

A PROOFS

Theorem 1. If $P_{\theta, \gamma, \lambda}$ is feature-MAR and label-MAR, then (2) simplifies to

$$\int_{\mathbf{X}^{miss}} P_{\gamma}(\mathbf{Y}|\mathbf{X}) P_{\theta}(\mathbf{X}^{miss}|\mathbf{X}^{obs}). \quad (6)$$

Proof.

$$\begin{aligned} P_{\theta, \gamma, \lambda}(\mathbf{Y}|\mathbf{X}, \mathbf{M}) &= P_{\lambda}(\mathbf{M}|\mathbf{X}, \mathbf{Y}) \frac{P_{\gamma}(\mathbf{Y}|\mathbf{X})}{P_{\gamma, \lambda}(\mathbf{M}|\mathbf{X})} \stackrel{(4)}{=} P_{\gamma}(\mathbf{Y}|\mathbf{X}) \\ P_{\theta, \gamma, \lambda}(\mathbf{X}^{miss}|\mathbf{X}^{obs}, \mathbf{M}) &= P_{\gamma, \lambda}(\mathbf{M}|\mathbf{X}^{obs}, \mathbf{X}^{miss}) \frac{P_{\theta}(\mathbf{X}^{miss}|\mathbf{X}^{obs})}{P_{\theta, \gamma, \lambda}(\mathbf{M}|\mathbf{X}^{obs})} \stackrel{(3)}{=} P_{\theta}(\mathbf{X}^{miss}|\mathbf{X}^{obs}) \end{aligned}$$

□

Theorem 2. Let $\mathbf{X} \in \mathbb{R}^{n \times d}$ and $\mathbf{Y} \in \mathcal{Y}^n$ be random variables, $\mathbf{M} \in \{0, 1\}^{n \times d}$ be a missingness mask and \mathbf{X}^{obs} denotes the observed (incomplete) data. We encode the pair $(\mathbf{X}^{obs}, \mathbf{M})$ with the random variable $\tilde{\mathbf{X}}$ with

$$\tilde{X}_{ij} = \begin{cases} X_{ij}, & M_{ij} = 0, \\ ?, & M_{ij} = 1. \end{cases}$$

Let the change in the information be defined as $\Delta := I(\mathbf{Y}; \tilde{\mathbf{X}}) - I(\mathbf{Y}; \mathbf{X})$, where $I(\cdot; \cdot)$ denotes the mutual information. Then,

1. If the missingness is label-MAR, then $\Delta \leq 0$.
2. If $\mathbf{X} \in \{0, 1\}^{n \times d}$ and the missingness is U-MCAR with missingness probability μ , and $s(\mathbf{X})$ is the sample sparsity as in Definition 2, then

$$-nd\mu h_2(\mathbb{E}[s(\mathbf{X})]) \leq \Delta \leq 0,$$

where $h_2(u) = -u \log u - (1 - u) \log(1 - u)$.

Proof. By construction $\tilde{\mathbf{X}} = g(\mathbf{X}, \mathbf{M})$ for some measurable g . Thus $(\mathbf{Y}) \rightarrow (\mathbf{X}, \mathbf{M}) \rightarrow \tilde{\mathbf{X}}$ is a Markov chain, and the data-processing inequality implies

$$I(\mathbf{Y}; \tilde{\mathbf{X}}) \leq I(\mathbf{Y}; \mathbf{X}, \mathbf{M}). \quad (7)$$

Moreover, for any three random elements (A, B, C) we have the chain-rule identities

$$I(A; B, C) = I(A; C) + I(A; B | C). \quad (8)$$

(1) Label-MAR $\Delta \leq 0$. Assume label-MAR: $\mathbb{P}(\mathbf{M} | \mathbf{X}, \mathbf{Y}) = \mathbb{P}(\mathbf{M} | \mathbf{X})$, which is equivalent to $\mathbf{Y} \perp \mathbf{M} | \mathbf{X}$. Applying equation 8 with $(A, B, C) = (\mathbf{Y}, \mathbf{X}, \mathbf{M})$,

$$I(\mathbf{Y}; \mathbf{X}, \mathbf{M}) = I(\mathbf{Y}; \mathbf{X}) + I(\mathbf{Y}; \mathbf{M} | \mathbf{X}).$$

Under label-MAR, $I(\mathbf{Y}; \mathbf{M} | \mathbf{X}) = 0$, hence

$$I(\mathbf{Y}; \mathbf{X}, \mathbf{M}) = I(\mathbf{Y}; \mathbf{X}). \quad (9)$$

Combining equation 7 and equation 9 yields

$$I(\mathbf{Y}; \tilde{\mathbf{X}}) \leq I(\mathbf{Y}; \mathbf{X}) \iff \Delta = I(\mathbf{Y}; \tilde{\mathbf{X}}) - I(\mathbf{Y}; \mathbf{X}) \leq 0.$$

(2) Two-sided bound under uniform MCAR and α - β sparsity. Assume uniform MCAR: $M_{ij} \sim \text{Bernoulli}(1 - \mu)$ independently of (\mathbf{X}, \mathbf{Y}) and i.i.d. across (i, j) , and that $\mathbb{P}(s(\mathbf{X}) \geq \alpha) \geq \beta$, where $s(\mathbf{X}) = \frac{1}{nd} \sum_{i,j} \mathbb{I}\{X_{ij} = 0\}$.

Upper side. MCAR implies label-MAR, so by part (1): $\Delta \leq 0$.

Lower side. We start from the chain-rule identity applied to $(A, B, C) = (\mathbf{Y}, \mathbf{X}, \tilde{\mathbf{X}})$:

$$I(\mathbf{Y}; \mathbf{X}, \tilde{\mathbf{X}}) = I(\mathbf{Y}; \tilde{\mathbf{X}}) + I(\mathbf{Y}; \mathbf{X} | \tilde{\mathbf{X}}) = I(\mathbf{Y}; \mathbf{X}) + I(\mathbf{Y}; \tilde{\mathbf{X}} | \mathbf{X}).$$

Rearranging gives

$$-\Delta = I(\mathbf{Y}; \mathbf{X}) - I(\mathbf{Y}; \tilde{\mathbf{X}}) = I(\mathbf{Y}; \mathbf{X} | \tilde{\mathbf{X}}) - I(\mathbf{Y}; \tilde{\mathbf{X}} | \mathbf{X}). \quad (10)$$

The second term on the right is nonnegative, hence

$$-\Delta \leq I(\mathbf{Y}; \mathbf{X} | \tilde{\mathbf{X}}). \quad (11)$$

Using the bound $I(U; V | W) \leq H(V | W)$, we get

$$-\Delta \leq H(\mathbf{X} | \tilde{\mathbf{X}}). \quad (12)$$

Index the matrix entries by a total order \prec on pairs (i, j) and apply the chain rule:

$$H(\mathbf{X} | \tilde{\mathbf{X}}) = \sum_{(i,j)} H(X_{ij} | \tilde{\mathbf{X}}, \{X_{kl} : (k, l) \prec (i, j)\}).$$

Since conditioning reduces entropy,

$$H(\mathbf{X} | \tilde{\mathbf{X}}) \leq \sum_{i,j} H(X_{ij} | \tilde{X}_{ij}). \quad (13)$$

Fix (i, j) and denote $\pi_{ij} = \Pr[X_{ij} = 1]$. Under uniform MCAR,

$$\Pr[\tilde{X}_{ij} = ?] = \mu, \quad \Pr[\tilde{X}_{ij} = x] = (1 - \mu) \Pr[X_{ij} = x], \quad x \in \{0, 1\}.$$

Hence: (i) if $\tilde{X}_{ij} \in \{0, 1\}$ then X_{ij} is revealed, so $H(X_{ij} | \tilde{X}_{ij} \in \{0, 1\}) = 0$; (ii) if $\tilde{X}_{ij} = ?$, then $\Pr[X_{ij} = 1 | \tilde{X}_{ij} = ?] = \pi_{ij}$ and $H(X_{ij} | \tilde{X}_{ij} = ?) = h_2(\pi_{ij})$. Averaging over \tilde{X}_{ij} gives

$$H(X_{ij} | \tilde{X}_{ij}) = \mu h_2(\pi_{ij}). \quad (14)$$

Combining equation 13 and equation 14

$$H(\mathbf{X} | \tilde{\mathbf{X}}) \leq \sum_{i,j} \mu h_2(\pi_{ij}) = nd \mu \cdot \frac{1}{nd} \sum_{i,j} h_2(\pi_{ij}) \leq nd \mu \cdot h_2\left(\frac{1}{nd} \sum_{i,j} \pi_{ij}\right),$$

since h_2 is concave. Note that

$$\frac{1}{nd} \sum_{i,j} \pi_{ij} = \frac{1}{nd} \sum_{i,j} \Pr[X_{ij} = 1] = \mathbb{E}\left[\frac{1}{nd} \sum_{i,j} \mathbb{I}\{X_{ij} = 1\}\right] = 1 - \mathbb{E}[s(\mathbf{X})].$$

Using the symmetry $h_2(u) = h_2(1 - u)$, we conclude

$$H(\mathbf{X} | \tilde{\mathbf{X}}) \leq nd \mu \cdot h_2(\mathbb{E}[s(\mathbf{X})]).$$

Combining with $-\Delta \leq H(\mathbf{X} | \tilde{\mathbf{X}})$ gives

$$-nd \mu h_2(\mathbb{E}[s(\mathbf{X})]) \leq \Delta \leq 0.$$

This concludes the proof. \square

B ADDITIONAL RESULTS ON BENCHMARKS AND PROPOSED DATASETS

This section presents the full plots of the results under the R1 regime introduced in Section 4

Figure 4 shows the complete set of results across all datasets, whose statistics are summarized in Table 3. The top three rows correspond to the classic benchmarks (CORA, CITESEER, PUBMED). Consistently with Proposition 2, models maintain nearly constant F1 scores up to extremely high missingness levels ($\sim 90\%$), confirming that these benchmarks are of limited value for evaluating robustness to missing features.

The bottom four rows correspond to our proposed datasets (SYNTHETIC, AIR, ELECTRIC, TAD-POLE). In these cases, performance degrades much earlier and more severely, highlighting the higher realism and difficulty of our benchmarks.

Table 3: Dataset statistics and feature sparsity. Classic benchmarks (CORA, CITESEER, PUBMED) exhibit extremely sparse bag-of-words features, while our proposed datasets (SYNTHETIC, AIR, ELECTRIC, TADPOLE) provide less sparse representations.

Dataset	#Nodes	#Features	Sparsity ↓	Type of features
CORA	2708	1433	0.9873	BoW (binary)
CITESEER	3327	3703	0.9915	BoW (binary)
PUBMED	19717	500	0.8998	BoW (binary)
SYNTHETIC	1000	5	0.0000	Gaussian
AIR	430	7	0.1615	Raw
ELECTRIC	2000	5	0.2000	Raw
TADPOLE	555	15	0.0000	Raw

C MORE CHALLENGING DATASETS

In Section 3, we introduced the synthetic and real-world datasets employed in our experiments. We now provide additional details on their construction and characteristics.

SYNTHETIC Synthetic dataset based on a Barabási–Albert graph topology. Each node is associated with five real-valued features sampled from a Gaussian distribution. Node labels are generated deterministically by applying a fixed two-layer GCN with hard-coded weights to the complete feature matrix. This construction ensures that the ground-truth labeling function is fully expressible by a GNN, allowing models to achieve near-perfect accuracy in the absence of missingness. The resulting task is a binary node classification problem, with classes separated according to structured feature combinations defined by the fixed GCN. This controlled setup provides a principled testbed to isolate and analyze the effects of different missingness mechanisms, while preserving a well-defined ground truth.

