AN EMPIRICIAL ASSESSMENT OF FIXED AND RANDOM PARAMETER LOGIT MODELS USING CRASH- AND NON-CRASH-SPECIFIC INJURY DATA

Panos Anastasopoulos and Fred Mannering

MOTIVATION

- Traditional crash-severity modeling uses detailed data gathered after a crash has occurred (number of vehicles involved, age of occupants, weather conditions at the time of the crash, types of vehicles involved, crash type, occupant restraint use, airbag deployment, etc.).
- But forecasting with post-crash data is near impossible
- Consider fixed vs. random parameters AND crash-specific and non-crash-specific data

METHODOLOGY

- Consider injury outcomes of no-injury, injury and fatality.

- Estimate fixed and random parameter models using individual crash data

- Estimate fixed and random parameter models using injury proportions data (as in Milton et al.)

DATA

- Crash data from rural interstate highways in Indiana were collected for a 5 year period (1995–1999 inclusive).
- The initial data sample consists of 5,795 police-reported crashes that occurred on 231 freeway segments.
- Of these 5,795 police-reported crashes, 4,658 resulted in no-injury, 1,084 in injury, and 53 resulted in fatality.
- Consider severity of most severely injured person in the crash.

Goodness-of-fit measures for the random and fixed parameter logit models

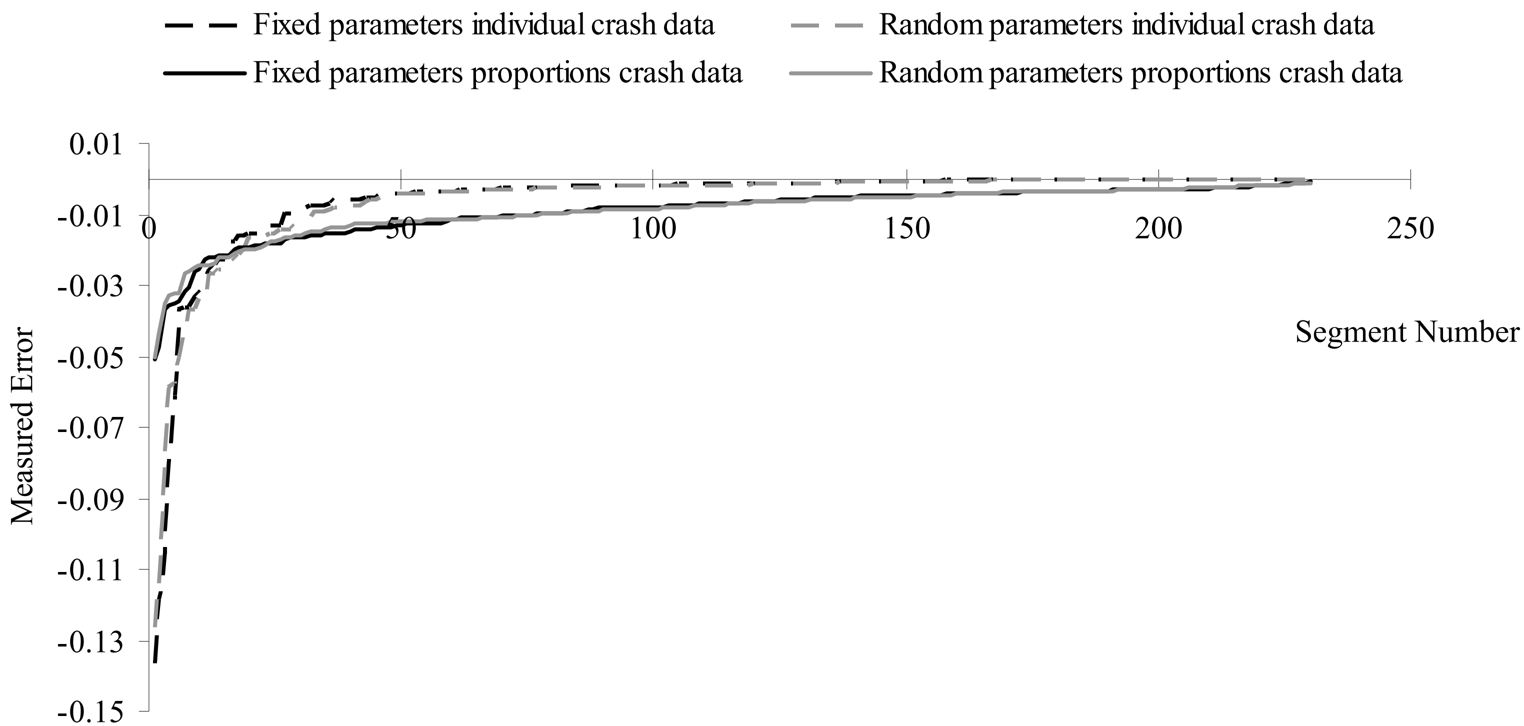
	Individual Crash Data Logit Models		Proportions Crash Data Logit Models	
	Random Parameters	Fixed Parameters	Random Parameters	Fixed Parameters
Number of parameters	26	24	16	14
Log-likelihood at zero, $LL(\mathbf{0})$	-774.14	-774.14	-5683.11	-5683.11
Log-likelihood at convergence, $LL(\boldsymbol{\beta})$	-250.33	-266.13	-2718.71	-2788.40
McFadden pseudo-ρ^2	0.677	0.656	0.522	0.509
McFadden adjusted pseudo-ρ^2	0.643	0.625	0.519	0.507
Likelihood-Ratio Test	Random versus Fixed Parameters		Random versus Fixed Parameters	
$X^2 = -2[LL(\boldsymbol{\beta}_{random}) - LL(\boldsymbol{\beta}_{fixed})]$	15.81		69.69	
Degrees of freedom	2		2	
Critical X^2 (0.999 level of confidence)	13.82		13.82	
Number of observations	5795		231	

