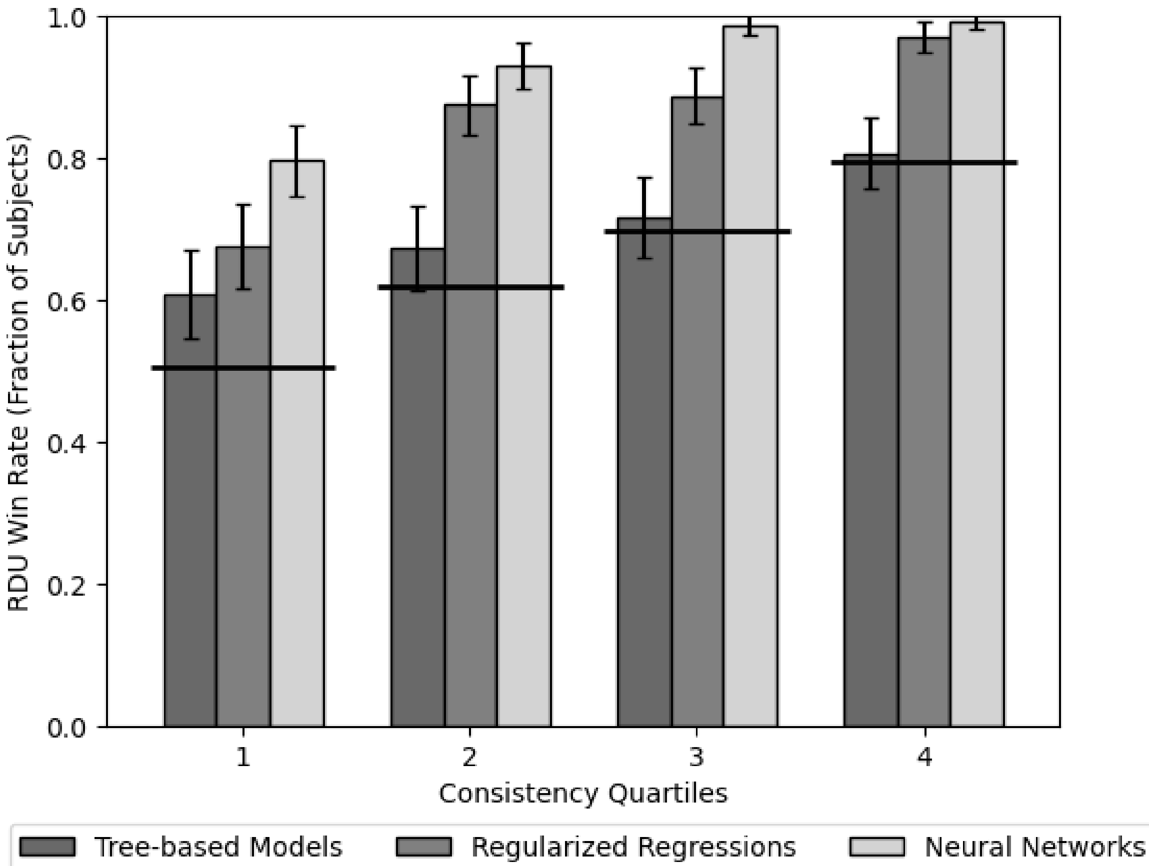
Online Appendix for "Predicting and Understanding Individual-Level Choice Under Risk"

Keaton Ellis, Shachar Kariv and Erkut Ozbay

September 30, 2024

Figure 1: RDU win rate over ML by quartiles of consistency scores with GARP and FOSD



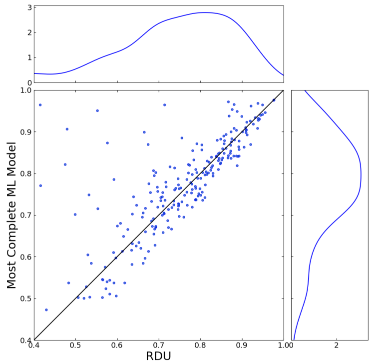
The fraction of subjects for whom RDU is more complete than the best regularized regression, tree-based, and neural network models, as well as the overall best ML model (indicated by black horizontal lines). The x-axis groups subjects by quartiles of consistency scores with GARP and FOSD, following the methods of Nishimura et al. (2017) and Polisson et al. (2020). This score measures the amount by which each budget constraint must be relaxed in order to remove all violations of GARP and FOSD and it is bounded between 0 and 1. A score closer to 1 indicates stronger consistency with GARP and FOSD. The quartile ranges are $[0, 0.83)$, $[0.83, 0.95)$, $[0.95, 0.99)$, and $[0.99, 1]$.

