ONLINE SUPPLEMENT TO

"When Random Assignment Is Not Enough:

Accounting for Item Selectivity in Experimental Research"

Online Supplement A: Auxiliary Tests of Model Specification

All three empirical applications relied on a particular model specification: Prior Ratings serving as an observed measure of individual-level preference (due to difficulties with full parametric heterogeneity based on relatively few choices per participant); a restricted error covariance structure; and inferences based on a lack of overt omitted regressors. We address each of these in turn via formal tests.

A1. Preference Heterogeneity

We stress the importance of accounting for "observed" (preference) heterogeneity when "unobserved" heterogeneity cannot be adequately captured due to the nature of collected data (e.g., few choices per consumer and/or many potentially heterogeneous regressors; see Andrews et al. 2008). To assess the role of the prior ratings, which serve as individual item preferences in each study, we re-estimate each of the "best" models omitting Prior Rating in the Selection submodel, the Prediction submodel, or both. For the RKK data (our Study 1), the "best" model yields LL of -2106.8 (Table 1; multiple ρ); removing Prior Rating from Selection alters this to -2174.3; out of Prediction, to -2369.9; out of both to -2683.4; *p*-values against the "best" model are infinitesimal in each case (Δ(df) is 1, 1, and 2, respectively, for the LR χ^2 test, and 2Δ(LL) is over 100 in all three), although comparatively less so for the Selection submodel. LL for Study 2 was -808.7 for the best model (Table 2; multiple ρ); removing Prior Rating from Selection yields -878.6; from Prediction, -839.0; from both, -909.5; *p*-values are again infinitesimal, although here the smaller difference was for Prediction. Lastly, for Study 3, LL was -916.0 for the best model (Table 3; multiple ρ); removing Prior Rating from Selection yields -943.3; from Prediction, -957.7; and from both, -977.3. Yet again, *p*-values are infinitesimal, but here the effects for Selection and Prediction are roughly equal. In all three studies, however, failing to account for observed preference heterogeneity in *either* submodel produces severe decrements in model fit. We would again stress that, for researchers estimating such many-parameter models on limited within-subject choice data, suitable measures of individual-level preference heterogeneity be collected and incorporated in both the Selection and Prediction portions of the model.

A2. Full Error Covariance

We find no evidence of significant error covariance in the multinomial choice model for any of our data sets; testing the full covariance probit model against the one used here yields the following results: RKK, $p = 0.996$; Study 2, $p = 0.327$; Study 3, $p = 0.753$. We cannot, of course, recommend that empirical researchers fail to account for error covariance in all applications, but the consistent lack of significance here contrasts sharply with the strong significance in all three studies for ρ, the central metric for the presence of item selectivity.

A3. Item Selectivity versus Omitted Variables

Researchers interested in substantive claims about item selectivity, must be careful about potential omitted regressors. General tests for omitted regressors date back to Ramsey (1969), who introduced the RESET procedure for linear models. Here, we follow Peters (2000), who demonstrates that a RESET-based procedure applies to a wide range of parametric and semiparametric model types, as well as how to perform associated (Likelihood Ratio) tests. The procedure is straightforward: for any particular model, calculate its predictions for the observed variables (e.g., ratings, in our context), $\hat{\boldsymbol{y}}$; then, re-estimate the model successively (cumulatively) including new regressors, $\hat{\boldsymbol{y}}^{\boldsymbol{k}}$, for $\boldsymbol{k} = \mathbf{2}, \mathbf{3}, \ldots$; note that $\boldsymbol{k} = \mathbf{1}$ corresponds to the regressors already included in the (Prediction) model.

Table A1 lists the results of this procedure for our three data sets, for $k = 1, \ldots, 5$ ($k > 5$ revealed no new phenomena). For each data set, all five values of k are run for the "multiple ρ" model (i.e., the best-fitting models, as listed in Tables 1, 2, and 3) and for an analogous model with ρ restricted to zero, that is, an ordinary regression; classical LL values are listed for all thirty models (3 data sets × 5 values of k × 2 conditions for ρ). Three types of p-values can therefore be calculated: each model vs. $k = 1$; each model vs. k being one smaller ("$k - 1$"); and each "multiple ρ" model vs. ρ = 0 (see footnotes to Table A1 for fuller descriptions). These appear in successive columns, and tell an especially clear 'story'. First, there is mild evidence of model misspecification (i.e., $.01 < p < .05$) only for the RKK data, and then only for k = 2 (i.e., by including $\hat{y}^2$ as a regressor); this does not hold for any k vs. k - 1 for $k > 2$. There is no evidence whatsoever of misspecification for Studies 2 and 3, with all 32 possible tests having $p > .23$.

