

Sensitivity Analysis on Trajectory with an Abundance of Zeros

Comparing Missings with Weighted Payments and zero payments
Problematic Residuals and Inferential Statistics with Zeros

Monthly Food and Leisure Costs
January through December Fake Data
SAS Proc Traj by Dr Jones
<https://www.andrew.cmu.edu/user/bjones/>

Trajectories are very robust with respect to zeros.

However statisticians should be cautious about inference.

With zeros standard errors and confidence intervals are very wide.

Individual Residuals are not symmetric or normal

Substituting missings for zeros and using weighted payments helps.

For large sample sizes distributions around means is better.

Sensitivity Analysis: Parameters and Standard Errors

Generate a three trajectory dataset with known trends

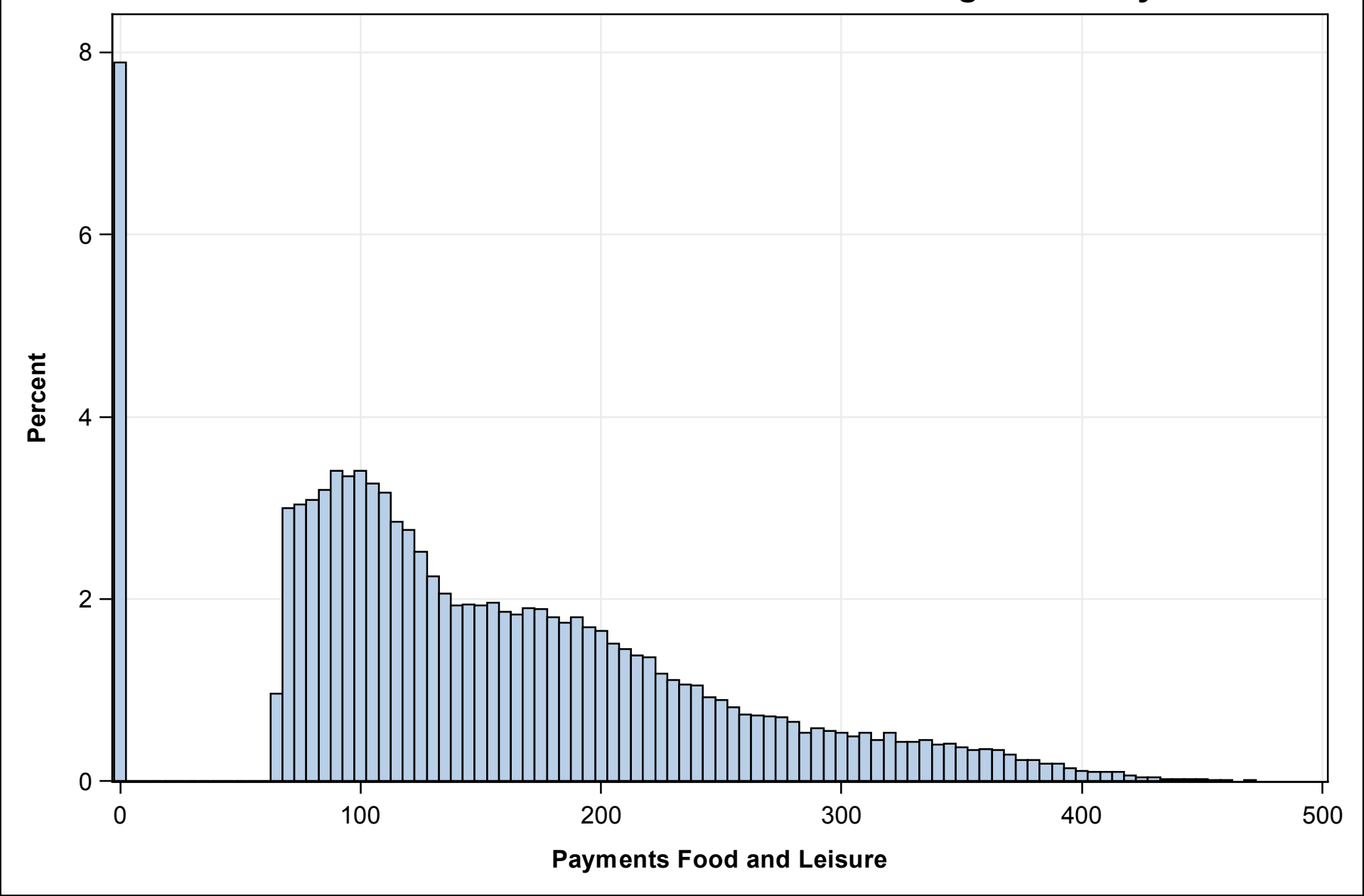
Three with 15% zeros and second set with 15% missing

$$f(\text{month}) = 76 + 2 * \text{month} + 0.1 * \text{month}^2 + \text{normal}(0, 15)$$