Table 1: The completeness and restrictiveness of RDU versus ML models

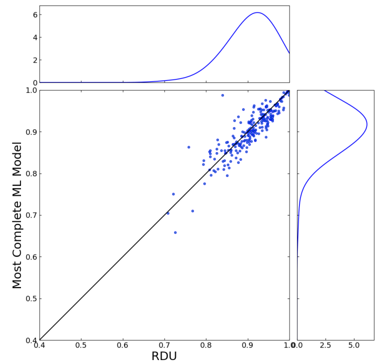
	Average	RDU's win rate	RDU's win rate against ML by e^{**} quartiles				Absolute completeness difference between RDU and ML by e^{**} quartiles				
Panel A: Model Classes	Completeness	against ML	1st	2nd	3rd	4th	1st	2nd	3rd	4th	Restrictiveness
RDU	89.2% [88.3%, 89.9%]	-	-	-	-	-	-	-	-	-	16.6%
Regularized Regressions	79.5% [77.8%, 80.5%]	85.1%	67.5%	87.4%	88.8%	97.0%	3.1%	7.5%	9.9%	18.0%	20.7%
Tree-based Models	89.1% [88.3%, 89.9%]	70.1%	60.8%	67.4%	71.7%	80.6%	-2.0%	0.6%	0.7%	0.8%	10.6%
Neural Networks	71.6% [68.8%, 73.7%]	92.6%	79.6%	92.9%	98.8%	99.2%	8.7%	14.4%	16.8%	30.7%	14.4%
Panel B: Regularized regressions											
Lasso	75.9% [74.2%, 76.9%]	89.6%	77.9%	90.8%	91.7%	98.3%	6.4%	11.5%	14.0%	21.3%	20.7%
OLS	70.2% [57.7%, 74.6%]	87.1%	70.8%	90.0%	90.0%	97.9%	10.6%	10.4%	15.9%	39.0%	20.7%
Ridge	70.6% [58.2%, 75.1%]	87.0%	70.8%	89.5%	90.0%	97.9%	10.5%	10.2%	15.6%	38.1%	20.7%
Panel C: Tree-based models											
Mean	86.6% [85.6%, 87.4%]	85.9%	77.5%	88.3%	86.3%	91.6%	2.4%	3.5%	2.4%	2.0%	12.4%
Linear	82.9% [81.7%, 84.0%]	86.5%	81.3%	85.8%	87.5%	91.6%	11.8%	5.8%	3.7%	3.6%	5.4%
SVR	85.7% [84.8%, 86.6%]	88.5%	80.4%	90.4%	87.9%	95.4%	3.5%	3.9%	2.9%	3.5%	10.7%
RF	88.0% [87.2%, 88.8%]	79.9%	70.0%	78.7%	80.8%	90.3%	-0.1%	1.5%	1.4%	1.7%	11.9%

The left column reports the average completeness of each model, as well the 95% confidence interval for average completeness, and the next column reports the win rate of RDU against each model (that is, the fraction of subjects for whom RDU is more complete). The next two blocks of four columns report the win rate of RDU against each model and its absolute completeness difference by quartiles of the consistency score with GARP and FOSD. The right column reports the restrictiveness of each model. Panel A reports the results for RDU and the three *families* of ML models—regularized regressions, tree-based, and neural networks. For regularized regressions and tree-based models, we report restrictiveness as weighted averages of the most complete model in the class for each subject. Panels B and C report the results for each regularized regression and tree-based model, respectively.

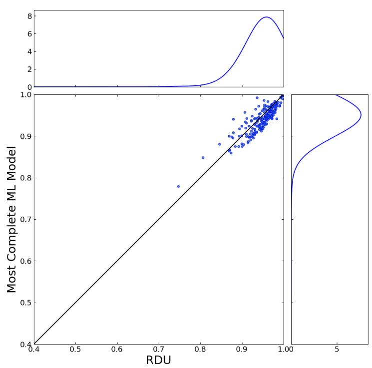
Figure 2: The individual-level completeness of RDU versus the most complete ML model by e^{**} quartile.



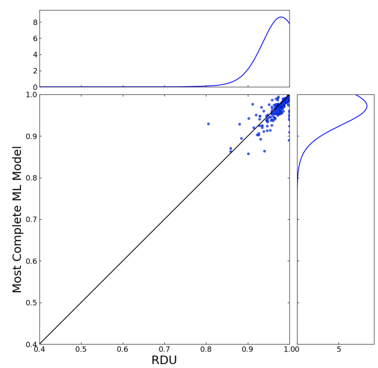
(a) e^{**} quartile 1



(b) e^{**} quartile 2



(c) e^{**} quartile 3



(d) e^{**} quartile 4

The four panels plot the completeness scores of all subjects for RDU and the best ML model. Panels refer to the quartile of consistency score; Panel (a) plots the subjects in the lowest quartile of e^{**}, Panel (b) the second quartile of e^{**}, and so on. The quartile ranges are $[0, 0.83)$, $[0.83, 0.95)$, $[0.95, 0.99)$, and $[0.99, 1]$. Each plotted point represents a subject. The horizontal axes are the completeness of RDU, and the vertical axes show the completeness of the best ML model. Each axis also provides a marginal kernel density estimate of completeness scores approximated using a Gaussian kernel.

References

Nishimura, H., E. A. Ok, and J. K.-H. Quah (2017): "A Comprehensive Approach to Revealed Preference Theory," *American Economic Review*, 107, 1239–63.

Polisson, M., J. K.-H. Quah, and L. Renou (2020): "Revealed Preferences over Risk and Uncertainty," *American Economic Review*, 110, 1782–1820.