[Table A1 about here]

While one might wonder about the RKK data in this regard, the lack of strong evidence for misspecification is shorn up greatly (and more so in Studies 2 and 3) by the last column of tests, against ρ = 0. In all cases, these are each in the range of $p \approx .001$. That is, *tests for including ρ come out at least 10 times stronger (100 times for Studies 2 and 3) than those for including various combinations of* $\hat{y}^k$. As such, this casts strong doubt on the omission of $\{\hat{y}^k\}$ as a *source* for claims of item selectivity, decisively for Studies 2 and 3 and arguably for RKK. We would strongly caution researchers who wish to test for item selectivity to run similar RESET-like tests on their resulting models in order to distinguish the strength of evidence for selectivity from that of possible omitted regressors.

<u>Additional Reference</u>:

Ramsey, J. B. (1969), Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis, *Journal of the Royal Statistical Society: Series B (Statistical Methodological)*, 350-371.

ONLINE TABLE A1: RAMSEY RESET TESTS FOR OMITTED REGRESSORS

***k* = 2, ..., 5 versus "*k* = 1", "*k* - 1", and "ρ = 0"**

Study	Model	*k*	LL	*p*-value vs. "*k* = 1"[a]	*p*-value vs. "*k* - 1"[b]	*p*-value vs. "ρ = 0"[c]
Study 1 (RKK)	**Multiple ρ**	**1**	**-2106.8**	---	---	0.0014
		2	**-2103.7**	0.0128	0.0128	0.0003
		3	**-2103.6**	0.0398	0.6201	0.0006
		4	**-2103.1**	0.0584	0.3117	0.0033
		5	**-2101.9**	0.0422	0.1192	0.0010
Study 1 (RKK)	**ρ = 0**	**1**	**-2113.4**	---	---	---
		2	**-2111.7**	0.0639	0.0639	---
		3	**-2111.1**	0.0969	0.2663	---
		4	**-2108.8**	0.0272	0.0340	---
		5	**-2108.8**	0.0553	0.7780	---
Study 2	**Multiple ρ**	**1**	**-808.7**	---	---	0.0010
		2	**-808.0**	0.2353	0.2353	0.0007
		3	**-808.0**	0.4847	0.8418	0.0007
		4	**-807.9**	0.6793	0.7997	0.0007
		5	**-807.9**	0.7979	0.7009	0.0007
Study 2	**ρ = 0**	**1**	**-815.6**	---	---	---
		2	**-815.3**	0.4351	0.4351	---
		3	**-815.2**	0.6874	0.7079	---
		4	**-815.2**	0.8514	0.8367	---
		5	**-815.2**	0.9388	0.9427	---
Study 3	**Multiple ρ**	**1**	**-916.0**	---	---	0.0011
		2	**-915.8**	0.5251	0.5251	0.0016
		3	**-915.3**	0.5043	0.3259	0.0016
		4	**-915.3**	0.6825	0.7185	0.0015
		5	**-914.8**	0.6696	0.3530	0.0016
Study 3	**ρ = 0**	**1**	**-924.0**	---	---	---
		2	**-923.4**	0.2811	0.2811	---
		3	**-923.0**	0.3586	0.3458	---
		4	**-923.0**	0.5616	0.9716	---
		5	**-922.5**	0.5403	0.3047	---

Notes.

[a] Tests the model with $\{\hat{y}, \ldots, \hat{y}^k\}$ against one with just $\hat{y}$.

[b] Tests the model with $\{\hat{y}, \ldots, \hat{y}^k\}$ against one with $\{\hat{y}, \ldots, \hat{y}^{k-1}\}$.

[c] Tests the "best" models ("Multiple ρ"; see Tables 1 - 3) with $\{\hat{y}, \ldots, \hat{y}^k\}$ against the analogous ones with the restriction of ρ = 0.