AIR Dataset (Zheng et al., 2015) built from a network of air quality monitoring stations deployed in an urban area. Each node corresponds to a station and is associated with a set of environmental measurements. The node features include both air pollutant concentrations (CO , NO_2 , PM_{10} , O_3 , SO_2) and meteorological variables (temperature, humidity, wind speed, wind direction). Edges are constructed based on the geographical distance between stations, with two nodes connected if their distance is below a given threshold. The target variable is derived from the $\text{PM}_{2.5}$ concentration, which is discretized into three balanced categories (low, medium, high) according to the distribution of observed values. This formulation allows us to frame the problem as a semi-supervised node classification task with three classes.

ELECTRIC Dataset (Birchfield et al., 2016; Baek & Birchfield, 2023) derived from a large-scale model of the Texas power grid. Nodes correspond to buses in the electrical network, each enriched with both structural and operational attributes. The node features include identifiers (area, zone), electrical measurements (voltage magnitude, voltage angle), and a topological property (betweenness centrality). Edges are constructed directly from the transmission lines specified in the raw grid data, connecting pairs of buses. The classification target is the nominal voltage level of each bus (base kV), which we discretize into three categories: low voltage (<100 kV), medium voltage (100–200 kV), and high voltage (>200 kV). This setup results in a three-class node classification problem reflecting operational conditions across the grid.

TADPOLE The TADPOLE dataset (Zhu et al., 2019) originates from the TADPOLE challenge, which provides longitudinal clinical and imaging data for patients at risk of developing Alzheimer’s disease. In our graph formulation, each node corresponds to a patient and is associated with a set of features encompassing clinical scores, cerebrospinal fluid (CSF) biomarkers, and neuroimaging measures such as MRI- and PET-derived variables. Since the original dataset does not provide graph connectivity, we construct edges using a k -nearest neighbors approach over the most informative biomarkers, so that patients with similar profiles are connected. The target variable is the diagnostic label, categorized into three classes (cognitively normal, mild cognitive impairment, Alzheimer’s disease). This results in a semi-supervised node classification problem where the goal is to pre-

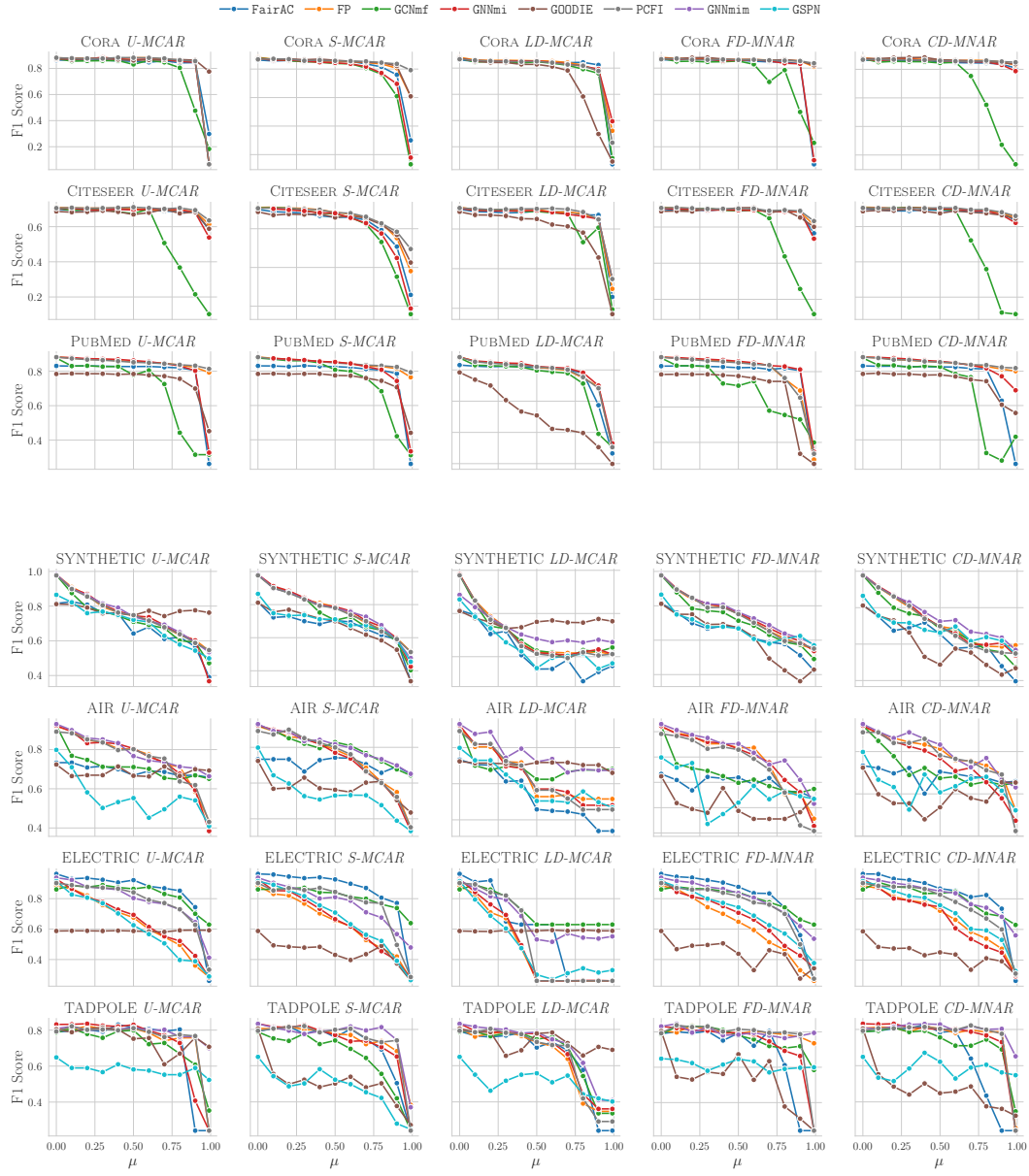


Figure 4: F1 score as a function of feature missingness (μ) for both classic benchmarks (top three rows) and our proposed datasets (bottom four rows), under the mechanisms described in Section 4. Classic benchmarks show almost no degradation until extremely high μ , while the proposed datasets reveal model weaknesses at more realistic missingness levels. Tables for numeric results are in App.

dict the diagnostic status of patients based on multimodal biomedical features and patient similarity structure.

Table 3 reports, for each dataset, the number of nodes, number of features, feature sparsity, and the type of features. While the number of nodes and features may seem small compared to standard benchmark graph datasets, we emphasize that using real features (as in AIR, ELECTRIC, and TADPOLE) is more realistic in the context of feature missingness. In fact, it is not meaningful to study missingness on pre-computed embeddings, since embeddings are typically high-dimensional representations mapped to wide feature spaces and are not expected to exhibit missingness in practice.

D EXPERIMENTAL DETAILS

All baseline and competitor methods are implemented using the official code released in their respective repositories, following the recommended training protocols and hyperparameter settings. For GNNmi and GNNmim, we adopt a standard GNN architecture where the convolutional layer type (Table 4), the number of layers (1-3), the learning rate (10^{-4} - 10^{-2}), and the weight decay (10^{-5} - 10^{-3}) are tuned via grid search on the validation set. All models are trained on the same data splits with early stopping to ensure a fair comparison.

Table 4: Best GNN encoder selected within GNNmim for each dataset and missingness mechanism.

Dataset	U-MCAR	S-MCAR	LD-MCAR	FD-MNAR	CD-MNAR
SYNTHETIC	GCN	GCN	GraphSAGE	GCN	GCN
AIR	GraphSAGE	GraphSAGE	GraphSAGE	GraphSAGE	GraphSAGE
ELECTRIC	GIN	GIN	GraphSAGE	GIN	GIN
TADPOLE	GCN	GraphSAGE	GraphSAGE	GraphSAGE	GCN

E SCALING THE SYNTHETIC DATASET

In this section, we analyze what happens when either the number of features or the number of nodes in the synthetic dataset is increased. To this end, we constructed three additional synthetic datasets (SYNTHETIC2, SYNTHETIC3, SYNTHETIC4) following the same design principles as SYNTHETIC. Table 5 reports their statistics.

As shown in Figure 5 the behavior of the models in this larger-scale setting is consistent with the one observed in our original setup. In this case, we experimented with the *uniform random missingness* mechanism, and we observe a monotonic decrease in performance for all models as the missingness rate μ increases. This confirms that dataset size does not affect the overall trend of performance degradation under feature missingness.

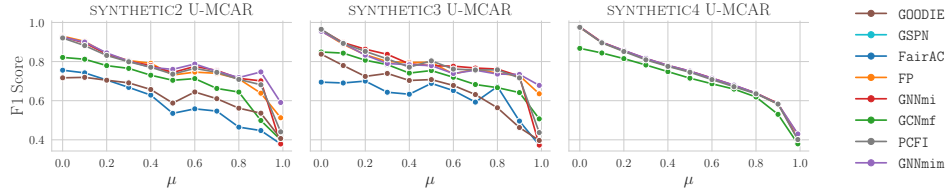


Figure 5: F1 score as a function of feature missingness (μ) for additional synthetic datasets generated with the same procedure as SYNTHETIC, but with either an increased number of nodes or features. For SYNTHETIC4, the FairAC model is not reported since training exceeded the 12-hour time limit, while GOODIE is excluded due to out-of-memory errors.

Table 5: Datasets information.

Dataset	#Nodes	#Features	Sparsity ↓	Type of features
SYNTHETIC	1000	5	0.0000	Gaussian
SYNTHETIC2	1000	20	0.0000	Gaussian
SYNTHETIC3	1000	50	0.0000	Gaussian
SYNTHETIC4	50000	5	0.0000	Gaussian

F COMPLETE RESULT TABLES – R1 REGIME

Table 6: F1 scores for CORA under mechanism *U-MCAR* and varying μ (GSPNis not reported as it is not designed for categorical features).

μ	GOODIE	FairAC	FP	GNNmi	GCNm	PCFI
0.00	0.875 (± 0.00)	0.863 (± 0.01)	0.882 (± 0.00)	0.873 (± 0.00)	0.875 (± 0.00)	0.882 (± 0.00)
0.10	0.867 (± 0.00)	0.866 (± 0.00)	0.877 (± 0.00)	0.876 (± 0.00)	0.856 (± 0.00)	0.878 (± 0.00)
0.20	0.875 (± 0.00)	0.862 (± 0.00)	0.878 (± 0.00)	0.873 (± 0.00)	0.858 (± 0.00)	0.877 (± 0.00)
0.30	0.873 (± 0.00)	0.865 (± 0.00)	0.881 (± 0.00)	0.885 (± 0.00)	0.860 (± 0.00)	0.876 (± 0.00)
0.40	0.869 (± 0.00)	0.857 (± 0.00)	0.878 (± 0.00)	0.873 (± 0.00)	0.860 (± 0.00)	0.884 (± 0.00)
0.50	0.861 (± 0.00)	0.856 (± 0.00)	0.882 (± 0.00)	0.867 (± 0.00)	0.831 (± 0.00)	0.882 (± 0.00)
0.60	0.866 (± 0.00)	0.847 (± 0.00)	0.882 (± 0.00)	0.871 (± 0.00)	0.862 (± 0.00)	0.881 (± 0.00)
0.70	0.866 (± 0.00)	0.858 (± 0.00)	0.869 (± 0.00)	0.865 (± 0.00)	0.847 (± 0.00)	0.877 (± 0.00)
0.80	0.868 (± 0.00)	0.843 (± 0.00)	0.864 (± 0.00)	0.854 (± 0.00)	0.805 (± 0.00)	0.863 (± 0.00)
0.90	0.864 (± 0.00)	0.845 (± 0.00)	0.860 (± 0.00)	0.848 (± 0.00)	0.476 (± 0.00)	0.856 (± 0.00)
0.99	0.776 (± 0.00)	0.298 (± 0.00)	0.066 (± 0.00)	0.066 (± 0.00)	0.183 (± 0.00)	0.065 (± 0.00)

Table 7: F1 scores for CORA under mechanism *S-MCAR* and varying μ (GSPNis not reported as it is not designed for categorical features).

