Supplemental Online Appendix

for

Leadership PAC Donations and Party Status: A Technical and Theoretical Extension

Hurdle or Zero-Inflation?

There could be some debate over whether a zero-inflated model or a hurdle model is the better choice for modeling the excess zeroes in the donation distribution. The difference between the two is how they conceptualize the $0 contributions. A hurdle model completely separates the $0 contributions from all positive contributions. So although the $0 contributions are modeled separately, if a race crosses the hurdle, it *must* receive some positive contribution. The theoretical process modeled by this empirical assumption is

1. First, LPACs decide *all* races to which they are not going to give any donations. Practically, when an LPAC thinks about distributing money, the first choice it makes is simultaneously deciding all races that will not receive a donation during the campaign cycle, which might be particularly difficult for races that are "on the fence" of receiving a donation. But all races that receive no donation are modeled together.

2. Second, of the remaining races, LPACs decide how much money to donate. All races remaining must receive some positive donation.

A zero-inflated model allows for a mixing of different types of $0 contributions. Some donations are $0 because the relevant race is *never* going to receive a donation; other donations are $0 because *during this cycle* the race happened not to receive a donation. The theoretical process modeled by this empirical assumption is

1. First, LPACs decide races to which they would *never*, under any circumstances, give a donation (the candidate is too extreme, the race is too unwinnable). This might be particularly easy for an LPAC to examine the electoral landscape for races that are unconditionally uninteresting to the LPAC. These races are eliminated from contention for a donation.

2. Second, of the remaining races, LPACs decide how much money to donate. Here, though, *some races could still not receive money.* Perhaps they would if the LPAC had more money to distribute, or if other (competing) races weren't more strategically valuable during this particular cycle. But during this cycle, we could observe *either* a positive contribution *or* a $0 contribution.

Which of these models better reflects the theoretical decisionmaking process of the LPAC is an open question, best answered by qualitative data from LPAC leaders, which we do not have here. We can, however, examine whether they perform similarly in the data. To that end, Table A.1 presents seven models. The first (Model 1) is the model in the main paper: the zero-inflated model with the theoretical triple interaction. The second (Model 2) recasts Model 1 as a hurdle model, rather than a zero-inflated model, to test whether there is an appreciable difference between the two. The third (Model 3) recasts Model 1 as a plain negative binomial regression with no zero-inflation (like ABLR's model from their main text) to test whether accounting for the excess zeroes is a statistical improvement.

The fourth and sixth models (Model 4 and Model 6) replicate ABLR's original hurdle models from the Supplemental Appendix of their (2017) work. The fifth and seventh models (Model 5 and Model 7) replicate ABLR's original hurdle models as zero-inflated models, again to test whether there is a preference for the hurdle specification or the zero-inflated specification.

The most immediate observation from Table A.1 is Model 3—the plain negative binomial regression, which is the only model in Table A.1 not to account for the excess zeroes—is clearly the worst choice. The AIC is much higher than any of the other models. This is independently important, as the plain negative binomial regression is reflective of the class of model ABLR interpret in their main text. A more critical look at the zero-generating process is warranted.