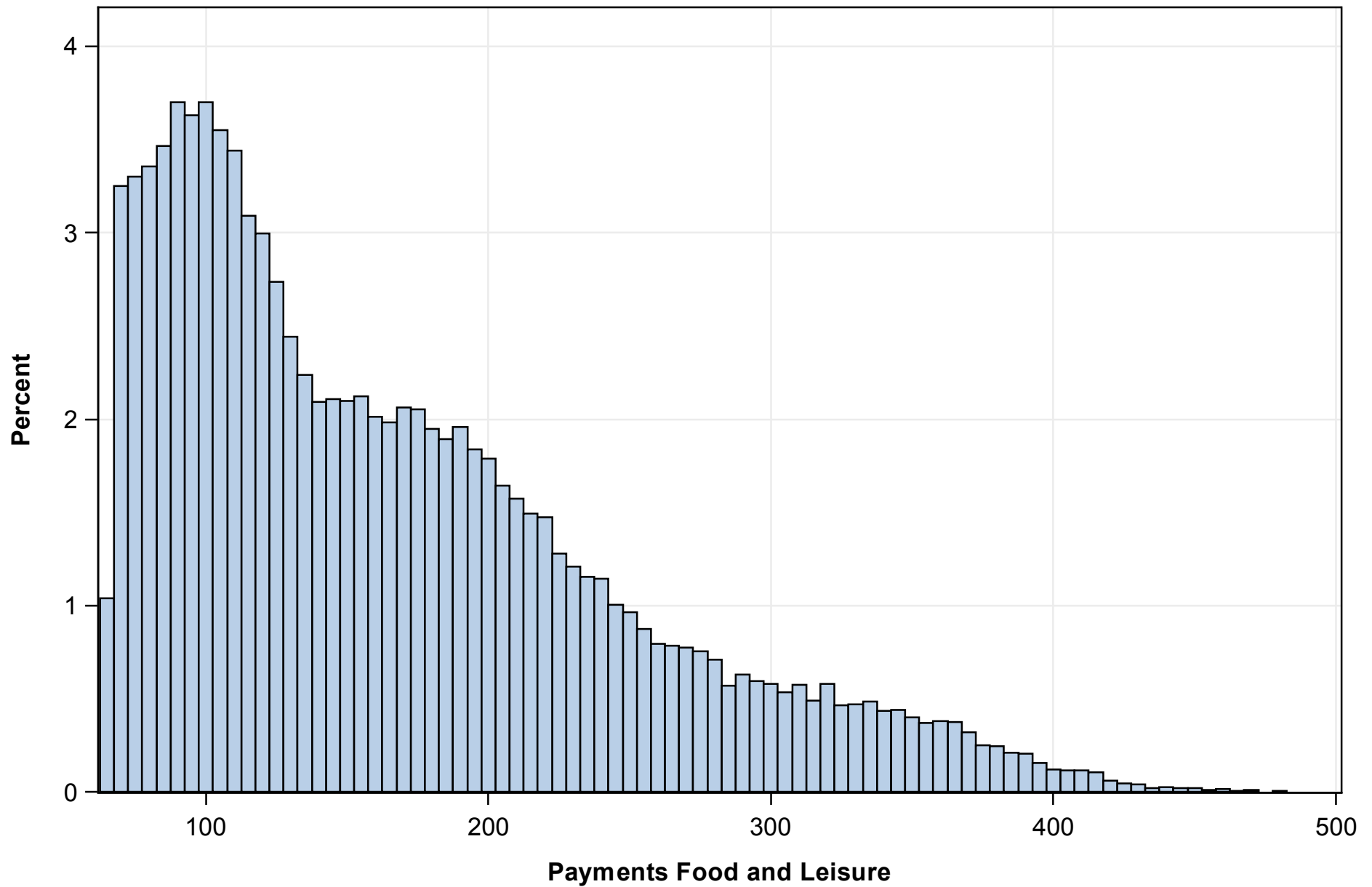
$$f(\text{month}) = 116 + 4 * \text{month} + 0.3 * \text{month}^2 + \text{normal}(0, 40)$$

$$f(\text{month}) = 146 + 6 * \text{month} + 1.0 * \text{month}^2 + \text{normal}(0, 40)$$