μ	GOODIE	FairAC	FP	GNNmi	GCNm	PCFI
0.00	0.875 (± 0.00)	0.863 (± 0.01)	0.882 (± 0.00)	0.872 (± 0.00)	0.875 (± 0.00)	0.868 (± 0.00)
0.10	0.868 (± 0.00)	0.857 (± 0.00)	0.869 (± 0.00)	0.862 (± 0.00)	0.869 (± 0.00)	0.872 (± 0.00)
0.20	0.872 (± 0.00)	0.860 (± 0.00)	0.863 (± 0.00)	0.863 (± 0.00)	0.858 (± 0.00)	0.869 (± 0.00)
0.30	0.865 (± 0.00)	0.850 (± 0.00)	0.854 (± 0.00)	0.855 (± 0.00)	0.852 (± 0.00)	0.858 (± 0.00)
0.40	0.870 (± 0.00)	0.857 (± 0.00)	0.859 (± 0.00)	0.848 (± 0.00)	0.848 (± 0.00)	0.862 (± 0.00)
0.50	0.862 (± 0.00)	0.854 (± 0.00)	0.854 (± 0.00)	0.844 (± 0.00)	0.839 (± 0.00)	0.858 (± 0.00)
0.60	0.855 (± 0.00)	0.854 (± 0.00)	0.853 (± 0.00)	0.837 (± 0.00)	0.837 (± 0.00)	0.856 (± 0.00)
0.70	0.847 (± 0.00)	0.836 (± 0.00)	0.845 (± 0.00)	0.817 (± 0.00)	0.807 (± 0.00)	0.854 (± 0.00)
0.80	0.845 (± 0.00)	0.815 (± 0.00)	0.836 (± 0.00)	0.772 (± 0.00)	0.764 (± 0.00)	0.845 (± 0.00)
0.90	0.822 (± 0.00)	0.760 (± 0.00)	0.806 (± 0.00)	0.696 (± 0.00)	0.610 (± 0.00)	0.836 (± 0.00)
0.99	0.609 (± 0.00)	0.300 (± 0.00)	0.606 (± 0.00)	0.179 (± 0.00)	0.132 (± 0.00)	0.792 (± 0.00)

Table 8: F1 scores for CORA under mechanism *LD-MCAR* and varying μ (GSPNis not reported as it is not designed for categorical features).

μ	GOODIE	FairAC	FP	GNNmi	GCNm	PCFI
0.00	0.875 (± 0.00)	0.863 (± 0.01)	0.882 (± 0.00)	0.873 (± 0.00)	0.875 (± 0.00)	0.868 (± 0.00)
0.10	0.852 (± 0.00)	0.851 (± 0.00)	0.862 (± 0.00)	0.857 (± 0.00)	0.846 (± 0.00)	0.860 (± 0.00)
0.20	0.843 (± 0.00)	0.854 (± 0.00)	0.859 (± 0.00)	0.854 (± 0.00)	0.850 (± 0.00)	0.855 (± 0.00)
0.30	0.843 (± 0.00)	0.856 (± 0.00)	0.859 (± 0.00)	0.855 (± 0.00)	0.846 (± 0.00)	0.852 (± 0.00)
0.40	0.828 (± 0.00)	0.854 (± 0.00)	0.858 (± 0.00)	0.853 (± 0.00)	0.838 (± 0.00)	0.849 (± 0.00)
0.50	0.828 (± 0.00)	0.854 (± 0.00)	0.855 (± 0.00)	0.855 (± 0.00)	0.848 (± 0.00)	0.852 (± 0.00)
0.60	0.812 (± 0.00)	0.847 (± 0.00)	0.853 (± 0.00)	0.844 (± 0.00)	0.837 (± 0.00)	0.841 (± 0.00)
0.70	0.782 (± 0.00)	0.841 (± 0.00)	0.842 (± 0.00)	0.831 (± 0.00)	0.822 (± 0.00)	0.827 (± 0.00)
0.80	0.584 (± 0.00)	0.844 (± 0.00)	0.822 (± 0.00)	0.815 (± 0.00)	0.792 (± 0.00)	0.818 (± 0.00)
0.90	0.297 (± 0.00)	0.824 (± 0.00)	0.777 (± 0.00)	0.793 (± 0.00)	0.760 (± 0.00)	0.778 (± 0.00)
0.99	0.088 (± 0.00)	0.066 (± 0.00)	0.322 (± 0.00)	0.395 (± 0.00)	0.113 (± 0.00)	0.231 (± 0.00)

Table 9: F1 scores for CORA under mechanism *FD-MNAR* and varying μ (GSPNis not reported as it is not designed for categorical features).

μ	GOODIE	FairAC	FP	GNNmi	GCNm	PCFI
0.00	0.875 (± 0.00)	0.863 (± 0.01)	0.882 (± 0.00)	0.873 (± 0.00)	0.875 (± 0.00)	0.868 (± 0.00)
0.10	0.872 (± 0.01)	0.862 (± 0.01)	0.873 (± 0.01)	0.868 (± 0.01)	0.851 (± 0.01)	0.873 (± 0.00)
0.20	0.879 (± 0.00)	0.870 (± 0.01)	0.874 (± 0.00)	0.865 (± 0.01)	0.853 (± 0.01)	0.863 (± 0.01)
0.30	0.880 (± 0.00)	0.864 (± 0.01)	0.869 (± 0.00)	0.867 (± 0.01)	0.847 (± 0.01)	0.864 (± 0.01)
0.40	0.869 (± 0.01)	0.855 (± 0.01)	0.864 (± 0.01)	0.856 (± 0.01)	0.849 (± 0.00)	0.866 (± 0.01)
0.50	0.865 (± 0.01)	0.860 (± 0.01)	0.866 (± 0.01)	0.859 (± 0.01)	0.854 (± 0.01)	0.863 (± 0.01)
0.60	0.866 (± 0.01)	0.853 (± 0.01)	0.865 (± 0.01)	0.863 (± 0.01)	0.829 (± 0.02)	0.864 (± 0.01)
0.70	0.859 (± 0.01)	0.847 (± 0.00)	0.862 (± 0.01)	0.853 (± 0.00)	0.695 (± 0.14)	0.860 (± 0.00)
0.80	0.865 (± 0.01)	0.845 (± 0.01)	0.861 (± 0.01)	0.837 (± 0.00)	0.785 (± 0.05)	0.857 (± 0.01)
0.90	0.854 (± 0.01)	0.833 (± 0.01)	0.855 (± 0.00)	0.833 (± 0.00)	0.465 (± 0.21)	0.854 (± 0.01)
0.99	0.822 (± 0.01)	0.066 (± 0.00)	0.810 (± 0.02)	0.098 (± 0.01)	0.230 (± 0.05)	0.837 (± 0.02)

Table 10: F1 scores for CORA under mechanism *CD-MNAR* and varying μ (GSPNis not reported as it is not designed for categorical features).

μ	GOODIE	FairAC	FP	GNNmi	GCNm	PCFI
0.00	0.875 (± 0.00)	0.863 (± 0.01)	0.882 (± 0.00)	0.873 (± 0.00)	0.875 (± 0.00)	0.868 (± 0.00)
0.10	0.875 (± 0.00)	0.864 (± 0.01)	0.870 (± 0.01)	0.862 (± 0.01)	0.850 (± 0.00)	0.869 (± 0.01)
0.20	0.881 (± 0.01)	0.865 (± 0.00)	0.874 (± 0.01)	0.868 (± 0.01)	0.856 (± 0.01)	0.869 (± 0.01)
0.30	0.882 (± 0.00)	0.858 (± 0.00)	0.873 (± 0.00)	0.871 (± 0.01)	0.854 (± 0.00)	0.866 (± 0.01)
0.40	0.884 (± 0.01)	0.862 (± 0.01)	0.870 (± 0.00)	0.864 (± 0.00)	0.853 (± 0.01)	0.865 (± 0.01)
0.50	0.867 (± 0.01)	0.852 (± 0.01)	0.867 (± 0.00)	0.861 (± 0.00)	0.844 (± 0.02)	0.861 (± 0.01)
0.60	0.864 (± 0.00)	0.847 (± 0.00)	0.860 (± 0.01)	0.856 (± 0.01)	0.849 (± 0.00)	0.857 (± 0.00)
0.70	0.860 (± 0.01)	0.845 (± 0.01)	0.864 (± 0.01)	0.852 (± 0.01)	0.753 (± 0.12)	0.856 (± 0.01)
0.80	0.853 (± 0.01)	0.844 (± 0.02)	0.862 (± 0.01)	0.852 (± 0.01)	0.551 (± 0.10)	0.861 (± 0.01)
0.90	0.848 (± 0.01)	0.835 (± 0.01)	0.852 (± 0.00)	0.831 (± 0.01)	0.271 (± 0.23)	0.855 (± 0.01)
0.99	0.836 (± 0.01)	0.810 (± 0.01)	0.828 (± 0.01)	0.788 (± 0.02)	0.135 (± 0.05)	0.849 (± 0.01)

Table 11: F1 scores for CITESEER under mechanism *U-MCAR* and varying μ (GSPNis not reported as it is not designed for categorical features).

μ	GOODIE	FairAC	FP	GNNmi	GCNm	PCFI
0.00	0.687 (± 0.00)	0.700 (± 0.00)	0.710 (± 0.02)	0.704 (± 0.02)	0.707 (± 0.00)	0.706 (± 0.02)
0.10	0.682 (± 0.00)	0.693 (± 0.00)	0.707 (± 0.00)	0.705 (± 0.00)	0.692 (± 0.00)	0.708 (± 0.00)
0.20	0.684 (± 0.00)	0.693 (± 0.00)	0.706 (± 0.00)	0.695 (± 0.00)	0.698 (± 0.00)	0.705 (± 0.00)
0.30	0.691 (± 0.00)	0.691 (± 0.00)	0.705 (± 0.00)	0.696 (± 0.00)	0.697 (± 0.00)	0.706 (± 0.00)
0.40	0.685 (± 0.00)	0.700 (± 0.00)	0.706 (± 0.00)	0.698 (± 0.00)	0.684 (± 0.00)	0.708 (± 0.00)
0.50	0.669 (± 0.00)	0.697 (± 0.00)	0.702 (± 0.00)	0.695 (± 0.00)	0.675 (± 0.00)	0.711 (± 0.00)
0.60	0.680 (± 0.00)	0.695 (± 0.00)	0.697 (± 0.00)	0.699 (± 0.00)	0.700 (± 0.00)	0.707 (± 0.00)
0.70	0.699 (± 0.00)	0.688 (± 0.00)	0.694 (± 0.00)	0.700 (± 0.00)	0.507 (± 0.00)	0.701 (± 0.00)
0.80	0.675 (± 0.00)	0.687 (± 0.00)	0.694 (± 0.00)	0.696 (± 0.00)	0.368 (± 0.00)	0.707 (± 0.00)
0.90	0.684 (± 0.00)	0.680 (± 0.00)	0.686 (± 0.00)	0.680 (± 0.00)	0.215 (± 0.00)	0.694 (± 0.00)
0.99	0.588 (± 0.00)	0.584 (± 0.00)	0.613 (± 0.00)	0.539 (± 0.00)	0.102 (± 0.00)	0.636 (± 0.00)

Table 12: F1 scores for CITESEER under mechanism *S-MCAR* and varying μ (GSPNis not reported as it is not designed for categorical features).