Table A.1: Alternative Specifications of the Main Model

Variable	(1) Zero-Infl., Interactive	(2) Hurdle, Interactive	(3) Negative Bin., Interactive	(4) Hurdle, Distance	(5) Zero-Infl., Distance	(6) Hurdle, Rating	(7) Zero-Infl., Rating
	Count Model						
Intercept	9.849*	9.849*	10.684*	9.396*	9.396*	9.941*	9.941*
	(0.140)	(0.140)	(0.422)	(0.096)	(0.096)	(0.114)	(0.114)
Ideo. Distance	-0.184	-0.184	-2.006*	-0.202	-0.202	-	-
	(0.334)	(0.334)	(0.501)	(0.137)	(0.137)		
Expected Gain	0.312*	0.312*	0.196	0.410*	0.410*	0.338*	0.338*
	(0.055)	(0.055)	(0.167)	(0.056)	(0.056)	(0.055)	(0.055)
Incumbent	0.501*	0.501*	0.896*	-0.028	-0.028	0.138*	0.138*
	(0.126)	(0.126)	(0.305)	(0.055)	(0.055)	(0.056)	(0.056)
Margin	0.028*	0.028*	0.040*	0.031*	0.031*	0.028*	0.028*
	(0.003)	(0.003)	(0.007)	(0.003)	(0.003)	(0.003)	(0.003)
Minority Party	-0.148	-0.148	0.361	-0.318*	-0.318*	-0.337*	-0.337*
	(0.135)	(0.135)	(0.325)	(0.053)	(0.053)	(0.052)	(0.052)
House Rating	-0.271*	-0.271*	-0.861*	-	-	-0.269*	-0.269*
	(0.032)	(0.032)	(0.109)			(0.032)	(0.032)
Republican	0.102	0.102	-0.026	0.043	0.043	0.163*	0.163*
	(0.053)	(0.053)	(0.159)	(0.053)	(0.053)	(0.051)	(0.051)
Mino. * Dist.	0.286	0.286	-0.819	-	-	-	-
	(0.432)	(0.432)	(0.679)				
Inc. * Dist	-0.507	-0.507	0.514	-	-	-	-
	(0.387)	(0.387)	(0.765)				
Inc. * Mino.	-0.482*	-0.482*	-1.211*	-	-	-	-
	(0.164)	(0.164)	(0.424)				
Inc. * Mino. * Dist.	0.333	0.333	1.197	-	-	-	-
	(0.540)	(0.540)	(1.099)				
Log(theta)	0.466*	0.466*	0.094	0.400*	0.400*	0.448*	0.448*
	(0.037)	(0.037)	(0.003)	(0.037)	(0.037)	(0.037)	(0.037)
	Zero Model						
Intercept	-6.126*	6.126*	-	-0.089	0.089	5.867*	-5.867*
	(0.525)	(0.525)		(0.172)	(0.172)	(0.482)	(0.482)
Ideo. Distance	1.462*	-1.462*	-	-1.673*	1.673*	-	-
	(0.663)	(0.663)		(0.216)	(0.216)		
Expected Gain	0.368*	-0.368*	-	-0.222*	0.222*	-0.408*	0.408*
	(0.121)	(0.121)		(0.107)	(0.107)	(0.119)	(0.119)
Incumbent	-1.172*	1.172*	-	0.162	-0.162	0.866*	-0.866*
	(0.265)	(0.265)		(0.094)	(0.094)	(0.109)	(0.109)
Margin	-0.019*	0.019*	-	0.034*	-0.034*	0.021*	-0.021*
	(0.005)	(0.005)		(0.005)	(0.005)	(0.005)	(0.005)
Minority Party	-1.129*	1.129*	-	-0.070	0.070	-0.107	0.107
	(0.313)	(0.313)		(0.096)	(0.096)	(0.106)	(0.106)
House Rating	2.407*	-2.407*	-	-	-	-2.400*	2.400*
	(0.162)	(0.162)				(0.160)	(0.160)
Republican	-0.364*	0.364*	-	0.075	-0.075	0.531*	-0.531*
	(0.118)	(0.118)		(0.103)	(0.103)	(0.113)	(0.113)
Mino. * Dist.	2.289*	-2.289*	-	-	-	-	-
	(0.900)	(0.900)					
Inc. * Dist	-0.109	0.109	-	-	-	-	-
	(0.773)	(0.773)					
Inc. * Mino.	1.498*	-1.498*	-	-	-	-	-
	(0.363)	(0.363)					
Inc. * Mino. * Dist.	-2.410*	2.410*	-	-	-	-	-
	(1.082)	(1.082)					
N	2160	2160	2160	2160	2160	2160	2160
AIC	29329.71	29329.71	32438.88	29932.39	29932.39	29423.88	29423.88

*$p < 0.05$.
AIC: Akaike information criterion.
Models estimated with all observations.
The unconditional hurdle models (just distance [4] or just House rating [6]) are replications of ABLR's Appendix.

The second observation from Table A.1 is that the hurdle models and the zero-inflated models are virtually equivalent, given the estimation results. The pairs of identical specifications, varying only the choice of the zero-inflation or hurdle (that is, Models 1 and 2, Models 4 and 5, and Models 6 and 7), are virtually identical in their estimated coefficients as well as their AIC. This implies the two distinct theoretical data-generating processes outlined above are observationally equivalent in the set of elections for which we have data. Given this equivalence, I interpret the zero-inflated model (Model 1) in the main text, but the results hold either way.

Lastly, Table A.1 sheds light on how the theoretical interaction in the main text improves our explanation over ABLR's original hurdle models. In their Supplemental Appendix, ABLR offer Models 4 and 6 (replicated in Table A.1) to account for the excess zeroes, but these models assume that the theoretical predictors are unconditional. Comparing Models 4 and 6 to the interactive specification, we observe the AIC of Model 1 (the triple interactive zero-inflated model) is lower, indicating a preference for the theoretical interaction. The worst of the three models is ABLR's hurdle model using *distance* (Model 4), followed by ABLR's hurdle model using *House rating* (Model 6), and finally the best-performing model is the triple interaction (Model 1). This preference for Model 1 is reinforced by the statistical significance of the triple interaction itself, as well as the constituent two-way interactions, especially in the zero equation.