Online Supplement B: Derivations of Simple Two-Step Test Procedure and Detailed Comparisons

Consider the residuals, $\{\varepsilon_s\}$, with unconditional IID Gumbel densities, $F(x) = \exp(-\exp(-x))$. We focus specifically on those for the chosen items. For an item to be chosen, it's utility must be the largest, that is $\Pr\left[v_{s,1} + \varepsilon_{s,1} > \{v_{s,i} + \varepsilon_{s,i}\}_{(i>1)}\right]$. We therefore wish to calculate the central tendency properties – mean, median, and or mode – of the following random variable, dropping the s subscript for clarity:

$$\left[\varepsilon_1 \mid v_1 + \varepsilon_1 > \{v_i + \varepsilon_i\}_{(i>1)}\right]$$

This stems directly from the cdf, which can be written as follows:

$$\begin{aligned} F_{\varepsilon_1}(x) &= \Pr\left[\varepsilon_1 < x \mid v_1 + \varepsilon_1 > \{v_i + \varepsilon_i\}_{(i>1)}\right] \\ &= \frac{\Pr\left[\{v_i - v_1 + \varepsilon_i\}_{(i>1)} < \varepsilon_1 < x\right]}{\Pr\left[\{v_i - v_1 + \varepsilon_i\}_{(i>1)} < \varepsilon_1\right]} \end{aligned}$$

The denominator is the usual logit probability and serves as a normalizing constant with respect to x, so we proceed by integrating the numerator over $w = \varepsilon_1$:

$$\begin{aligned} &\Pr\left[\{v_i - v_1 + \varepsilon_i\}_{(i>1)} < \varepsilon_1 < x\right] \\ &= \int_{w=-\infty}^{x} \Pr\left[\max\{v_i - v_1 + \varepsilon_i\}_{(i>1)} < w\right] dF(w)\, dw \\ &= \int_{w=-\infty}^{x} \left(\prod_{i>1} \Pr(\varepsilon_i < w - [v_i - v_1])\right) \left[\exp(-\exp(-w))\right] \exp(-w)\, dw \\ &= \int_{w=-\infty}^{x} \left(\prod_{i>1} \exp\left(-\exp(-(w - [v_i - v_1]))\right)\right) \left[\exp(-\exp(-w))\right] \exp(-w)\, dw \end{aligned}$$

This can be simplified by noticing that the product of two Gumbel CDFs is again Gumbel; in this case, two Gumbels with different location parameters a and b, but the same scale parameter, 1:

$$\begin{aligned} &\exp\left(-\exp(-(w-a))\right) \exp\left(-\exp(-(w-b))\right) \\ &= \exp(-\exp(-w)\,[\exp(a) + \exp(b)]) \\ &= \exp\left(-\exp\left(-\left[w - \ln(\exp(a) + \exp(b))\right]\right)\right) \end{aligned}$$

The resulting Gumbel has location parameter $\ln(\exp(a) + \exp(b))$; this generalizes to any number of Gumbels. The proof is completed by noting that the last integrand above is a product of CDFs with parameters $\{v_i - v_1\}$ and 1, so the kernel of the integral corresponds to a Gumbel CDF with location parameter:

$$\begin{aligned} &\ln\left(1 + \sum_{i>1} \exp(v_i - v_1)\right) \\ &= \ln\left(\frac{\exp(v_1) + \sum_{i>1} \exp(v_i)}{\exp(v_1)}\right) \\ &= -\ln(p_1) \end{aligned}$$

where $p_1 = \exp(v_1)/\sum_i \exp(v_i)$ is the usual logit probability.

The mean, median, and mode, respectively, of this distribution are given by the following, with $\mu = -\ln(p_1)$, $\beta = 1$, and γ the Euler-Mascheroni constant, approximately 0.5772:

$$\text{Mean:} \quad \mu + \beta\gamma = -\ln(p_1) + \gamma$$
$$\text{Median:} \quad \mu - \beta\ln(\ln(2)) = -\ln(p_1) - \ln(\ln(2))$$
$$\text{Mode:} \quad \mu = -\ln(p_1)$$

To use the results above for a sample selection model requires converting these measures of central tendency into their corresponding points on the *standard normal* CDF. This is readily accomplished by simply computing the standard Gumbel CDF and then the inverse standard normal:

$$\text{Mean:} \quad \Phi^{-1}\left[\exp\left(-p_1\exp(-\gamma)\right)\right]$$
$$\text{Median:} \quad \Phi^{-1}\left[\exp\left(-p_1\ln(2)\right)\right]$$
$$\text{Mode:} \quad \Phi^{-1}[\exp(-p_1)]$$

These can all be seen as variants on $\Phi^{-1}\left(\exp(-ap_1)\right)$, with $a \approx 0.5615$, 0.6932, and 1, respectively. Below, we empirically test the performance of all three (see Tables B1-B3).