μ	GOODIE	FairAC	FP	GNNmi	GCNmf	PCFI
0.00	0.687 (± 0.00)	0.700 (± 0.00)	0.710 (± 0.02)	-	0.707 (± 0.00)	0.706 (± 0.02)
0.10	0.670 (± 0.00)	0.688 (± 0.00)	0.711 (± 0.00)	0.703 (± 0.00)	0.708 (± 0.00)	0.708 (± 0.00)
0.20	0.675 (± 0.00)	0.685 (± 0.00)	0.707 (± 0.00)	0.697 (± 0.00)	0.707 (± 0.00)	0.706 (± 0.00)
0.30	0.673 (± 0.00)	0.681 (± 0.00)	0.705 (± 0.00)	0.692 (± 0.00)	0.693 (± 0.00)	0.701 (± 0.00)
0.40	0.677 (± 0.00)	0.667 (± 0.00)	0.698 (± 0.00)	0.682 (± 0.00)	0.682 (± 0.00)	0.698 (± 0.00)
0.50	0.658 (± 0.00)	0.659 (± 0.00)	0.685 (± 0.00)	0.680 (± 0.00)	0.676 (± 0.00)	0.683 (± 0.00)
0.60	0.667 (± 0.00)	0.659 (± 0.00)	0.676 (± 0.00)	0.656 (± 0.00)	0.659 (± 0.00)	0.680 (± 0.00)
0.70	0.655 (± 0.00)	0.646 (± 0.00)	0.656 (± 0.00)	0.629 (± 0.00)	0.624 (± 0.00)	0.662 (± 0.00)
0.80	0.621 (± 0.00)	0.593 (± 0.00)	0.629 (± 0.00)	0.575 (± 0.00)	0.531 (± 0.00)	0.628 (± 0.00)
0.90	0.568 (± 0.00)	0.508 (± 0.00)	0.552 (± 0.00)	0.449 (± 0.00)	0.352 (± 0.00)	0.584 (± 0.00)
0.99	0.425 (± 0.00)	0.258 (± 0.00)	0.381 (± 0.00)	0.188 (± 0.00)	0.159 (± 0.00)	0.495 (± 0.00)

Table 13: F1 scores for CITESEER under mechanism *LD-MCAR* and varying μ (GSPNis not reported as it is not designed for categorical features).

μ	GOODIE	FairAC	FP	GNNmi	GCNmf	PCFI
0.00	0.687 (± 0.00)	0.700 (± 0.00)	0.710 (± 0.02)	0.704 (± 0.02)	0.707 (± 0.00)	0.706 (± 0.02)
0.10	0.671 (± 0.00)	0.687 (± 0.00)	0.698 (± 0.00)	0.694 (± 0.00)	0.693 (± 0.00)	0.702 (± 0.00)
0.20	0.670 (± 0.00)	0.686 (± 0.00)	0.699 (± 0.00)	0.691 (± 0.00)	0.696 (± 0.00)	0.698 (± 0.00)
0.30	0.666 (± 0.00)	0.682 (± 0.00)	0.697 (± 0.00)	0.691 (± 0.00)	0.694 (± 0.00)	0.699 (± 0.00)
0.40	0.652 (± 0.00)	0.683 (± 0.00)	0.698 (± 0.00)	0.691 (± 0.00)	0.688 (± 0.00)	0.701 (± 0.00)
0.50	0.650 (± 0.00)	0.690 (± 0.00)	0.699 (± 0.00)	0.693 (± 0.00)	0.688 (± 0.00)	0.702 (± 0.00)
0.60	0.622 (± 0.00)	0.686 (± 0.00)	0.685 (± 0.00)	0.685 (± 0.00)	0.681 (± 0.00)	0.704 (± 0.00)
0.70	0.613 (± 0.00)	0.687 (± 0.00)	0.686 (± 0.00)	0.674 (± 0.00)	0.677 (± 0.00)	0.700 (± 0.00)
0.80	0.582 (± 0.00)	0.671 (± 0.00)	0.677 (± 0.00)	0.664 (± 0.00)	0.534 (± 0.00)	0.686 (± 0.00)
0.90	0.456 (± 0.00)	0.671 (± 0.00)	0.650 (± 0.00)	0.650 (± 0.00)	0.607 (± 0.00)	0.648 (± 0.00)
0.99	0.171 (± 0.00)	0.257 (± 0.00)	0.298 (± 0.00)	0.346 (± 0.00)	0.195 (± 0.00)	0.348 (± 0.00)

Table 14: F1 scores for CITESEER under mechanism *FD-MNAR* and varying μ (GSPNis not reported as it is not designed for categorical features).

μ	GOODIE	FairAC	FP	GNNmi	GCNmf	PCFI
0.00	0.687 (± 0.00)	0.700 (± 0.00)	0.710 (± 0.02)	0.704 (± 0.02)	0.707 (± 0.00)	0.706 (± 0.02)
0.10	0.689 (± 0.03)	0.691 (± 0.03)	0.706 (± 0.02)	0.699 (± 0.02)	0.699 (± 0.02)	0.708 (± 0.03)
0.20	0.686 (± 0.02)	0.698 (± 0.03)	0.703 (± 0.02)	0.697 (± 0.02)	0.696 (± 0.02)	0.704 (± 0.02)
0.30	0.701 (± 0.04)	0.690 (± 0.03)	0.701 (± 0.03)	0.693 (± 0.02)	0.704 (± 0.02)	0.700 (± 0.03)
0.40	0.696 (± 0.04)	0.699 (± 0.04)	0.695 (± 0.02)	0.695 (± 0.02)	0.692 (± 0.03)	0.701 (± 0.03)
0.50	0.707 (± 0.03)	0.688 (± 0.04)	0.698 (± 0.03)	0.693 (± 0.03)	0.690 (± 0.02)	0.702 (± 0.03)
0.60	0.708 (± 0.02)	0.694 (± 0.03)	0.691 (± 0.03)	0.693 (± 0.03)	0.696 (± 0.02)	0.702 (± 0.03)
0.70	0.678 (± 0.04)	0.688 (± 0.03)	0.688 (± 0.03)	0.686 (± 0.02)	0.649 (± 0.03)	0.690 (± 0.04)
0.80	0.695 (± 0.03)	0.689 (± 0.04)	0.689 (± 0.02)	0.685 (± 0.02)	0.437 (± 0.27)	0.694 (± 0.03)
0.90	0.653 (± 0.03)	0.681 (± 0.04)	0.682 (± 0.02)	0.687 (± 0.03)	0.257 (± 0.17)	0.689 (± 0.02)
0.99	0.601 (± 0.01)	0.566 (± 0.01)	0.611 (± 0.01)	0.535 (± 0.02)	0.118 (± 0.04)	0.633 (± 0.01)

Table 15: F1 scores for CITESEER under mechanism *CD-MNAR* and varying μ (GSPNis not reported as it is not designed for categorical features).

μ	GOODIE	FairAC	FP	GNNmi	GCNm	PCFI
0.00	0.687 (± 0.00)	0.700 (± 0.05)	0.710 (± 0.02)	0.704 (± 0.02)	0.707 (± 0.00)	0.706 (± 0.02)
0.10	0.692 (± 0.04)	0.696 (± 0.04)	0.708 (± 0.02)	0.705 (± 0.02)	0.702 (± 0.03)	0.705 (± 0.02)
0.20	0.690 (± 0.04)	0.689 (± 0.04)	0.703 (± 0.03)	0.702 (± 0.02)	0.705 (± 0.02)	0.704 (± 0.02)
0.30	0.700 (± 0.02)	0.689 (± 0.04)	0.708 (± 0.03)	0.706 (± 0.02)	0.708 (± 0.02)	0.705 (± 0.02)
0.40	0.687 (± 0.04)	0.695 (± 0.04)	0.707 (± 0.03)	0.704 (± 0.02)	0.703 (± 0.03)	0.704 (± 0.03)
0.50	0.675 (± 0.03)	0.692 (± 0.03)	0.699 (± 0.03)	0.700 (± 0.03)	0.697 (± 0.02)	0.706 (± 0.03)
0.60	0.689 (± 0.03)	0.689 (± 0.03)	0.702 (± 0.03)	0.699 (± 0.03)	0.693 (± 0.03)	0.706 (± 0.03)
0.70	0.681 (± 0.03)	0.685 (± 0.03)	0.692 (± 0.03)	0.691 (± 0.03)	0.522 (± 0.20)	0.696 (± 0.03)
0.80	0.676 (± 0.05)	0.685 (± 0.03)	0.690 (± 0.03)	0.689 (± 0.02)	0.359 (± 0.15)	0.696 (± 0.04)
0.90	0.665 (± 0.02)	0.681 (± 0.03)	0.677 (± 0.03)	0.666 (± 0.03)	0.113 (± 0.06)	0.681 (± 0.03)
0.99	0.645 (± 0.03)	0.631 (± 0.02)	0.652 (± 0.02)	0.621 (± 0.02)	0.104 (± 0.06)	0.660 (± 0.02)

Table 16: F1 scores for PUBMED under mechanism *U-MCAR* and varying μ (GSPNis not reported as it is not designed for categorical features).

μ	GOODIE	FairAC	FP	GNNmi	GCNm	PCFI
0.00	0.784 (± 0.01)	0.831 (± 0.00)	0.883 (± 0.00)	0.881 (± 0.00)	0.877 (± 0.00)	0.882 (± 0.00)
0.10	0.787 (± 0.00)	0.830 (± 0.00)	0.877 (± 0.00)	0.879 (± 0.00)	0.830 (± 0.00)	0.874 (± 0.00)
0.20	0.786 (± 0.00)	0.831 (± 0.00)	0.868 (± 0.00)	0.873 (± 0.00)	0.832 (± 0.00)	0.868 (± 0.00)
0.30	0.785 (± 0.00)	0.830 (± 0.00)	0.870 (± 0.00)	0.872 (± 0.00)	0.827 (± 0.00)	0.864 (± 0.00)
0.40	0.782 (± 0.00)	0.828 (± 0.00)	0.861 (± 0.00)	0.869 (± 0.00)	0.828 (± 0.00)	0.858 (± 0.00)
0.50	0.784 (± 0.00)	0.827 (± 0.00)	0.856 (± 0.00)	0.862 (± 0.00)	0.778 (± 0.00)	0.852 (± 0.00)
0.60	0.777 (± 0.00)	0.828 (± 0.00)	0.851 (± 0.00)	0.855 (± 0.00)	0.805 (± 0.00)	0.849 (± 0.00)
0.70	0.772 (± 0.00)	0.824 (± 0.00)	0.847 (± 0.00)	0.845 (± 0.00)	0.726 (± 0.00)	0.844 (± 0.00)
0.80	0.756 (± 0.00)	0.819 (± 0.00)	0.836 (± 0.00)	0.832 (± 0.00)	0.443 (± 0.00)	0.837 (± 0.00)
0.90	0.700 (± 0.00)	0.806 (± 0.00)	0.822 (± 0.00)	0.803 (± 0.00)	0.315 (± 0.00)	0.832 (± 0.00)
0.99	0.452 (± 0.00)	0.262 (± 0.00)	0.793 (± 0.00)	0.327 (± 0.00)	0.315 (± 0.00)	0.814 (± 0.00)

Table 17: F1 scores for PUBMED under mechanism *S-MCAR* and varying μ (GSPNis not reported as it is not designed for categorical features).

