A Appendix

A.1 Model Robustness

We run the best embedding-based model ConvE and Ours(ConvE) on all datasets using 5 different random seeds with all other hyperparameters fixed. Table 6 reports the mean and standard deviation of each model. We observe that both models demonstrate a small standard deviation (< 0.01) on all datasets.

Dataset	ConvE	Ours(ConveE)
UMLS	95.5±0.4	93.7±0.2
Kinship	86.9±0.3	86.2±0.7
FB15k-237	43.5±0.1	40.7±0.2
WN18RR	45.3±0.4	44.7±0.2
NELL-995	76.2±0.3	72.7±0.4

Table 6: Test set MRR×100 mean and standard deviation across five runs on all datasets.

A.2 Development Set Evaluation Using Complete KGs

Comparing Table 2 and Table 3 reveals that the dev set MRRs are significantly lower than the test set MRRs on some datasets (UMLS, Kinship and FB15k-237). Such discrepancies are caused by the multi-answer queries in these datasets. As most benchmark datasets randomly split the KG triples into train/dev/test sets, the queries that have multiple answers may fall into multiple splits. Because we hide all triples in the test set during the dev set evaluation, some predictions generated during dev set evaluation were wrongly punished as false negatives. In contrast, the test set evaluation metrics are computed using the complete KGs. Access to the complete KG eliminates most of the false negatives cases and hence increases the performance.

Model	UMLS	Kinship	FB15k237	WN18RR	NELL995
Ours(ConvE)	**95.1**	**86.8**	**41.8**	44.1	**78.8**
−RS	85.6	75.7	37.1	**46.1**	78.4
−AD	76.2	75.9	32.4	39.3	76.1

Table 7: Comparison of dev set MRR computed using the complete KGs of Ours(ConvE) and models without reward shaping and action dropout.

Table 7 shows the dev set MRR of the same systems shown in Table 3 with the MRRs computed using the complete KGs. On four of the datasets, the evaluation metrics increases significantly to the level that is comparable to those on the test set, with the relative improvement correlated with the average node fan-out in the KG (Table 1).

Notice that Table 7 is generated after all hyperparameters were fixed and the purpose is to show the effects of such dataset peculiarities. To avoid potential test set leakage, hyperparameter search should be done with the test set triples hidden (Table 1) instead of with the full KG.

A.3 Action Dropout Rates Used for Different KGs

Table 8 show the action dropout rates used for all KG datasets in our experiments. In general, larger action dropout rates are necessary for KGs that are densely connected. We find a positive correlation between the optimal action dropout rate and the average node fan-out (Table 1).

For UMLS and Kinship, we tried setting the action dropout rate to 1.0 (completely random sampling) and observed small but significant performance drop. Random sampling performs reasonably well on these two datasets possibly due to the fact that they are small. For larger KGs (FB15k-237, WN18RR, NELL-995), policy-guided sampling is necessary.

Dataset	α
UMLS	0.95
Kinship	0.9
FB15k-237	0.5
WN18RR	0.1
NELL-995	0.1

Table 8: Action dropout rates used in our experiments.