We can get a visual sense of this preference by examining the predictions and errors of the models. To that end, Figure A.1 plots the predicted versus actual contributions from Models 1 and 2 (to demonstrate that the zero-inflated and hurdle specifications produce similar inferences) alongside ABLR's predictions from Models 4 and 6. The top row contains the predictions versus actual donations from the two triple-interactive models (Models 1 and 2). The bottom-left panel contains the predictions versus actual donations from ABLR's hurdle model using only distance (Model 4); the bottom-right panel contains the predictions versus

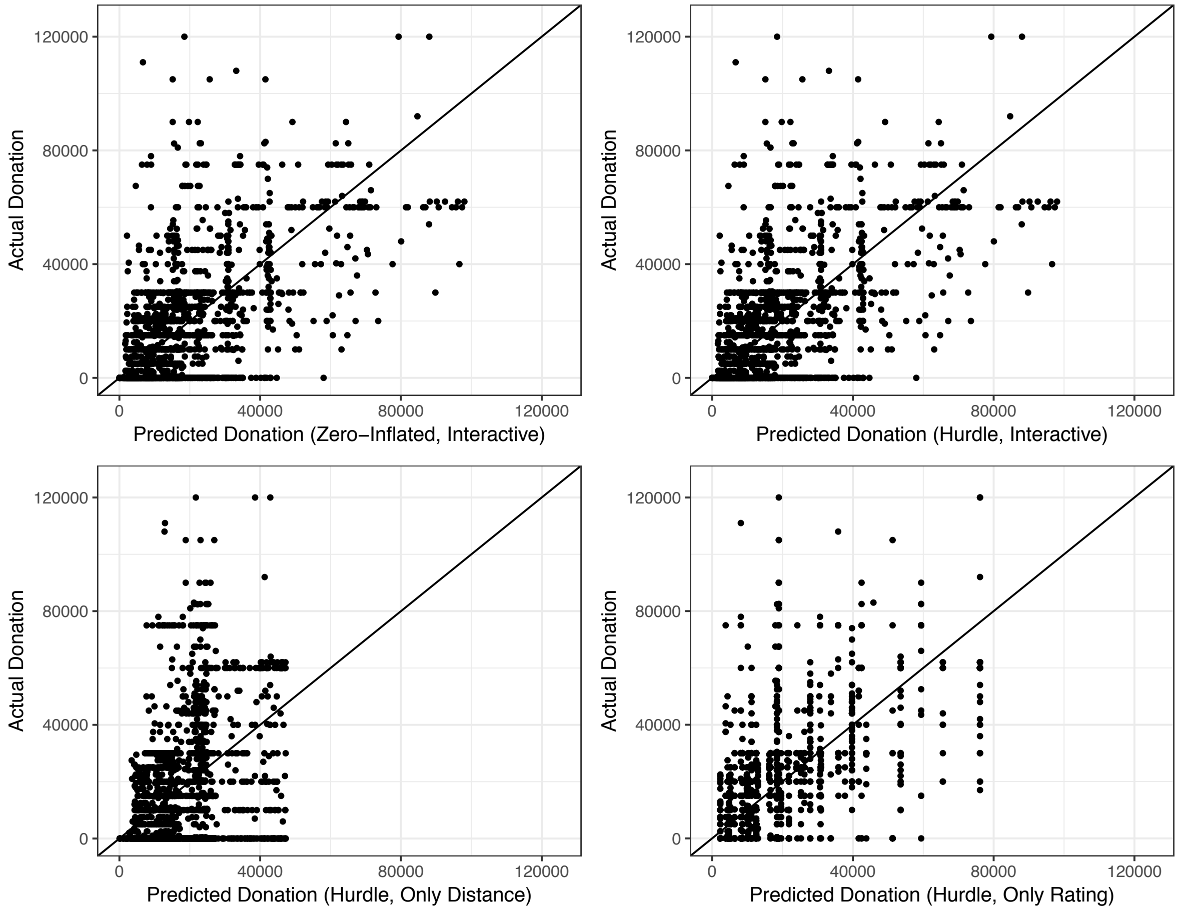

Figure A.1: Prediction Comparisons: Interactive Model (Zero-Inflated and Hurdle) Versus ABLR Original Hurdle Models.

actual donations from ABLR's hurdle model using only House rating (Model 6). Figure A.1 omits the two multi-million dollar contributions discussed in the main text to help keep the axes compact and interpretable.

Start with the bottom-left panel. The hurdle model using only ideological distance makes no predictions for donations over \$50,000: in fact, the maximum predicted donation is about \$47,250. This is well short of the range of observed donations which routinely measure over \$50,000. The bottom-right panel, the hurdle model using only House rating, performs better, but it too caps predictions, this time at about \$76,100. Both of the interactive models in the top row allow for more variation on the x-axis, better fitting the observed donating behavior by better matching the range of observed values.