Two-Step Estimation of ρ

To estimate ρ, we augment the Prediction model using some estimate of the residuals, as above, for the chosen items, which we refer to simply as X :

$$Y_p \;=\; X_p\beta_p + \varepsilon_p \;=\; X_p\beta_p + X\beta + \varepsilon_{new}$$

Now, ρ can be estimated by definition as the correlation between the error in the selection model – represented by the new variable (X) – and ε_p, which has been decomposed into $X\beta + \varepsilon_{new}$. That is:

$$\rho \approx corr(X, X\beta + \varepsilon_{new})$$

Because X and ε_{new} are are in the same regression, they are uncorrelated. We can therefore compute the correlation above as:

$$\rho = corr(X, X\beta + \varepsilon_{new}) = \frac{Cov(X, X\beta + \varepsilon_{new})}{\sqrt{Var(X)\,Var(X\beta + \varepsilon_{new})}}$$
$$= \frac{\beta\sigma_X^2}{\sqrt{\sigma_X^2[\beta^2\sigma_X^2 + Var(\varepsilon_{new})]}}$$
$$\approx \frac{b}{\sqrt{b^2 + MSE/\sigma_X^2}}$$

where $b = \hat{\beta}$ is the estimated coefficient for X , σ_X^2 is the variance of X, and MSE is the estimated mean-square-error of the regression.

ONLINE TABLES B1-3: TESTING APPROXIMATE ESTIMATION PROCEDURES

Correct Full Bayesian Results vs. {Logit, Probit} Selection × {OLS, Bayesian} Regression × {Mean, Median, Mode} Variable

ONLINE TABLE B1: MEAN Variable, $X = \Phi^{-1}[\exp(-p_1 \exp(-\gamma))]$, $\gamma = 0.5772$

		Probit Selection				Logit Selection: MEAN					Probit Selection: MEAN				
		Correct Full Bayesian				OLS		Bayesian			OLS		Bayesian		
Study	**Condition**	**Rho**	**p-value**	**CI Lower bound**	**CI upper bound**	**Est'd rho**	**p-value**	**Est'd (mean) rho**	**CI Lower bound**	**CI upper bound**	**Est'd rho**	**p-value**	**Est'd (mean) rho**	**CI Lower bound**	**CI upper bound**
1 (RKK)	ALL	0.043	0.714	-0.289	0.332	0.012	0.82	0.029	-0.045	0.126	0.022	0.665	0.014	-0.077	0.110
	Small set size	-0.125	0.001	-0.404	0.138	-0.068	0.211	-0.052	-0.146	0.050	-0.068	0.212	-0.063	-0.172	0.037
	Large set size	0.571		0.353	0.727	0.110	0.086	0.135	0.039	0.236	0.117	0.069	0.140	0.041	0.244
2	ALL	0.413	0.021	0.055	0.703	0.294	0.009	0.283	0.082	0.476	0.325	0.004	0.321	0.126	0.513
	Small set size	-0.273	0.001	-0.731	0.433	0.067	0.602	0.070	-0.165	0.305	0.077	0.556	0.065	-0.151	0.290
	Large set size	0.568		0.280	0.787	0.473	0	0.468	0.294	0.628	0.467	0	0.457	0.266	0.613
3	ALL	0.543	0.002	0.190	0.788	0.633	0	0.639	0.570	0.712	0.630	0	0.654	0.590	0.709
	Bunchy Unattr	0.470		0.040	0.823	0.355	0.158	0.464	0.290	0.637	0.322	0.213	0.354	0.221	0.495
	Not Bunchy	0.835	0.001	0.578	0.959	0.755	0	0.732	0.617	0.821	0.757	0	0.756	0.675	0.833
	Bunchy Attr	0.450		0.028	0.740	0.583	0	0.586	0.449	0.691	0.591	0	0.598	0.508	0.680

ONLINE TABLE B2: MEDIAN Variable, $\mathbf{X} = \boldsymbol{\Phi}^{-1}\left[\exp\left(-\boldsymbol{p}_1 \ln(2)\right)\right]$