μ	GOODIE	FairAC	FP	GNNmi	GCNm	PCFI
0.00	0.784 (± 0.01)	0.831 (± 0.00)	0.883 (± 0.00)	-	0.877 (± 0.00)	0.882 (± 0.00)
0.10	0.786 (± 0.00)	0.831 (± 0.00)	0.875 (± 0.00)	0.875 (± 0.00)	0.870 (± 0.00)	0.871 (± 0.00)
0.20	0.783 (± 0.00)	0.827 (± 0.00)	0.869 (± 0.00)	0.870 (± 0.00)	0.861 (± 0.00)	0.867 (± 0.00)
0.30	0.785 (± 0.00)	0.832 (± 0.00)	0.863 (± 0.00)	0.865 (± 0.00)	0.861 (± 0.00)	0.863 (± 0.00)
0.40	0.785 (± 0.00)	0.828 (± 0.00)	0.856 (± 0.00)	0.857 (± 0.00)	0.848 (± 0.00)	0.856 (± 0.00)
0.50	0.775 (± 0.00)	0.827 (± 0.00)	0.853 (± 0.00)	0.854 (± 0.00)	0.808 (± 0.00)	0.848 (± 0.00)
0.60	0.774 (± 0.00)	0.822 (± 0.00)	0.843 (± 0.00)	0.845 (± 0.00)	0.798 (± 0.00)	0.843 (± 0.00)
0.70	0.760 (± 0.00)	0.813 (± 0.00)	0.832 (± 0.00)	0.827 (± 0.00)	0.762 (± 0.00)	0.836 (± 0.00)
0.80	0.744 (± 0.00)	0.806 (± 0.00)	0.828 (± 0.00)	0.808 (± 0.00)	0.683 (± 0.00)	0.832 (± 0.00)
0.90	0.706 (± 0.00)	0.786 (± 0.00)	0.815 (± 0.00)	0.743 (± 0.00)	0.421 (± 0.00)	0.825 (± 0.00)
0.99	0.441 (± 0.00)	0.259 (± 0.00)	0.765 (± 0.00)	0.333 (± 0.00)	0.310 (± 0.00)	0.794 (± 0.00)

Table 18: F1 scores for PUBMED under mechanism *LD-MCAR* and varying μ (GSPNis not reported as it is not designed for categorical features).

μ	GOODIE	FairAC	FP	GNNmi	GCNmf	PCFI
0.00	0.784 (± 0.01)	0.831 (± 0.00)	0.883 (± 0.00)	0.881 (± 0.00)	0.877 (± 0.00)	0.882 (± 0.00)
0.10	0.738 (± 0.00)	0.824 (± 0.00)	0.855 (± 0.00)	0.857 (± 0.00)	0.830 (± 0.00)	0.852 (± 0.00)
0.20	0.700 (± 0.00)	0.820 (± 0.00)	0.845 (± 0.00)	0.851 (± 0.00)	0.828 (± 0.00)	0.844 (± 0.00)
0.30	0.607 (± 0.00)	0.823 (± 0.00)	0.843 (± 0.00)	0.844 (± 0.00)	0.823 (± 0.00)	0.836 (± 0.00)
0.40	0.534 (± 0.00)	0.821 (± 0.00)	0.834 (± 0.00)	0.842 (± 0.00)	0.818 (± 0.00)	0.830 (± 0.00)
0.50	0.509 (± 0.00)	0.814 (± 0.00)	0.818 (± 0.00)	0.823 (± 0.00)	0.797 (± 0.00)	0.820 (± 0.00)
0.60	0.422 (± 0.00)	0.812 (± 0.00)	0.808 (± 0.00)	0.816 (± 0.00)	0.787 (± 0.00)	0.812 (± 0.00)
0.70	0.415 (± 0.00)	0.802 (± 0.00)	0.797 (± 0.00)	0.811 (± 0.00)	0.779 (± 0.00)	0.801 (± 0.00)
0.80	0.396 (± 0.00)	0.779 (± 0.00)	0.749 (± 0.00)	0.783 (± 0.00)	0.713 (± 0.00)	0.754 (± 0.00)
0.90	0.306 (± 0.00)	0.574 (± 0.00)	0.693 (± 0.00)	0.700 (± 0.00)	0.391 (± 0.00)	0.683 (± 0.00)
0.99	0.198 (± 0.00)	0.266 (± 0.00)	0.303 (± 0.00)	0.330 (± 0.00)	0.306 (± 0.00)	0.305 (± 0.00)

Table 19: F1 scores for PUBMED under mechanism *FD-MNAR* and varying μ (GSPNis not reported as it is not designed for categorical features).

μ	GOODIE	FairAC	FP	GNNmi	GCNmf	PCFI
0.00	0.784 (± 0.01)	0.831 (± 0.00)	0.883 (± 0.00)	0.881 (± 0.00)	0.877 (± 0.00)	0.882 (± 0.00)
0.10	0.785 (± 0.02)	0.832 (± 0.00)	0.876 (± 0.01)	0.880 (± 0.01)	0.834 (± 0.00)	0.874 (± 0.01)
0.20	0.785 (± 0.02)	0.834 (± 0.00)	0.869 (± 0.00)	0.875 (± 0.00)	0.832 (± 0.00)	0.869 (± 0.01)
0.30	0.785 (± 0.02)	0.830 (± 0.00)	0.865 (± 0.00)	0.870 (± 0.00)	0.829 (± 0.00)	0.860 (± 0.00)
0.40	0.780 (± 0.01)	0.827 (± 0.00)	0.860 (± 0.00)	0.866 (± 0.00)	0.733 (± 0.11)	0.856 (± 0.00)
0.50	0.775 (± 0.02)	0.822 (± 0.00)	0.853 (± 0.00)	0.859 (± 0.00)	0.720 (± 0.12)	0.850 (± 0.00)
0.60	0.763 (± 0.02)	0.824 (± 0.01)	0.847 (± 0.01)	0.850 (± 0.00)	0.746 (± 0.04)	0.842 (± 0.00)
0.70	0.745 (± 0.03)	0.813 (± 0.00)	0.836 (± 0.00)	0.834 (± 0.00)	0.579 (± 0.25)	0.837 (± 0.00)
0.80	0.745 (± 0.03)	0.819 (± 0.00)	0.759 (± 0.04)	0.829 (± 0.00)	0.555 (± 0.14)	0.764 (± 0.00)
0.90	0.336 (± 0.01)	0.806 (± 0.00)	0.693 (± 0.01)	0.812 (± 0.00)	0.529 (± 0.13)	0.653 (± 0.00)
0.99	0.278 (± 0.01)	0.282 (± 0.01)	0.303 (± 0.05)	0.347 (± 0.00)	0.399 (± 0.33)	0.335 (± 0.01)

Table 20: F1 scores for PUBMED under mechanism *CD-MNAR* and varying μ (GSPNis not reported as it is not designed for categorical features).

μ	GOODIE	FairAC	FP	GNNmi	GCNmf	PCFI
0.00	0.784 (± 0.01)	0.831 (± 0.00)	0.883 (± 0.00)	0.881 (± 0.00)	0.877 (± 0.00)	0.882 (± 0.00)
0.10	0.789 (± 0.02)	0.829 (± 0.00)	0.878 (± 0.00)	0.880 (± 0.00)	0.835 (± 0.00)	0.877 (± 0.00)
0.20	0.783 (± 0.01)	0.830 (± 0.00)	0.870 (± 0.00)	0.876 (± 0.00)	0.834 (± 0.00)	0.867 (± 0.01)
0.30	0.783 (± 0.02)	0.828 (± 0.00)	0.863 (± 0.00)	0.871 (± 0.00)	0.823 (± 0.00)	0.866 (± 0.00)
0.40	0.777 (± 0.02)	0.826 (± 0.00)	0.858 (± 0.00)	0.863 (± 0.00)	0.830 (± 0.00)	0.857 (± 0.01)
0.50	0.779 (± 0.01)	0.825 (± 0.00)	0.853 (± 0.00)	0.858 (± 0.00)	0.826 (± 0.00)	0.853 (± 0.00)
0.60	0.769 (± 0.02)	0.824 (± 0.00)	0.847 (± 0.01)	0.848 (± 0.01)	0.784 (± 0.04)	0.848 (± 0.00)
0.70	0.752 (± 0.03)	0.816 (± 0.00)	0.837 (± 0.00)	0.835 (± 0.00)	0.765 (± 0.02)	0.837 (± 0.00)
0.80	0.742 (± 0.03)	0.813 (± 0.00)	0.828 (± 0.00)	0.817 (± 0.00)	0.323 (± 0.10)	0.836 (± 0.00)
0.90	0.605 (± 0.13)	0.628 (± 0.24)	0.812 (± 0.00)	0.770 (± 0.00)	0.280 (± 0.05)	0.823 (± 0.00)
0.99	0.557 (± 0.14)	0.260 (± 0.00)	0.800 (± 0.00)	0.689 (± 0.01)	0.418 (± 0.04)	0.818 (± 0.00)

Table 21: F1 scores for SYNTHETIC under mechanism *U-MCAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.812 (± 0.00)	0.865 (± 0.00)	0.815 (± 0.00)	0.980 (± 0.00)	0.982 (± 0.00)	0.978 (± 0.00)	0.977 (± 0.00)	0.983 (± 0.01)
0.10	0.810 (± 0.00)	0.822 (± 0.00)	0.825 (± 0.00)	0.910 (± 0.00)	0.902 (± 0.00)	0.875 (± 0.00)	0.898 (± 0.00)	0.901 (± 0.00)
0.20	0.792 (± 0.00)	0.759 (± 0.00)	0.808 (± 0.00)	0.863 (± 0.00)	0.870 (± 0.00)	0.790 (± 0.00)	0.855 (± 0.00)	0.861 (± 0.00)
0.30	0.758 (± 0.00)	0.768 (± 0.00)	0.762 (± 0.00)	0.795 (± 0.00)	0.808 (± 0.00)	0.770 (± 0.00)	0.805 (± 0.00)	0.815 (± 0.00)
0.40	0.758 (± 0.00)	0.749 (± 0.00)	0.759 (± 0.00)	0.764 (± 0.00)	0.771 (± 0.00)	0.745 (± 0.00)	0.763 (± 0.00)	0.791 (± 0.00)
0.50	0.747 (± 0.00)	0.721 (± 0.00)	0.642 (± 0.00)	0.745 (± 0.00)	0.745 (± 0.00)	0.710 (± 0.00)	0.748 (± 0.00)	0.739 (± 0.00)
0.60	0.773 (± 0.00)	0.708 (± 0.00)	0.680 (± 0.00)	0.720 (± 0.00)	0.737 (± 0.00)	0.692 (± 0.00)	0.717 (± 0.00)	0.714 (± 0.00)
0.70	0.742 (± 0.00)	0.629 (± 0.00)	0.611 (± 0.00)	0.683 (± 0.00)	0.689 (± 0.00)	0.673 (± 0.00)	0.678 (± 0.00)	0.693 (± 0.00)
0.80	0.771 (± 0.00)	0.579 (± 0.00)	0.621 (± 0.00)	0.632 (± 0.00)	0.638 (± 0.00)	0.601 (± 0.00)	0.638 (± 0.00)	0.649 (± 0.00)
0.90	0.776 (± 0.00)	0.544 (± 0.00)	0.567 (± 0.00)	0.605 (± 0.00)	0.602 (± 0.00)	0.592 (± 0.00)	0.588 (± 0.00)	0.590 (± 0.00)
0.99	0.762 (± 0.00)	0.499 (± 0.00)	0.391 (± 0.00)	0.542 (± 0.00)	0.367 (± 0.00)	0.471 (± 0.00)	0.547 (± 0.00)	0.535 (± 0.00)