Lastly, we can look at the errors themselves (actual donation minus predicted donation) for each of the four models. These errors are presented in Figure A.2. Again, Figure A.2 omits the two multi-million dollar contributions discussed in the main text to help keep the axes compact and interpretable. Begin with the bottom-left panel. The hurdle model using only ideological distance (Model 4) has an appreciably wider spread of errors than any of the other three models. When we examine ABLR's hurdle model using only House rating (Model 6) in the bottom-right panel, we find that the errors are on average closer to zero, but are more likely to be negative than positive (echoing the low maximum prediction observed in Figure A.1). In comparison, the errors from both the interactive specifications (Models 1 and 2, both the zero-inflated model and the hurdle model) are closer to zero on average, and the distribution of those errors is approximately normal.

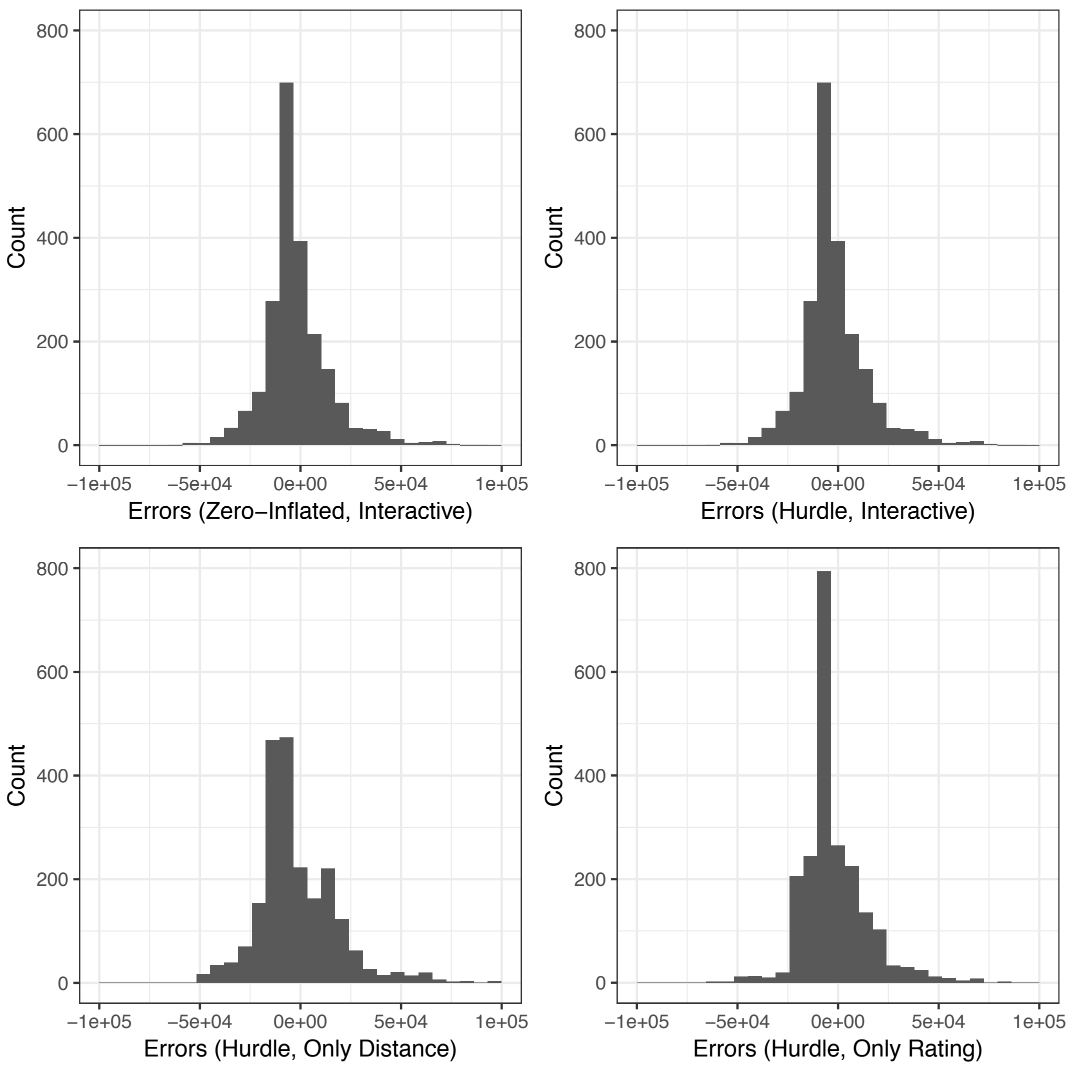

Figure A.2: Error Comparisons: Interactive Model (Zero-Inflated and Hurdle) Versus ABLR Original Hurdle Models.