Slide 1200 Problematic Overall 12 Month Histogram of Payments



Slide 1300 Better Histogram Proc Traj ignores missing



```

*;
* PROC TRAJ MULTIPLE MODELS;
*;
%let cmpMdl=%sysfunc(compress(&mdl));

%macro cmmi_md1chk(mdl,x=);
%let cmpMdl=%sysfunc(compress(&mdl));

%pdfbeg(pdf=d:/taj/pdf/&pgm._tbl1510&x.pdf);
proc traj data = taj.taj_simulate&x

outplot =    taj.taj_md1Plot12_&cmpMdl&x
outest  =    taj.taj_md1Est12_&cmpMdl&x

outstat =    taj.taj_md1Stat12_&cmpMdl&x
out = taj.taj_md1Detail12_&cmpMdl&x ci95M;

model order&cmpmdl;
id id;
var _1-_12 ;
indep t1-t12;
order &mdl;

min 0;
max 500;
model cnorm;
run;quit;
%pdfend;

%mend cmmi_md1chk;
%*cmmi_md1chk(2 2 2);
%*cmmi_md1chk(2 2 2,x=X);
,

```

Model (222) Three Quadratics With Zeros

Maximum Likelihood Estimates

Model: Censored Normal (CNORM)

Group	Parameter	Estimate	Standard	T for H0:	Prob > T
			Error	Parameter=0	
1	Intercept	67.52848	1.52777	44.201	0.0000
	Linear	1.77536	0.53766	3.302	0.0010
	Quadratic	0.18853	0.04024	4.686	0.0000
2	Intercept	124.07592	1.58282	78.389	0.0000
	Linear	5.25813	0.56155	9.364	0.0000
	Quadratic	1.21352	0.04207	28.846	0.0000
3	Intercept	99.73415	1.61344	61.814	0.0000
	Linear	4.42900	0.57170	7.747	0.0000
	Quadratic	0.38208	0.04286	8.915	0.0000
	Sigma	57.21588	0.18814	304.115	0.0000
Group membership					
1	(%)	35.64032	0.73588	48.432	0.0000
2	(%)	31.97513	0.70096	45.616	0.0000
3	(%)	32.38455	0.72385	44.740	0.0000
BIC=-282854.2 (N=54000) BIC=-282839.3 (N=4500) AIC=-282800.8 L=-282788.8					

Model (222) Three Quadratics With Missings

Maximum Likelihood Estimates

Model: Censored Normal (CNORM)

Group	Parameter	Estimate	Standard	T for H0:	Prob > T
			Error	Parameter=0	
1	Intercept	83.74628	0.95634	87.569	0.0000
	Linear	0.79796	0.33294	2.397	0.0165
	Quadratic	0.14538	0.02470	5.885	0.0000
2	Intercept	121.19329	0.96161	126.031	0.0000
	Linear	2.82945	0.33501	8.446	0.0000
	Quadratic	0.36342	0.02487	14.615	0.0000
3	Intercept	146.09933	0.95210	153.449	0.0000
	Linear	5.92656	0.33236	17.832	0.0000
	Quadratic	1.00975	0.02469	40.896	0.0000
	Sigma	33.57330	0.10648	315.289	0.0000
Group membership					
1	(%)	33.48212	0.70444	47.530	0.0000
2	(%)	33.18438	0.70287	47.213	0.0000
3	(%)	33.33349	0.70281	47.429	0.0000
BIC=-250351.3 (N=49740) BIC=-250336.8 (N=4500) AIC=-250298.4 L=-250286.4					