		Probit Selection				Logit Selection: MEDIAN					Probit Selection: MEDIAN				
		Correct Full Bayesian				OLS		Bayesian			OLS		Bayesian		
Study	**Condition**	**Rho**	**p-value**	**CI Lower bound**	**CI upper bound**	**Est'd rho**	**p-value**	**Est'd (mean) rho**	**CI Lower bound**	**CI upper bound**	**Est'd rho**	**p-value**	**Est'd (mean) rho**	**CI Lower bound**	**CI upper bound**
1	ALL	0.043	0.714	-0.289	0.332	0.011	0.829	0.015	-0.078	0.121	0.022	0.674	0.023	-0.078	0.119
(RKK)	Small set size	-0.125	0.001	-0.404	0.138	-0.069	0.209	-0.067	-0.178	0.052	-0.069	0.209	-0.073	-0.178	0.035
	Large set size	0.571		0.353	0.727	0.110	0.086	0.119	0.011	0.238	0.117	0.069	0.146	0.028	0.282
2	ALL	0.413	0.021	0.055	0.703	0.294	0.009	0.289	0.089	0.467	0.326	0.004	0.328	0.138	0.535
	Small set size	-0.273	0.001	-0.731	0.433	0.068	0.598	0.075	-0.168	0.300	0.078	0.55	0.076	-0.164	0.296
	Large set size	0.568		0.280	0.787	0.474	0	0.470	0.306	0.632	0.468	0	0.456	0.286	0.628
3	ALL	0.543	0.002	0.190	0.788	0.636	0	0.674	0.603	0.741	0.634	0	0.613	0.474	0.715
	Bunchy Unattr	0.470		0.040	0.823	0.360	0.153	0.469	0.161	0.685	0.328	0.206	0.386	0.203	0.563
	Not Bunchy	0.835	0.001	0.578	0.959	0.756	0	0.711	0.643	0.777	0.758	0	0.762	0.698	0.822
	Bunchy Attr	0.450		0.028	0.740	0.584	0	0.665	0.604	0.726	0.593	0	0.578	0.506	0.650

ONLINE TABLE B3: MODE Variable, $X = \Phi^{-1}[\exp(-p_1)]$

		Probit Selection				Logit Selection: MODE					Probit Selection: MODE				
		Correct Full Bayesian				OLS		Bayesian			OLS		Bayesian		
Study	Condition	Rho	p-value	CI Lower bound	CI upper bound	Est'd rho	p-value	Est'd (mean) rho	CI Lower bound	CI upper bound	Est'd rho	p-value	Est'd (mean) rho	CI Lower bound	CI upper bound
1	ALL	0.043	0.714	-0.289	0.332	0.010	0.846	0.007	-0.092	0.106	0.020	0.691	0.020	-0.078	0.116
(RKK)	Small set size	-0.125	0.001	-0.404	0.138	-0.069	0.204	-0.043	-0.145	0.054	-0.070	0.205	-0.074	-0.179	0.021
	Large set size	0.571		0.353	0.727	0.110	0.087	0.147	0.053	0.256	0.117	0.069	0.119	0.019	0.226
2	ALL	0.413	0.021	0.055	0.703	0.294	0.009	0.282	0.090	0.482	0.327	0.004	0.328	0.117	0.513
	Small set size	-0.273	0.001	-0.731	0.433	0.069	0.592	0.086	-0.159	0.317	0.080	0.539	0.086	-0.155	0.317
	Large set size	0.568		0.280	0.787	0.475	0	0.473	0.302	0.628	0.470	0	0.466	0.297	0.631
3	ALL	0.543	0.002	0.190	0.788	0.641	0	0.714	0.632	0.805	0.640	0	0.610	0.492	0.706
	Bunchy Unattr	0.470	0.001	0.040	0.823	0.368	0.145	0.430	0.199	0.654	0.339	0.192	0.313	0.131	0.468
	Not Bunchy	0.835		0.578	0.959	0.756	0	0.811	0.762	0.857	0.760	0	0.671	0.582	0.753
	Bunchy Attr	0.450		0.028	0.740	0.586	0	0.540	0.415	0.645	0.595	0	0.618	0.539	0.687

Online Supplement C: Robustness to Alternative Specifications of the Selection Model