Table 22: F1 scores for SYNTHETIC under mechanism *S-MCAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.812 (± 0.00)	0.865 (± 0.00)	0.815 (± 0.00)	0.980 (± 0.00)	0.982 (± 0.00)	0.978 (± 0.00)	0.977 (± 0.00)	0.983 (± 0.01)
0.10	0.756 (± 0.00)	0.748 (± 0.00)	0.723 (± 0.00)	0.903 (± 0.00)	0.912 (± 0.00)	0.903 (± 0.00)	0.900 (± 0.00)	0.898 (± 0.00)
0.20	0.769 (± 0.00)	0.733 (± 0.00)	0.727 (± 0.00)	0.883 (± 0.00)	0.883 (± 0.00)	0.872 (± 0.00)	0.870 (± 0.00)	0.875 (± 0.00)
0.30	0.742 (± 0.00)	0.737 (± 0.00)	0.700 (± 0.00)	0.830 (± 0.00)	0.842 (± 0.00)	0.841 (± 0.00)	0.831 (± 0.00)	0.833 (± 0.00)
0.40	0.716 (± 0.00)	0.712 (± 0.00)	0.683 (± 0.00)	0.810 (± 0.00)	0.798 (± 0.00)	0.752 (± 0.00)	0.793 (± 0.00)	0.799 (± 0.00)
0.50	0.700 (± 0.00)	0.711 (± 0.00)	0.704 (± 0.00)	0.785 (± 0.00)	0.788 (± 0.00)	0.705 (± 0.00)	0.780 (± 0.00)	0.779 (± 0.00)
0.60	0.658 (± 0.00)	0.674 (± 0.00)	0.695 (± 0.00)	0.747 (± 0.00)	0.761 (± 0.00)	0.726 (± 0.00)	0.738 (± 0.00)	0.756 (± 0.00)
0.70	0.618 (± 0.00)	0.675 (± 0.00)	0.652 (± 0.00)	0.687 (± 0.00)	0.703 (± 0.00)	0.665 (± 0.00)	0.700 (± 0.00)	0.727 (± 0.00)
0.80	0.584 (± 0.00)	0.649 (± 0.00)	0.616 (± 0.00)	0.653 (± 0.00)	0.667 (± 0.00)	0.645 (± 0.00)	0.638 (± 0.00)	0.676 (± 0.00)
0.90	0.527 (± 0.00)	0.588 (± 0.00)	0.589 (± 0.00)	0.597 (± 0.00)	0.597 (± 0.00)	0.578 (± 0.00)	0.591 (± 0.00)	0.582 (± 0.00)
0.99	0.337 (± 0.00)	0.455 (± 0.00)	0.338 (± 0.00)	0.515 (± 0.00)	0.425 (± 0.00)	0.403 (± 0.00)	0.513 (± 0.00)	0.477 (± 0.00)

Table 23: F1 scores for SYNTHETIC under mechanism *LD-MCAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.812 (± 0.00)	0.865 (± 0.00)	0.815 (± 0.00)	0.980 (± 0.00)	0.982 (± 0.00)	0.978 (± 0.00)	0.977 (± 0.00)	0.886 (± 0.00)
0.10	0.778 (± 0.00)	0.785 (± 0.00)	0.792 (± 0.00)	0.860 (± 0.00)	0.857 (± 0.00)	0.845 (± 0.00)	0.860 (± 0.00)	0.829 (± 0.00)
0.20	0.760 (± 0.00)	0.731 (± 0.00)	0.705 (± 0.00)	0.788 (± 0.00)	0.770 (± 0.00)	0.741 (± 0.00)	0.772 (± 0.00)	0.780 (± 0.00)
0.30	0.730 (± 0.00)	0.666 (± 0.00)	0.718 (± 0.00)	0.736 (± 0.00)	0.733 (± 0.00)	0.730 (± 0.00)	0.734 (± 0.00)	0.738 (± 0.00)
0.40	0.736 (± 0.00)	0.625 (± 0.00)	0.607 (± 0.00)	0.661 (± 0.00)	0.659 (± 0.00)	0.673 (± 0.00)	0.649 (± 0.00)	0.703 (± 0.00)
0.50	0.761 (± 0.00)	0.547 (± 0.00)	0.542 (± 0.00)	0.619 (± 0.00)	0.618 (± 0.00)	0.628 (± 0.00)	0.613 (± 0.00)	0.682 (± 0.00)
0.60	0.768 (± 0.00)	0.594 (± 0.00)	0.543 (± 0.00)	0.621 (± 0.00)	0.613 (± 0.00)	0.619 (± 0.00)	0.605 (± 0.00)	0.667 (± 0.00)
0.70	0.759 (± 0.00)	0.603 (± 0.00)	0.586 (± 0.00)	0.617 (± 0.00)	0.607 (± 0.00)	0.591 (± 0.00)	0.594 (± 0.00)	0.675 (± 0.00)
0.80	0.758 (± 0.00)	0.613 (± 0.00)	0.486 (± 0.00)	0.617 (± 0.00)	0.622 (± 0.00)	0.631 (± 0.00)	0.620 (± 0.00)	0.666 (± 0.00)
0.90	0.775 (± 0.00)	0.544 (± 0.00)	0.529 (± 0.00)	0.623 (± 0.00)	0.633 (± 0.00)	0.623 (± 0.00)	0.606 (± 0.00)	0.678 (± 0.00)
0.99	0.764 (± 0.00)	0.569 (± 0.00)	0.557 (± 0.00)	0.609 (± 0.00)	0.611 (± 0.00)	0.643 (± 0.00)	0.612 (± 0.00)	0.667 (± 0.00)

Table 24: F1 scores for SYNTHETIC under mechanism *FD-MNAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.812 (± 0.00)	0.865 (± 0.00)	0.815 (± 0.00)	0.980 (± 0.00)	0.982 (± 0.00)	0.978 (± 0.00)	0.977 (± 0.00)	0.983 (± 0.01)
0.10	0.751 (± 0.05)	0.750 (± 0.03)	0.761 (± 0.02)	0.893 (± 0.01)	0.900 (± 0.02)	0.878 (± 0.02)	0.895 (± 0.01)	0.895 (± 0.01)
0.20	0.750 (± 0.03)	0.721 (± 0.01)	0.699 (± 0.04)	0.836 (± 0.02)	0.845 (± 0.02)	0.785 (± 0.04)	0.847 (± 0.02)	0.843 (± 0.04)
0.30	0.691 (± 0.04)	0.678 (± 0.02)	0.667 (± 0.03)	0.810 (± 0.01)	0.812 (± 0.01)	0.771 (± 0.03)	0.789 (± 0.01)	0.812 (± 0.01)
0.40	0.693 (± 0.03)	0.678 (± 0.03)	0.682 (± 0.03)	0.791 (± 0.02)	0.798 (± 0.00)	0.763 (± 0.02)	0.791 (± 0.00)	0.806 (± 0.01)
0.50	0.673 (± 0.04)	0.668 (± 0.01)	0.676 (± 0.03)	0.753 (± 0.01)	0.758 (± 0.02)	0.713 (± 0.03)	0.752 (± 0.01)	0.763 (± 0.01)
0.60	0.620 (± 0.02)	0.608 (± 0.02)	0.610 (± 0.02)	0.708 (± 0.01)	0.715 (± 0.00)	0.685 (± 0.02)	0.702 (± 0.02)	0.727 (± 0.01)
0.70	0.494 (± 0.07)	0.580 (± 0.06)	0.588 (± 0.02)	0.651 (± 0.03)	0.670 (± 0.04)	0.631 (± 0.03)	0.653 (± 0.04)	0.688 (± 0.02)
0.80	0.425 (± 0.07)	0.607 (± 0.04)	0.577 (± 0.01)	0.611 (± 0.01)	0.627 (± 0.02)	0.589 (± 0.03)	0.596 (± 0.01)	0.639 (± 0.02)
0.90	0.362 (± 0.02)	0.625 (± 0.02)	0.512 (± 0.05)	0.575 (± 0.02)	0.595 (± 0.02)	0.573 (± 0.02)	0.582 (± 0.01)	0.612 (± 0.00)
0.99	0.429 (± 0.13)	0.570 (± 0.02)	0.423 (± 0.11)	0.547 (± 0.02)	0.536 (± 0.01)	0.490 (± 0.05)	0.551 (± 0.01)	0.576 (± 0.02)

Table 25: F1 scores for SYNTHETIC under mechanism *CD-MNAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.812 (± 0.00)	0.865 (± 0.00)	0.815 (± 0.00)	0.980 (± 0.00)	0.982 (± 0.00)	0.978 (± 0.00)	0.977 (± 0.00)	0.983 (± 0.01)
0.10	0.756 (± 0.04)	0.757 (± 0.02)	0.752 (± 0.02)	0.913 (± 0.02)	0.918 (± 0.02)	0.882 (± 0.02)	0.912 (± 0.01)	0.913 (± 0.02)
0.20	0.730 (± 0.05)	0.718 (± 0.02)	0.674 (± 0.05)	0.856 (± 0.03)	0.868 (± 0.03)	0.800 (± 0.04)	0.861 (± 0.04)	0.865 (± 0.03)
0.30	0.663 (± 0.05)	0.716 (± 0.02)	0.689 (± 0.03)	0.803 (± 0.02)	0.820 (± 0.02)	0.768 (± 0.03)	0.810 (± 0.03)	0.830 (± 0.03)
0.40	0.530 (± 0.16)	0.678 (± 0.01)	0.718 (± 0.03)	0.744 (± 0.01)	0.749 (± 0.00)	0.753 (± 0.01)	0.739 (± 0.03)	0.776 (± 0.01)
0.50	0.487 (± 0.12)	0.662 (± 0.03)	0.655 (± 0.04)	0.697 (± 0.03)	0.695 (± 0.03)	0.683 (± 0.04)	0.699 (± 0.04)	0.725 (± 0.01)
0.60	0.575 (± 0.06)	0.696 (± 0.03)	0.577 (± 0.02)	0.683 (± 0.03)	0.658 (± 0.03)	0.666 (± 0.02)	0.645 (± 0.03)	0.731 (± 0.03)
0.70	0.553 (± 0.03)	0.616 (± 0.03)	0.583 (± 0.02)	0.613 (± 0.02)	0.600 (± 0.04)	0.617 (± 0.04)	0.592 (± 0.05)	0.668 (± 0.01)
0.80	0.486 (± 0.06)	0.638 (± 0.03)	0.592 (± 0.03)	0.588 (± 0.02)	0.596 (± 0.03)	0.570 (± 0.02)	0.563 (± 0.03)	0.655 (± 0.02)
0.90	0.432 (± 0.08)	0.618 (± 0.05)	0.479 (± 0.10)	0.586 (± 0.04)	0.607 (± 0.03)	0.556 (± 0.03)	0.553 (± 0.01)	0.635 (± 0.04)
0.99	0.468 (± 0.03)	0.545 (± 0.06)	0.396 (± 0.08)	0.594 (± 0.01)	0.537 (± 0.01)	0.475 (± 0.06)	0.549 (± 0.03)	0.568 (± 0.01)