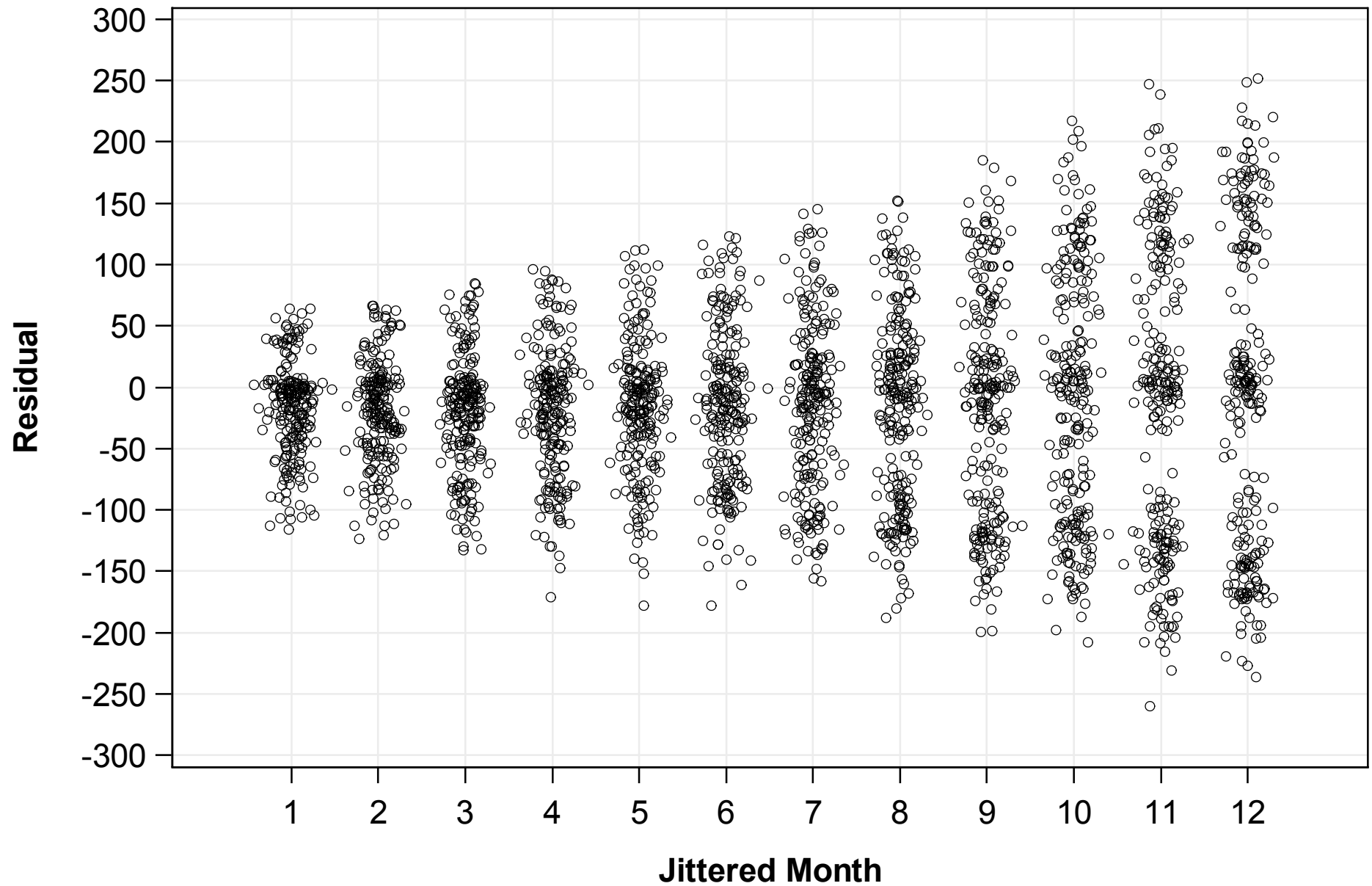
Figure 1900 Sensitivity Analysis on Model Parameters
Using 15% Zero or Missing

Parameter	Theory	Missing	Zeros
Intercept	76	83.74	67.52
Linear	2	0.8	1.78
Quadratic	0.1	0.15	0.19
Intercept	116	121.2	24.08
Linear	4	2.83	5.26
Quadratic	0.3	0.36	1.21
Intercept	146	146.09	99.73
Linear	6	5.93	4.42
Quadratic	1	1.01	0.39

Figure 2000 Sensitivity Analysis on Model Standard Errors
Using 15% Zero or Missing

Standard Error	Missing	Zeros
Intercept	0.95	1.53
Linear	0.33	0.54
Quadratic	0.02	0.04
Intercept	0.96	1.58
Linear	0.33	0.56
Quadratic	0.02	0.04
Intercept	0.95	1.61
Linear	0.33	0.57
Quadratic	0.02	0.04

Figure 2100 Missings Time jittered Residual Plot Three Quadratics(222)



