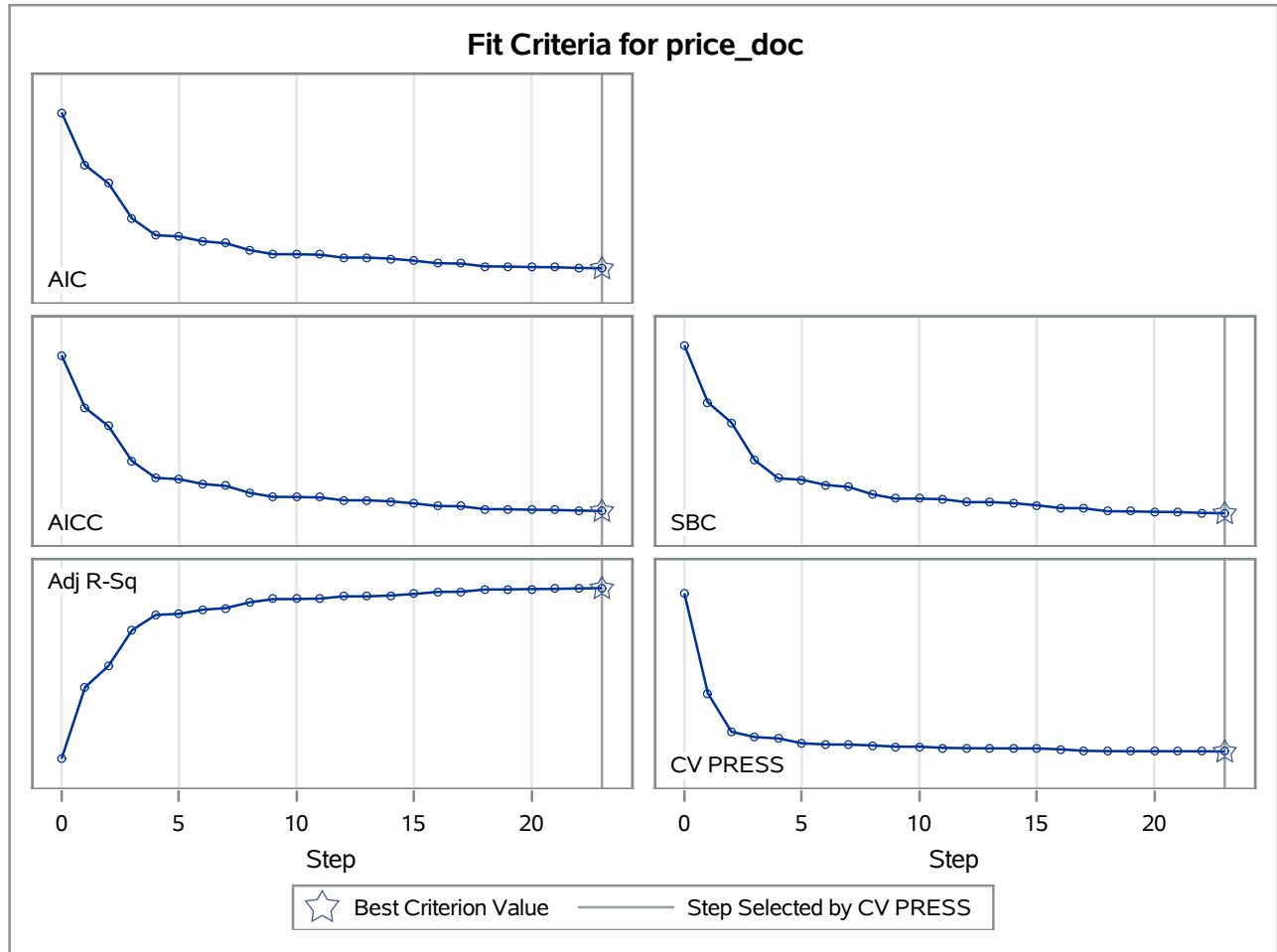
Selection stopped at a local minimum of the cross validation PRESS.

Stop Details				
Candidate For	Effect	Candidate CV PRESS		Compare CV PRESS
Entry	public_transport_sta	1.31352E17	>	1.31351E17