ONLINE TABLE C1: STUDY 1 (RKK) AUXILIARY MODEL COMPARISONS

Posterior Means for Selection, Prediction, ρ, and σ

	Parameter Estimates (Std. Err.)			
Model	**Focal Model: Multiple ρ**	**M1a**	**M1b**	**M1c**
Selection Model				
Prior Rating	**0.432 (.038)**	**0.374 (.037)**	**0.474 (.033)**	
Choice Lag	**-0.481 (.077)**	**-0.352 (.072)**	-0.099 (.060)	**-0.460 (.074)**
Frequency	**0.168 (.061)**	**0.361 (.057)**		**0.532 (.056)**
Choice Lag × Prior Rating	**-0.226 (.067)**	-0.113 (.063)	-0.095 (.062)	
Choice Lag × Frequency	**0.791 (.096)**			**0.472 (.089)**
Prediction Model				
Intercept	**-0.129 (.054)**	-0.084 (.057)	-0.125 (.074)	-0.106 (.053)
Prior Rating	**0.747 (.023)**	**0.739 (.023)**	**0.748 (.025)**	**0.716 (.021)**
Choice Lag	**0.155 (.058)**	**0.159 (.060)**	**0.166 (.058)**	**0.151 (.058)**
Frequency	**0.167 (.029)**	**0.162 (.030)**	**0.150 (.027)**	**0.170 (.030)**
Set Size	**-0.322 (.112)**	-0.261 (.137)	-0.246 (.133)	**-0.657 (.169)**
Frequency × Prior Rating	**-0.147 (.022)**	**-0.143 (.022)**	**-0.146 (.022)**	**-0.133 (.022)**
Frequency × Set Size	**0.099 (.047)**	0.093 (.049)	0.052 (.044)	**0.188 (.057)**
Sigma, σ	**0.651 (.100)**	**0.643 (.019)**	**0.645 (.019)**	**0.684 (.027)**
ρ_{Small}, Small Set Size[a]	-0.125 [-.404, .138]	-0.149 [-.419, .134]	-0.058 [-.355, .253]	**-0.448 [-.739, -.051]**
ρ_{Large}, Large Set Size[a]	**0.571 [.353, .727]**	**0.456 [.167, .661]**	**0.489 [.102, .686]**	**0.679 [.477, .811]**
Number of parameters	15	14	13	13
Log Likelihood	-2106.836	-2141.98	-2163.24	-2186.92
LR Test: p-value vs. focal model		<.001	<.001	<.001

Notes. Bold denotes statistical significance.
[a] Numbers in brackets represent the 95% Bayesian Highest Density Region (HDR).

ONLINE TABLE C2: STUDY 2 AUXILIARY MODEL COMPARISONS

Posterior Means for Selection, Prediction, ρ, and σ

	Parameter Estimates (Std. Err.)		
Model	**Focal Model: Multiple ρ**	**M2a**	**M2b**
Selection Model			
Prior Rating	**0.527 (.054)**	**0.528 (.054)**	
Favorite	**0.315 (.109)**	**0.281 (.108)**	**1.029 (.133)**
Choice Lag	**-0.289 (.138)**	**-0.335 (.141)**	-0.184 (.208)
Choice Lag × SEQ	**0.970 (.267)**	**1.021 (.273)**	**0.869 (.394)**
Choice Lag × Favorite	**-0.792 (.299)**		-0.476 (.430)
Prediction Model			
Intercept	-0.185 (.187)	-0.145 (.163)	-0.001 (.255)
Prior Rating	**0.493 (.064)**	**0.486 (.063)**	**0.474 (.077)**
Favorite	**0.273 (.134)**	0.250 (.130)	0.116 (.225)
Choice Lag	0.286 (.153)	0.289 (.154)	0.397 (.271)
Set Size	**-0.881 (.322)**	**-0.985 (.353)**	-0.799 (.497)
Choice Lag × Favorite	**-0.711 (.310)**	**-0.674 (.300)**	**-0.768 (.311)**
Sigma, σ	**0.901 (.048)**	**.905 (.050)**	**0.891 (.058)**
ρ_{Small}, Small Set Size[a]	-0.273 [-.731, .433]	-0.364 [-.799, .276]	-0.320 [-.748, .175]
ρ_{Large}, Large Set Size[a]	**0.568 [.280, .787]**	**0.573 [.249, .786]**	0.346 [-.152, .682]
Number of parameters	14	13	13
Log Likelihood	-808.703	-812.214	-878.644
LR Test: p-value vs. focal model		.008	<.001

Notes. Bold denotes statistical significance.
[a] Numbers in brackets represent the 95% Bayesian Highest Density Region (HDR).