**Figure 2200 Missings Time jittered Standardized Residual Plot Three
Quadratics(222)**

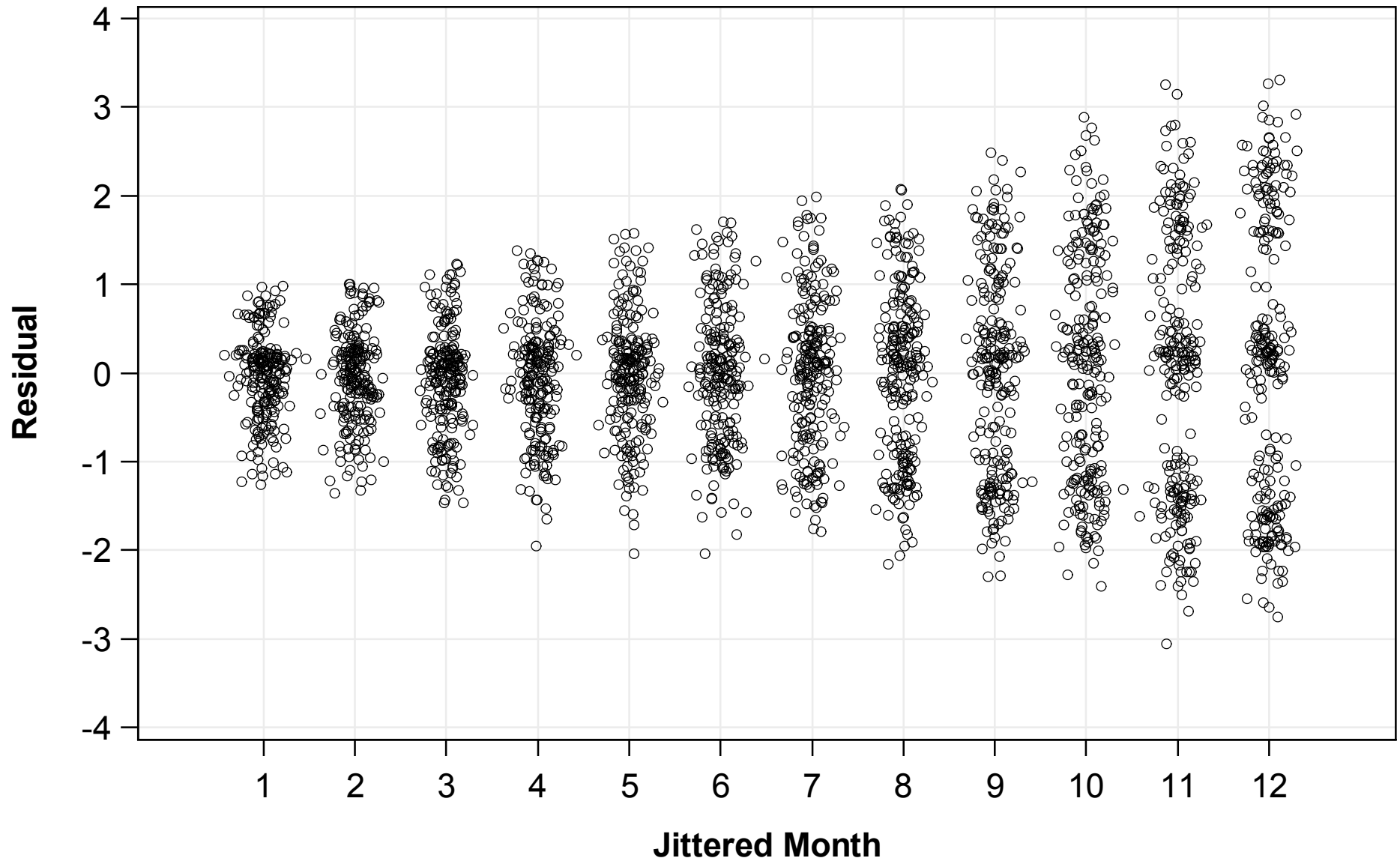
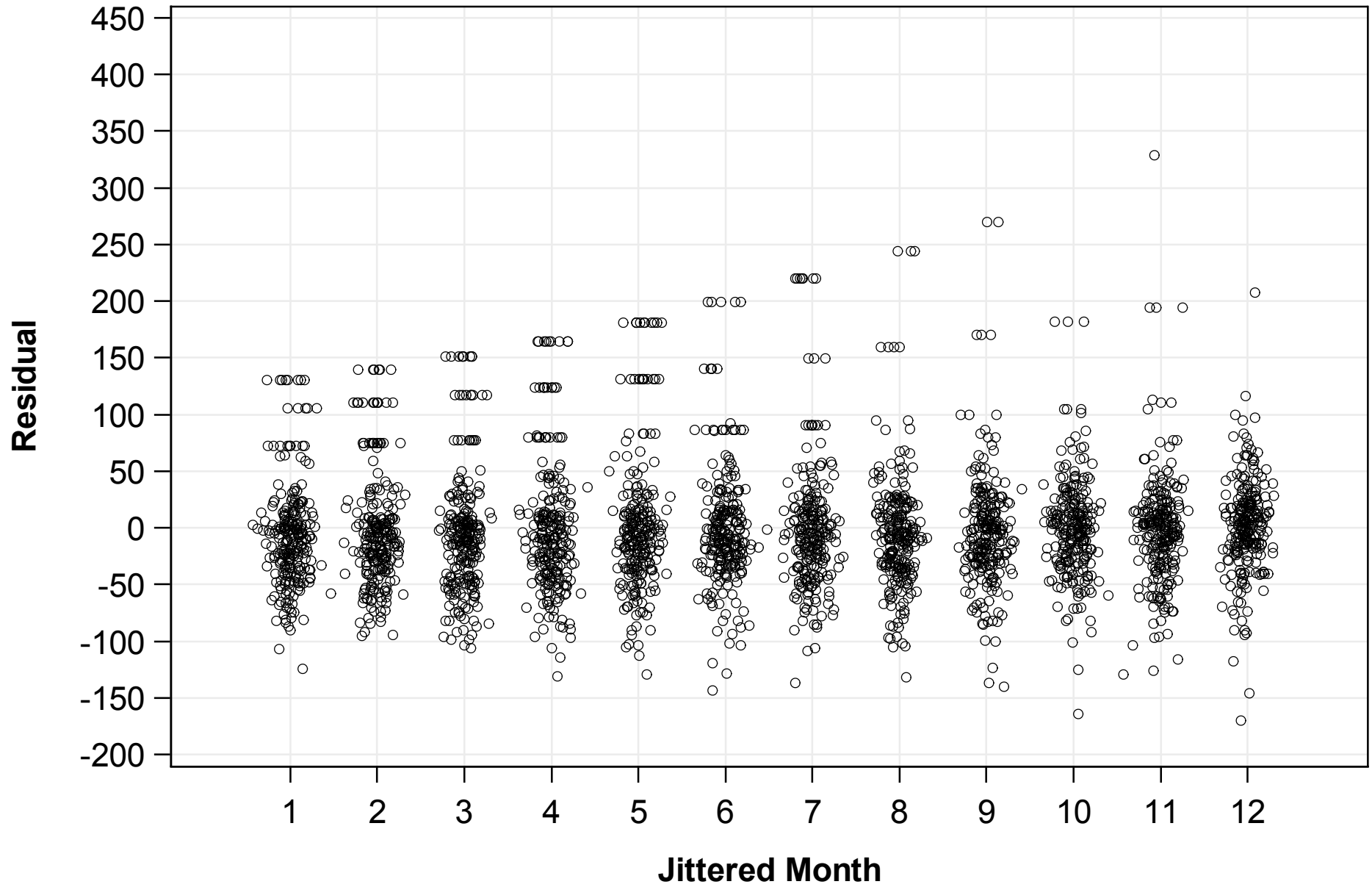