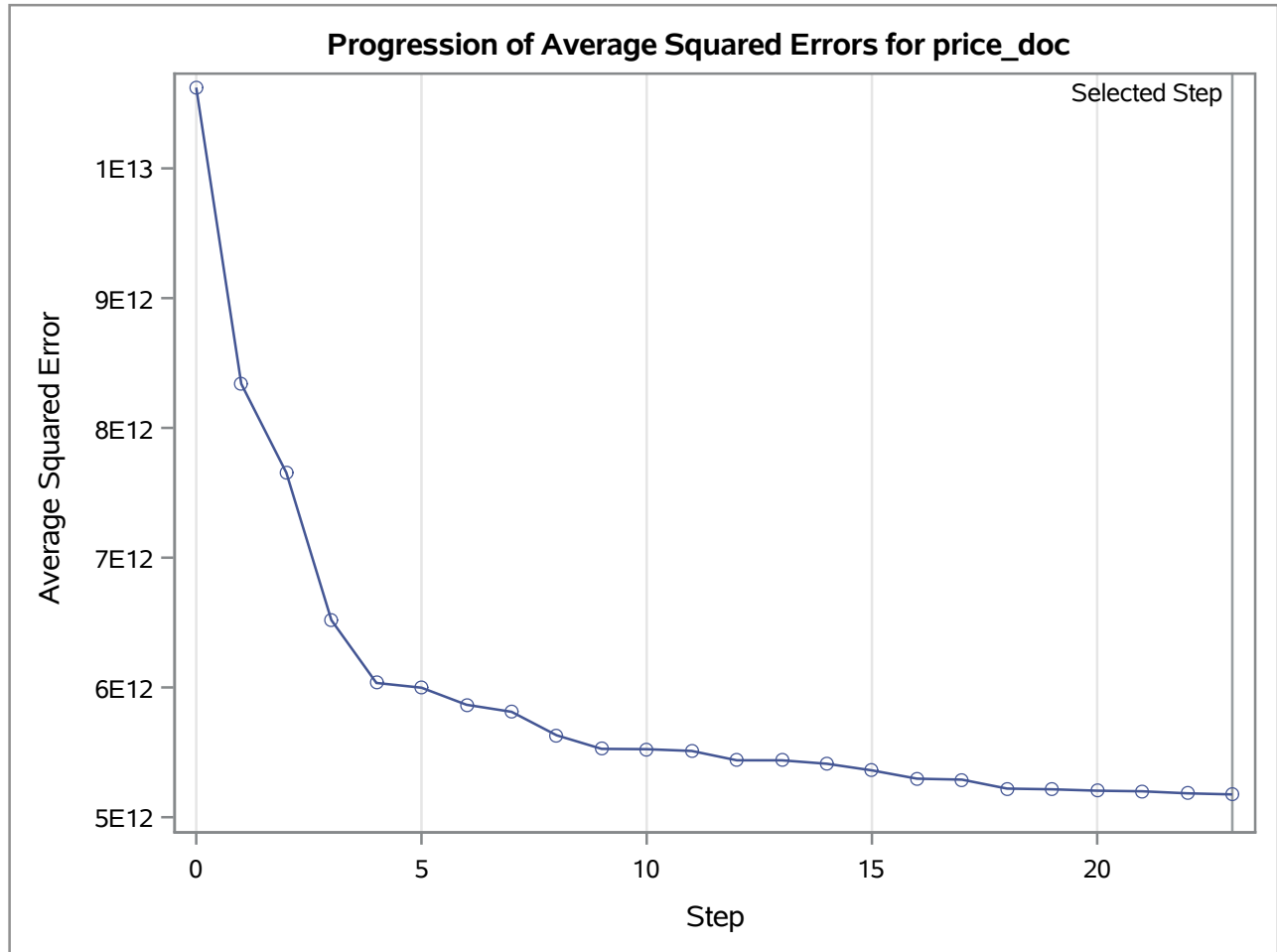


# Selection Method LASSO Using Cross Validation

## The GLMSELECT Procedure





**Selection Method LASSO Using Cross Validation****The GLMSELECT Procedure**



## Selection Method LASSO Using Cross Validation

### The GLMSELECT Procedure Selected Model

The selected model, based on Cross Validation, is the model at Step 23.

<b>Effects:</b>	Intercept timestamp full_sq floor hospital_beds_raion raion_popul kitch_sq healthcare_centers_r university_top_20_ra build_count_wood kremlin_km big_road1_km big_road2_km railroad_km big_market_km metro_km_walk ice_rink_km swim_pool_km fitness_km university_km
-----------------	--

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Value
Model	19	1.38631E17	7.296368E15	1408.69
Error	25432	1.317261E17	5.179542E12	
Corrected Total	25451	2.703571E17		

Root MSE	2275861
Dependent Mean	6782908
R-Square	0.5128
Adj R-Sq	0.5124
AIC	770600
AICC	770600
SBC	745309
CV PRESS	1.313514E17

Cross Validation Details			
Index	Observations		CV PRESS
	Fitted	Left Out	
1	20444	5008	2.5672E16
2	20319	5133	2.58299E16
3	20473	4979	2.63066E16
4	20224	5228	2.66356E16
5	20348	5104	2.69073E16
Total			1.31351E17



## Selection Method LASSO Using Cross Validation

### The GLMSELECT Procedure Selected Model

Parameter Estimates		
Parameter	DF	Estimate
Intercept	1	-47862642
timestamp	1	1195.997367
full_sq	1	107699
floor	1	35857
hospital_beds_raion	1	59.610384
raion_popul	1	3.120163
kitch_sq	1	37357
healthcare_centers_r	1	53150
university_top_20_ra	1	119827
build_count_wood	1	-629.871244
kremlin_km	1	-70606
big_road1_km	1	-123655
big_road2_km	1	-13514
railroad_km	1	11301
big_market_km	1	28137
metro_km_walk	1	-35773
ice_rink_km	1	55985
swim_pool_km	1	-111995
fitness_km	1	-164805
university_km	1	-42403