ONLINE TABLE C3: STUDY 3 AUXILIARY MODEL COMPARISONS

Posterior Means for Selection, Prediction, ρ, and σ

	Parameter Estimates (Std. Err.)			
Model	**Focal Model: Multiple ρ**	**M3a**	**M3b**	**M3c**
Selection Model				
Prior Rating	**0.454 (.065)**	**0.492 (.062)**	**0.656 (.050)**	
Favorite	**0.517 (.118)**	**0.460 (.114)**		**1.212 (.087)**
Choice Lag	-0.003 (.117)	0.030 (.115)	0.111 (.110)	0.160 (.116)
Prior Rating × SEQ	**-0.196 (.092)**	**-0.185 (.093)**	**-0.199 (.097)**	
Choice Lag × SEQ	**0.883 (.233)**	**0.855 (.229)**	**0.827 (.225)**	**0.778 (.218)**
Choice Lag × Prior Rating	**-0.266 (.116)**	**-0.262 (.115)**	-0.191 (.113)	
Bunchy Attractive × Favorite	**-0.497 (.206)**			**-0.741 (.201)**
Bunchy Unattractive × Favorite	-0.075 (.207)			0.342 (.193)
Prediction Model				
Intercept	**-0.607 (.135)**	**-0.618 (.128)**	**-0.679 (.126)**	**-0.321 (.120)**
Prior Rating	**0.546 (.060)**	**0.552 (.059)**	**0.584 (.061)**	**0.473 (.057)**
Favorite	**0.640 (.153)**	**0.627 (.152)**	**0.513 (.136)**	**0.547 (.161)**
Choice Lag	-0.062 (.108)	-0.055 (.106)	-0.018 (.108)	-0.063 (.104)
SEQ	-0.075 (.084)	-0.075 (.085)	-0.075 (.083)	-0.074 (.085)
Bunchy Attractive	0.073 (.204)	0.108 (.193)	0.089 (.185)	0.336 (.253)
Bunchy Unattractive	-0.093 (.202)	-0.066 (.208)	-0.017 (.197)	-0.126 (.259)
Prior Rating × SEQ	**-0.240 (.089)**	**-0.244 (.089)**	**-0.249 (.090)**	**-0.193 (.086)**
Choice Lag × SEQ	**0.536 (.222)**	**0.546 (.221)**	**0.563 (.221)**	0.386 (.213)
Bunchy Attractive × Favorite	**-1.093 (.259)**	**-0.923 (.249)**	**-0.846 (.240)**	**-1.186 (.297)**
Bunchy Unattractive × Favorite	0.504 (.362)	0.572 (.355)	**0.689 (.336)**	0.346 (.413)
Bunchy Attractive × Prior Rating	**0.288 (.136)**	**0.280 (.137)**	0.265 (.135)	**0.364 (.138)**
Bunchy Unattractive × Prior Rating	-0.256 (.138)	-0.262 (.140)	**-0.279 (.138)**	-0.159 (.138)
Sigma, σ	**0.891 (.047)**	**0.896 (.047)**	**0.907 (.048)**	**0.838 (.036)**
$\rho_{BunchyAttr}$, Bunchy Attractive Set[a]	**0.450 [.028, .740]**	**0.465 [.052, .747]**	**0.528[.156, .786]**	-0.039 [-.355, .354]
$\rho_{BunchyUnattr}$, Bunchy Unattractive Set[a]	**0.470 [.040, .823]**	**0.474 [.043, .824]**	**0.500 [.052, .835]**	0.236 [-.179, .635]
$\rho_{NotBunchy}$, Not Bunchy Set[a]	**0.835 [.578, .959]**	**0.854 [.633, .962]**	**0.893 [.727, .962]**	**0.565 [.238, .806]**
Number of parameters	25	23	22	22
Log Likelihood	-916.022	-919.179	-927.410	-965.452
LR Test: p-value vs. focal model		.046	<.001	<.001

Notes. Bold denotes statistical significance.
[a] Numbers in brackets represent the 95% Bayesian Highest Density Region (HDR).