Figure 2400 With Zeros: Time jittered Residual Plot Three QuadraticsModel



**Figure 2400 With Zeros: Time jittered Standardized Residual Plot Three
QuadraticsModel**

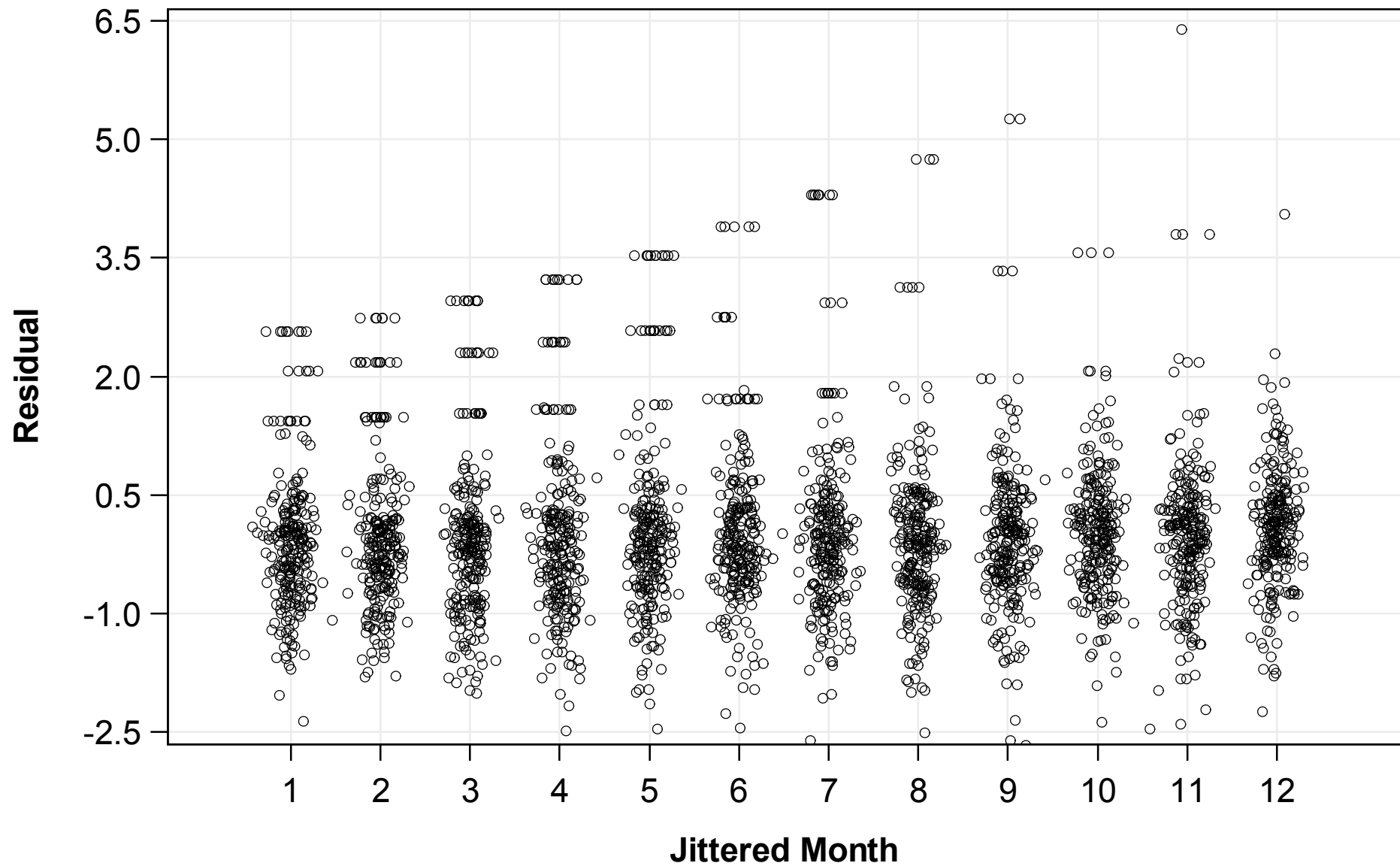
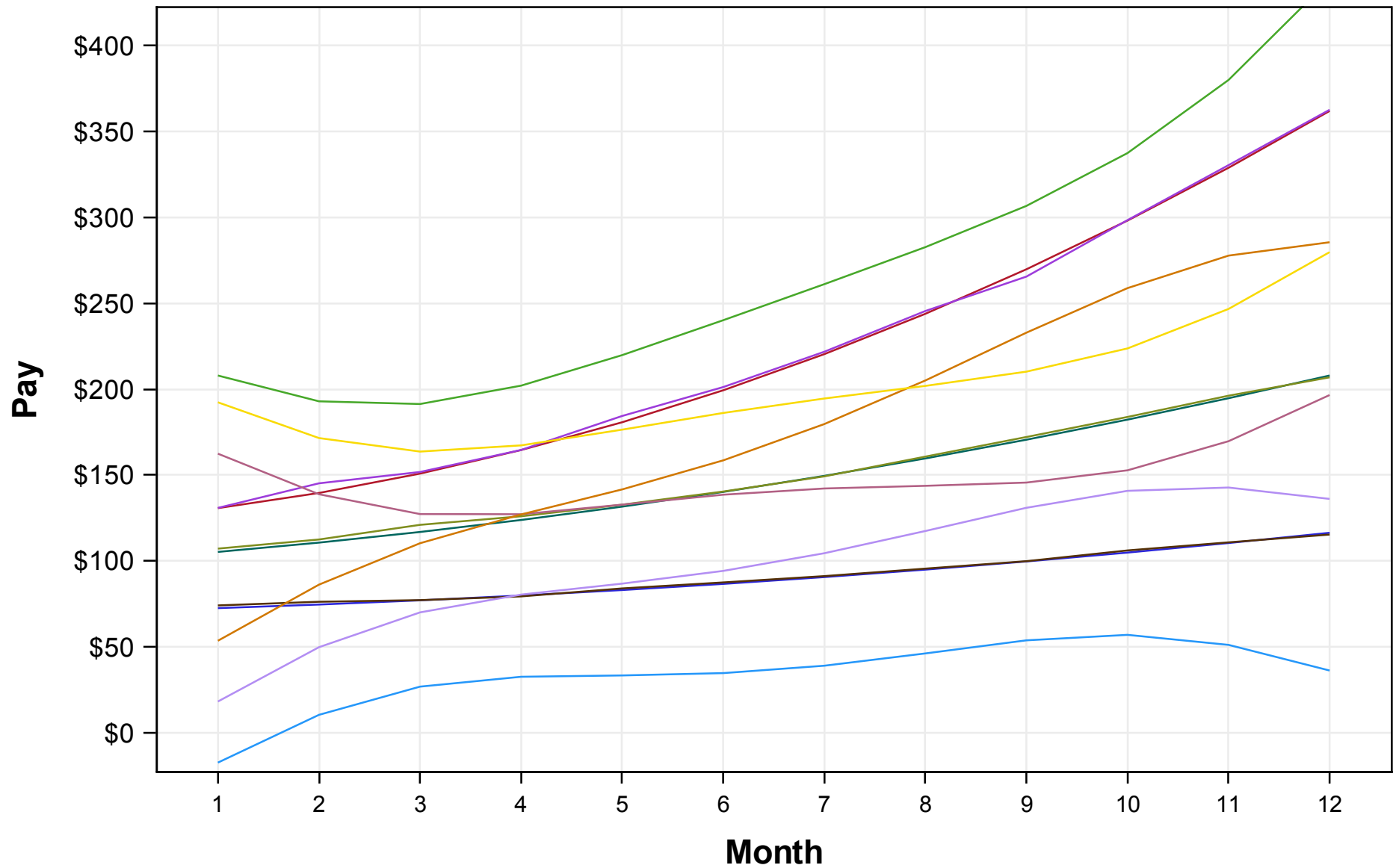