Table 26: F1 scores for AIR under mechanism *U-MCAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.724 (± 0.00)	0.798 (± 0.02)	0.733 (± 0.00)	0.918 (± 0.00)	0.922 (± 0.01)	0.922 (± 0.00)	0.891 (± 0.00)	0.930 (± 0.00)
0.10	0.665 (± 0.00)	0.710 (± 0.00)	0.733 (± 0.00)	0.895 (± 0.00)	0.891 (± 0.00)	0.768 (± 0.00)	0.883 (± 0.00)	0.899 (± 0.00)
0.20	0.669 (± 0.00)	0.582 (± 0.00)	0.709 (± 0.00)	0.848 (± 0.00)	0.833 (± 0.00)	0.747 (± 0.00)	0.852 (± 0.00)	0.859 (± 0.00)
0.30	0.669 (± 0.00)	0.502 (± 0.00)	0.715 (± 0.00)	0.836 (± 0.00)	0.837 (± 0.00)	0.712 (± 0.00)	0.836 (± 0.00)	0.852 (± 0.00)
0.40	0.714 (± 0.00)	0.532 (± 0.00)	0.700 (± 0.00)	0.805 (± 0.00)	0.829 (± 0.00)	0.712 (± 0.00)	0.797 (± 0.00)	0.833 (± 0.00)
0.50	0.666 (± 0.00)	0.553 (± 0.00)	0.669 (± 0.00)	0.801 (± 0.00)	0.805 (± 0.00)	0.711 (± 0.00)	0.802 (± 0.00)	0.767 (± 0.00)
0.60	0.663 (± 0.00)	0.452 (± 0.00)	0.691 (± 0.00)	0.775 (± 0.00)	0.762 (± 0.00)	0.701 (± 0.00)	0.767 (± 0.00)	0.744 (± 0.00)
0.70	0.714 (± 0.00)	0.495 (± 0.00)	0.686 (± 0.00)	0.724 (± 0.00)	0.736 (± 0.00)	0.656 (± 0.00)	0.754 (± 0.00)	0.736 (± 0.00)
0.80	0.666 (± 0.00)	0.559 (± 0.00)	0.667 (± 0.00)	0.712 (± 0.00)	0.677 (± 0.00)	0.647 (± 0.00)	0.637 (± 0.00)	0.713 (± 0.00)
0.90	0.700 (± 0.00)	0.541 (± 0.00)	0.670 (± 0.00)	0.585 (± 0.00)	0.593 (± 0.00)	0.669 (± 0.00)	0.619 (± 0.00)	0.705 (± 0.00)
0.99	0.693 (± 0.00)	0.409 (± 0.00)	0.658 (± 0.00)	0.436 (± 0.00)	0.384 (± 0.00)	0.651 (± 0.00)	0.431 (± 0.00)	0.664 (± 0.00)

Table 27: F1 scores for AIR under mechanism *S-MCAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.724 (± 0.00)	0.798 (± 0.02)	0.733 (± 0.00)	0.918 (± 0.00)	0.922 (± 0.01)	0.922 (± 0.00)	0.891 (± 0.00)	0.930 (± 0.00)
0.10	0.568 (± 0.00)	0.644 (± 0.00)	0.733 (± 0.00)	0.891 (± 0.00)	0.899 (± 0.00)	0.895 (± 0.00)	0.872 (± 0.00)	0.891 (± 0.00)
0.20	0.573 (± 0.00)	0.597 (± 0.00)	0.733 (± 0.00)	0.860 (± 0.00)	0.883 (± 0.00)	0.851 (± 0.00)	0.899 (± 0.00)	0.890 (± 0.00)
0.30	0.630 (± 0.00)	0.527 (± 0.00)	0.665 (± 0.00)	0.850 (± 0.00)	0.847 (± 0.00)	0.820 (± 0.00)	0.852 (± 0.00)	0.835 (± 0.00)
0.40	0.571 (± 0.00)	0.508 (± 0.00)	0.728 (± 0.00)	0.819 (± 0.00)	0.819 (± 0.00)	0.795 (± 0.00)	0.826 (± 0.00)	0.842 (± 0.00)
0.50	0.562 (± 0.00)	0.530 (± 0.00)	0.742 (± 0.00)	0.787 (± 0.00)	0.770 (± 0.00)	0.829 (± 0.00)	0.799 (± 0.00)	0.817 (± 0.00)
0.60	0.549 (± 0.00)	0.532 (± 0.00)	0.739 (± 0.00)	0.750 (± 0.00)	0.737 (± 0.00)	0.809 (± 0.00)	0.761 (± 0.00)	0.797 (± 0.00)
0.70	0.603 (± 0.00)	0.532 (± 0.00)	0.706 (± 0.00)	0.686 (± 0.00)	0.661 (± 0.00)	0.767 (± 0.00)	0.666 (± 0.00)	0.756 (± 0.00)
0.80	0.610 (± 0.00)	0.476 (± 0.00)	0.657 (± 0.00)	0.607 (± 0.00)	0.605 (± 0.00)	0.721 (± 0.00)	0.601 (± 0.00)	0.734 (± 0.00)
0.90	0.504 (± 0.00)	0.389 (± 0.00)	0.692 (± 0.00)	0.549 (± 0.00)	0.505 (± 0.00)	0.677 (± 0.00)	0.522 (± 0.00)	0.699 (± 0.00)
0.99	0.435 (± 0.00)	0.332 (± 0.00)	0.652 (± 0.00)	0.350 (± 0.00)	0.333 (± 0.00)	0.643 (± 0.00)	0.353 (± 0.00)	0.652 (± 0.00)

Table 28: F1 scores for AIR under mechanism *LD-MCAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.724 (± 0.00)	0.798 (± 0.02)	0.733 (± 0.00)	0.918 (± 0.00)	0.922 (± 0.01)	0.922 (± 0.00)	0.891 (± 0.00)	0.930 (± 0.00)
0.10	0.714 (± 0.00)	0.730 (± 0.00)	0.706 (± 0.00)	0.804 (± 0.00)	0.819 (± 0.00)	0.700 (± 0.00)	0.820 (± 0.00)	0.876 (± 0.00)
0.20	0.714 (± 0.00)	0.730 (± 0.00)	0.703 (± 0.00)	0.804 (± 0.00)	0.819 (± 0.00)	0.677 (± 0.00)	0.820 (± 0.00)	0.887 (± 0.00)
0.30	0.710 (± 0.00)	0.651 (± 0.00)	0.613 (± 0.00)	0.721 (± 0.00)	0.697 (± 0.00)	0.696 (± 0.00)	0.726 (± 0.00)	0.744 (± 0.00)
0.40	0.701 (± 0.00)	0.587 (± 0.00)	0.617 (± 0.00)	0.717 (± 0.00)	0.687 (± 0.00)	0.691 (± 0.00)	0.701 (± 0.00)	0.794 (± 0.00)
0.50	0.717 (± 0.00)	0.504 (± 0.00)	0.458 (± 0.00)	0.528 (± 0.00)	0.571 (± 0.00)	0.625 (± 0.00)	0.564 (± 0.00)	0.722 (± 0.00)
0.60	0.717 (± 0.00)	0.504 (± 0.00)	0.450 (± 0.00)	0.528 (± 0.00)	0.571 (± 0.00)	0.625 (± 0.00)	0.564 (± 0.00)	0.737 (± 0.00)
0.70	0.717 (± 0.00)	0.498 (± 0.00)	0.446 (± 0.00)	0.540 (± 0.00)	0.553 (± 0.00)	0.668 (± 0.00)	0.518 (± 0.00)	0.662 (± 0.00)
0.80	0.703 (± 0.00)	0.557 (± 0.00)	0.430 (± 0.00)	0.515 (± 0.00)	0.481 (± 0.00)	0.676 (± 0.00)	0.457 (± 0.00)	0.680 (± 0.00)
0.90	0.703 (± 0.00)	0.498 (± 0.00)	0.338 (± 0.00)	0.515 (± 0.00)	0.481 (± 0.00)	0.676 (± 0.00)	0.457 (± 0.00)	0.674 (± 0.00)
0.99	0.660 (± 0.00)	0.468 (± 0.00)	0.338 (± 0.00)	0.515 (± 0.00)	0.481 (± 0.00)	0.682 (± 0.00)	0.457 (± 0.00)	0.673 (± 0.00)

Table 29: F1 scores for AIR under mechanism *FD-MNAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.724 (± 0.00)	0.798 (± 0.02)	0.733 (± 0.00)	0.918 (± 0.00)	0.922 (± 0.01)	0.922 (± 0.00)	0.891 (± 0.00)	0.930 (± 0.00)
0.10	0.618 (± 0.10)	0.758 (± 0.05)	0.709 (± 0.03)	0.895 (± 0.01)	0.891 (± 0.04)	0.772 (± 0.02)	0.883 (± 0.03)	0.906 (± 0.02)
0.20	0.595 (± 0.10)	0.776 (± 0.05)	0.668 (± 0.08)	0.883 (± 0.03)	0.879 (± 0.01)	0.756 (± 0.03)	0.867 (± 0.02)	0.887 (± 0.01)
0.30	0.580 (± 0.12)	0.536 (± 0.15)	0.721 (± 0.01)	0.852 (± 0.03)	0.859 (± 0.01)	0.745 (± 0.03)	0.833 (± 0.02)	0.875 (± 0.01)
0.40	0.677 (± 0.03)	0.575 (± 0.09)	0.716 (± 0.02)	0.852 (± 0.02)	0.855 (± 0.03)	0.725 (± 0.02)	0.840 (± 0.04)	0.852 (± 0.02)
0.50	0.587 (± 0.13)	0.620 (± 0.08)	0.719 (± 0.02)	0.837 (± 0.01)	0.832 (± 0.03)	0.698 (± 0.04)	0.829 (± 0.04)	0.852 (± 0.03)
0.60	0.556 (± 0.16)	0.686 (± 0.05)	0.692 (± 0.02)	0.837 (± 0.02)	0.808 (± 0.06)	0.711 (± 0.03)	0.793 (± 0.02)	0.817 (± 0.03)
0.70	0.556 (± 0.16)	0.634 (± 0.02)	0.717 (± 0.02)	0.769 (± 0.03)	0.779 (± 0.05)	0.685 (± 0.01)	0.750 (± 0.04)	0.770 (± 0.03)
0.80	0.556 (± 0.16)	0.665 (± 0.02)	0.665 (± 0.03)	0.654 (± 0.05)	0.709 (± 0.03)	0.667 (± 0.03)	0.660 (± 0.04)	0.786 (± 0.02)
0.90	0.582 (± 0.09)	0.645 (± 0.04)	0.662 (± 0.01)	0.658 (± 0.05)	0.661 (± 0.02)	0.659 (± 0.03)	0.530 (± 0.05)	0.710 (± 0.05)
0.99	0.638 (± 0.05)	0.635 (± 0.02)	0.637 (± 0.04)	0.557 (± 0.04)	0.528 (± 0.03)	0.674 (± 0.02)	0.508 (± 0.07)	0.616 (± 0.05)