Observed vs. model-predicted probabilities

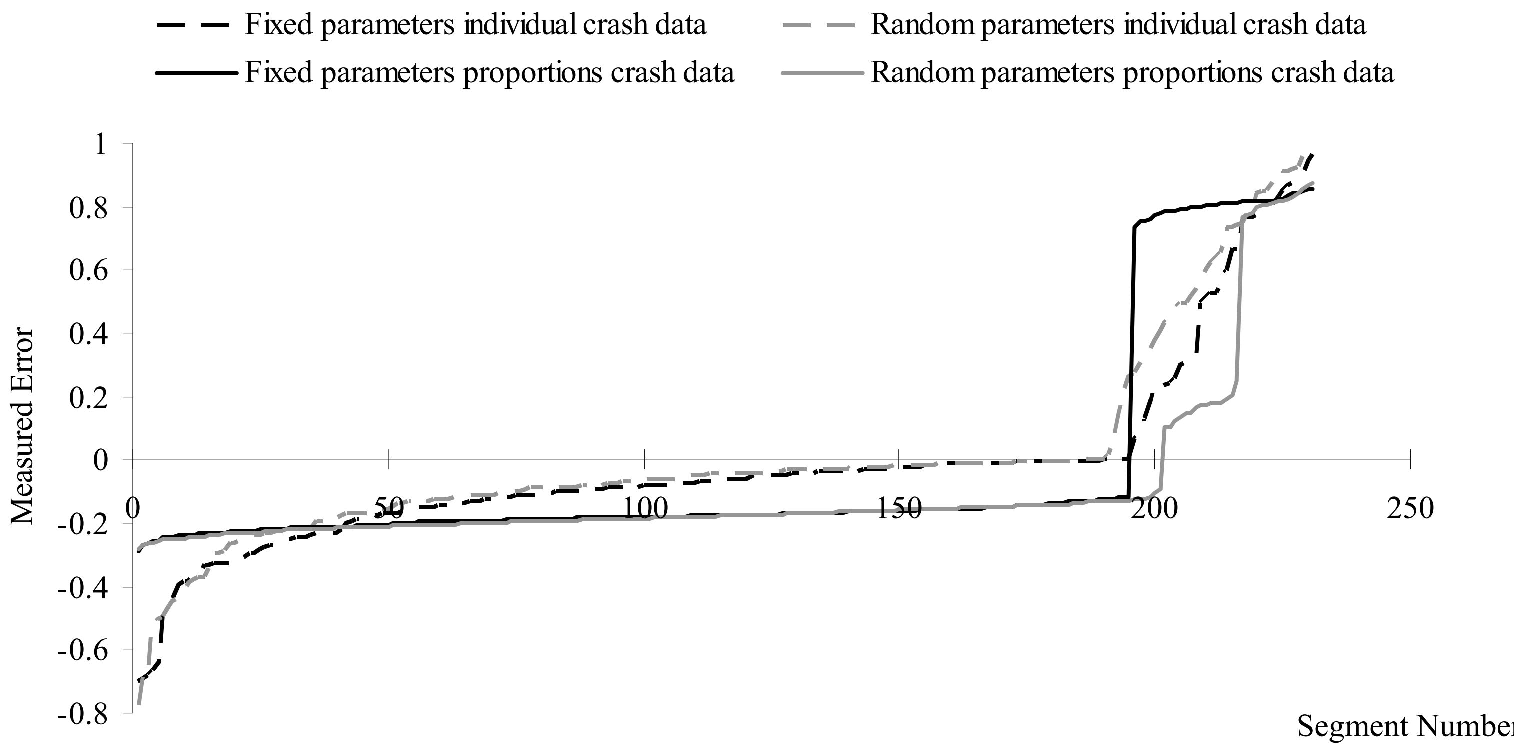
<table>
<tr><td rowspan="2"></td><td colspan="6">Observed Versus Model-Predicted Probabilities</td></tr>
<tr><td colspan="2">Fatal</td><td colspan="2">Injury</td><td colspan="2">No-Injury</td></tr>
<tr><td>Model Type</td><td>Observed</td><td>Predicted</td><td>Observed</td><td>Predicted</td><td>Observed</td><td>Predicted</td></tr>
<tr><td>Fixed Parameters individual crash data</td><td rowspan="2">0.0071</td><td>0.0094</td><td rowspan="2">0.1716</td><td>0.1787</td><td rowspan="2">0.8213</td><td>0.8119</td></tr>
<tr><td>Random Parameters individual crash data</td><td>0.0071</td><td>0.1688</td><td>0.8241</td></tr>
<tr><td>Fixed Parameters proportions crash data</td><td rowspan="2">0.0063</td><td>0.0120</td><td rowspan="2">0.1418</td><td>0.2219</td><td rowspan="2">0.8519</td><td>0.7660</td></tr>
<tr><td>Random Parameters proportions crash data</td><td>0.0089</td><td>0.1850</td><td>0.8062</td></tr>
</table>

Observed vs. model-predicted injury-severity proportions of the 231 road segments predicted with the individual crash data and the proportions crash data