Figure 2500 With Zeros: Fit Analysis and Residuals

Model 222 Three Quadratics

Figure 2600 Model 222 With Zeros Best Fit Trajectories Payments
Individual Obsevation 95% CI



**Figure 2700 Model 222 With Missing Best Fit Trajectories Three Quadratic
Individual Obsevation 95% CI**

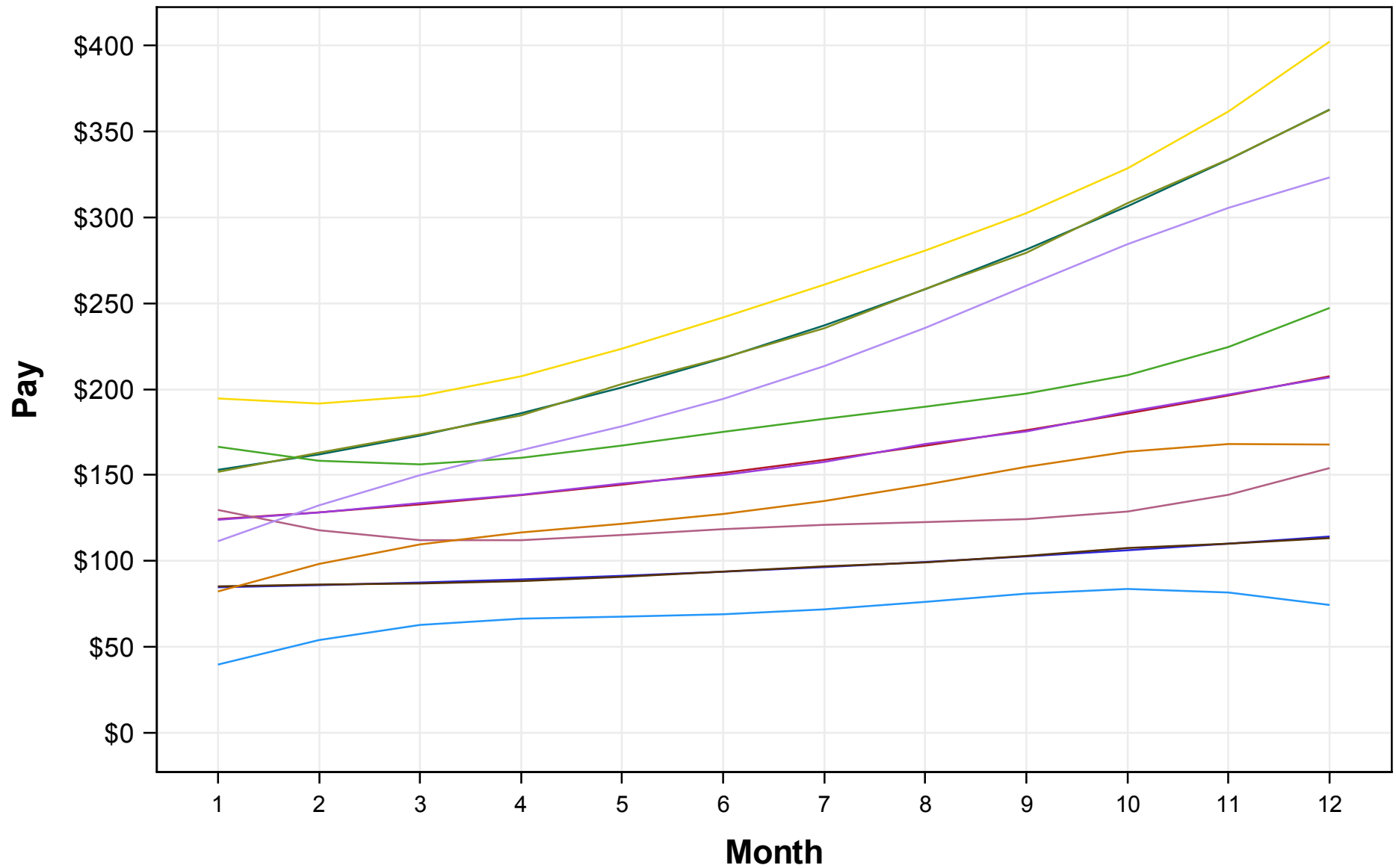


Figure 2800 Classification a little more Accurate for the missing model

	Miss Classification Probabilities				
Models	Classified	Trajectory	Theoretical Trajectory	Frequency	Probability
Missings	Miss Classified	1. Low	2. Moderate	4	0.0027
Zeros	Miss Classified	1. Low	2. Moderate	80	0.0533
		2. Moderate	3. High	58	0.0387