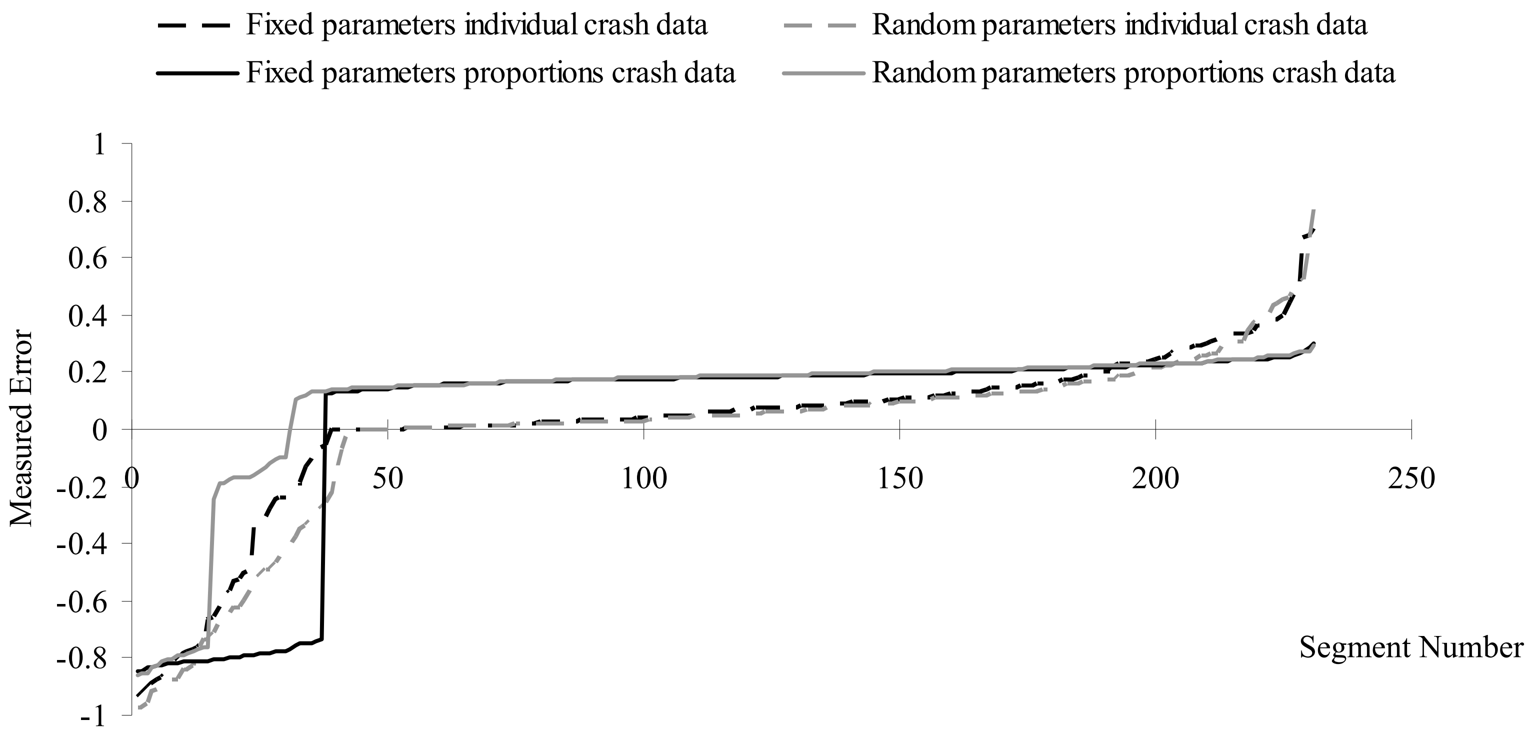
	Observed Versus Model-Predicted Injury-Severity Proportions of the 231 Road segments					
	Fatal		Injury		No-Injury	
Model Type	Observed	Predicted	Observed	Predicted	Observed	Predicted
Fixed Parameters individual crash data	0.0063	0.0080	0.1418	0.1762	0.8519	0.8157
Random Parameters individual crash data		0.0063		0.1555		0.8382
Fixed Parameters proportions crash data		0.0120		0.2219		0.7660
Random Parameters proportions crash data		0.0089		0.1850		0.8062



Measured error (observed minus model-predicted proportions) of fatalities for 226 road segments (ordered from most negative to most positive). Note that 5 of the 231 roadway segments had high positive measurement error and were not included in this figure to allow more scale detail.



Measured error (observed minus model-predicted proportions) of injuries for the 231 road segments (ordered from most negative to most positive)



Measured error (observed minus model-predicted proportions) of no-injuries for the 231 road segments (ordered from most negative to most positive)

Models' ability to identifying critical needs:

- Identify the 5% of roadway segments (12) that have the highest fatality proportions predicted by the random parameters model on individual crash data.

- The random parameters model on proportions data identifies 11 of these 12 segments with it predicted fatality proportions.

Conclusions

- Random parameter models give better fit and better prediction
- Models based on individual crash data gives better fit
- However, proportions crash-data models provide reasonable accuracy for practical use.