Table 30: F1 scores for AIR under mechanism *CD-MNAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.724 (± 0.00)	0.798 (± 0.02)	0.733 (± 0.00)	0.918 (± 0.00)	0.922 (± 0.01)	0.922 (± 0.00)	0.891 (± 0.00)	0.930 (± 0.00)
0.10	0.598 (± 0.11)	0.667 (± 0.02)	0.722 (± 0.02)	0.888 (± 0.05)	0.887 (± 0.04)	0.851 (± 0.01)	0.891 (± 0.03)	0.895 (± 0.04)
0.20	0.556 (± 0.16)	0.632 (± 0.19)	0.697 (± 0.02)	0.864 (± 0.02)	0.848 (± 0.06)	0.778 (± 0.04)	0.841 (± 0.01)	0.864 (± 0.01)
0.30	0.556 (± 0.16)	0.526 (± 0.13)	0.722 (± 0.03)	0.845 (± 0.01)	0.825 (± 0.04)	0.689 (± 0.02)	0.841 (± 0.03)	0.891 (± 0.04)
0.40	0.480 (± 0.16)	0.691 (± 0.14)	0.601 (± 0.12)	0.833 (± 0.02)	0.805 (± 0.03)	0.722 (± 0.02)	0.860 (± 0.03)	0.860 (± 0.03)
0.50	0.536 (± 0.16)	0.607 (± 0.09)	0.705 (± 0.02)	0.813 (± 0.02)	0.769 (± 0.04)	0.674 (± 0.01)	0.783 (± 0.04)	0.833 (± 0.03)
0.60	0.622 (± 0.06)	0.636 (± 0.04)	0.694 (± 0.01)	0.758 (± 0.05)	0.708 (± 0.07)	0.681 (± 0.01)	0.766 (± 0.06)	0.766 (± 0.06)
0.70	0.580 (± 0.10)	0.672 (± 0.07)	0.681 (± 0.01)	0.757 (± 0.03)	0.724 (± 0.04)	0.644 (± 0.02)	0.753 (± 0.05)	0.726 (± 0.05)
0.80	0.563 (± 0.12)	0.681 (± 0.05)	0.676 (± 0.01)	0.733 (± 0.02)	0.655 (± 0.02)	0.658 (± 0.02)	0.712 (± 0.01)	0.769 (± 0.03)
0.90	0.655 (± 0.03)	0.615 (± 0.04)	0.653 (± 0.01)	0.693 (± 0.04)	0.579 (± 0.04)	0.643 (± 0.04)	0.692 (± 0.06)	0.668 (± 0.02)
0.99	0.654 (± 0.03)	0.522 (± 0.04)	0.660 (± 0.05)	0.524 (± 0.07)	0.473 (± 0.05)	0.650 (± 0.06)	0.424 (± 0.06)	0.631 (± 0.07)

Table 31: F1 scores for ELECTRIC under mechanism *U-MCAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.588 (± 0.00)	0.915 (± 0.00)	0.963 (± 0.01)	0.885 (± 0.00)	0.929 (± 0.00)	0.861 (± 0.00)	0.903 (± 0.00)	0.938 (± 0.01)
0.10	0.589 (± 0.00)	0.827 (± 0.00)	0.931 (± 0.00)	0.865 (± 0.00)	0.864 (± 0.00)	0.887 (± 0.00)	0.889 (± 0.00)	0.923 (± 0.00)
0.20	0.589 (± 0.00)	0.806 (± 0.00)	0.935 (± 0.00)	0.821 (± 0.00)	0.807 (± 0.00)	0.876 (± 0.00)	0.877 (± 0.00)	0.877 (± 0.00)
0.30	0.588 (± 0.00)	0.770 (± 0.00)	0.924 (± 0.00)	0.758 (± 0.00)	0.780 (± 0.00)	0.889 (± 0.00)	0.872 (± 0.00)	0.868 (± 0.00)
0.40	0.590 (± 0.00)	0.703 (± 0.00)	0.906 (± 0.00)	0.711 (± 0.00)	0.728 (± 0.00)	0.874 (± 0.00)	0.865 (± 0.00)	0.859 (± 0.00)
0.50	0.587 (± 0.00)	0.626 (± 0.00)	0.922 (± 0.00)	0.676 (± 0.00)	0.693 (± 0.00)	0.864 (± 0.00)	0.841 (± 0.00)	0.804 (± 0.00)
0.60	0.584 (± 0.00)	0.567 (± 0.00)	0.881 (± 0.00)	0.598 (± 0.00)	0.614 (± 0.00)	0.877 (± 0.00)	0.793 (± 0.00)	0.779 (± 0.00)
0.70	0.582 (± 0.00)	0.506 (± 0.00)	0.868 (± 0.00)	0.548 (± 0.00)	0.553 (± 0.00)	0.831 (± 0.00)	0.771 (± 0.00)	0.766 (± 0.00)
0.80	0.592 (± 0.00)	0.397 (± 0.00)	0.852 (± 0.00)	0.496 (± 0.00)	0.522 (± 0.00)	0.807 (± 0.00)	0.730 (± 0.00)	0.728 (± 0.00)
0.90	0.593 (± 0.00)	0.389 (± 0.00)	0.744 (± 0.00)	0.361 (± 0.00)	0.423 (± 0.00)	0.701 (± 0.00)	0.628 (± 0.00)	0.646 (± 0.00)
0.99	0.592 (± 0.00)	0.289 (± 0.00)	0.260 (± 0.00)	0.285 (± 0.00)	0.282 (± 0.00)	0.630 (± 0.00)	0.333 (± 0.00)	0.412 (± 0.00)

Table 32: F1 scores for ELECTRIC under mechanism *S-MCAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.588 (± 0.00)	0.915 (± 0.00)	0.963 (± 0.01)	0.885 (± 0.00)	0.929 (± 0.00)	0.861 (± 0.00)	0.903 (± 0.00)	0.938 (± 0.01)
0.10	0.493 (± 0.00)	0.891 (± 0.00)	0.959 (± 0.00)	0.831 (± 0.00)	0.853 (± 0.00)	0.862 (± 0.00)	0.854 (± 0.00)	0.904 (± 0.00)
0.20	0.484 (± 0.00)	0.855 (± 0.00)	0.945 (± 0.00)	0.821 (± 0.00)	0.851 (± 0.00)	0.867 (± 0.00)	0.870 (± 0.00)	0.878 (± 0.00)
0.30	0.478 (± 0.00)	0.816 (± 0.00)	0.935 (± 0.00)	0.768 (± 0.00)	0.796 (± 0.00)	0.872 (± 0.00)	0.856 (± 0.00)	0.855 (± 0.00)
0.40	0.483 (± 0.00)	0.756 (± 0.00)	0.940 (± 0.00)	0.703 (± 0.00)	0.734 (± 0.00)	0.842 (± 0.00)	0.871 (± 0.00)	0.801 (± 0.00)
0.50	0.431 (± 0.00)	0.708 (± 0.00)	0.926 (± 0.00)	0.656 (± 0.00)	0.665 (± 0.00)	0.839 (± 0.00)	0.844 (± 0.00)	0.810 (± 0.00)
0.60	0.397 (± 0.00)	0.632 (± 0.00)	0.898 (± 0.00)	0.619 (± 0.00)	0.617 (± 0.00)	0.813 (± 0.00)	0.808 (± 0.00)	0.787 (± 0.00)
0.70	0.435 (± 0.00)	0.563 (± 0.00)	0.870 (± 0.00)	0.528 (± 0.00)	0.545 (± 0.00)	0.799 (± 0.00)	0.776 (± 0.00)	0.711 (± 0.00)
0.80	0.490 (± 0.00)	0.522 (± 0.00)	0.806 (± 0.00)	0.475 (± 0.00)	0.455 (± 0.00)	0.764 (± 0.00)	0.770 (± 0.00)	0.676 (± 0.00)
0.90	0.374 (± 0.00)	0.392 (± 0.00)	0.771 (± 0.00)	0.420 (± 0.00)	0.394 (± 0.00)	0.738 (± 0.00)	0.496 (± 0.00)	0.567 (± 0.00)
0.99	0.260 (± 0.00)	0.265 (± 0.00)	0.260 (± 0.00)	0.269 (± 0.00)	0.277 (± 0.00)	0.639 (± 0.00)	0.285 (± 0.00)	0.479 (± 0.00)

Table 33: F1 scores for ELECTRIC under mechanism *LD-MCAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.588 (± 0.00)	0.915 (± 0.00)	0.963 (± 0.01)	0.885 (± 0.00)	0.929 (± 0.00)	0.861 (± 0.00)	0.903 (± 0.00)	0.920 (± 0.00)
0.10	0.585 (± 0.00)	0.794 (± 0.00)	0.910 (± 0.00)	0.828 (± 0.00)	0.843 (± 0.00)	0.890 (± 0.00)	0.894 (± 0.00)	0.867 (± 0.00)
0.20	0.584 (± 0.00)	0.687 (± 0.00)	0.920 (± 0.00)	0.710 (± 0.00)	0.762 (± 0.00)	0.860 (± 0.00)	0.842 (± 0.00)	0.815 (± 0.00)
0.30	0.591 (± 0.00)	0.604 (± 0.00)	0.650 (± 0.00)	0.672 (± 0.00)	0.693 (± 0.00)	0.815 (± 0.00)	0.820 (± 0.00)	0.793 (± 0.00)
0.40	0.587 (± 0.00)	0.475 (± 0.00)	0.630 (± 0.00)	0.475 (± 0.00)	0.494 (± 0.00)	0.729 (± 0.00)	0.723 (± 0.00)	0.685 (± 0.00)
0.50	0.589 (± 0.00)	0.301 (± 0.00)	0.630 (± 0.00)	0.260 (± 0.00)	0.260 (± 0.00)	0.630 (± 0.00)	0.260 (± 0.00)	0.532 (± 0.00)
0.60	0.593 (± 0.00)	0.271 (± 0.00)	0.630 (± 0.00)	0.260 (± 0.00)	0.260 (± 0.00)	0.630 (± 0.00)	0.260 (± 0.00)	0.517 (± 0.00)
0.70	0.589 (± 0.00)	0.310 (± 0.00)	0.260 (± 0.00)	0.260 (± 0.00)	0.260 (± 0.00)	0.629 (± 0.00)	0.260 (± 0.00)	0.571 (± 0.00)
0.80	0.593 (± 0.00)	0.343 (± 0.00)	0.260 (± 0.00)	0.260 (± 0.00)	0.263 (± 0.00)	0.630 (± 0.00)	0.260 (± 0.00)	0.544 (± 0.00)
0.90	0.589 (± 0.00)	0.315 (± 0.00)	0.260 (± 0.00)	0.260 (± 0.00)	0.263 (± 0.00)	0.630 (± 0.00)	0.260 (± 0.00)	0.538 (± 0.00)
0.99	0.589 (± 0.00)	0.330 (± 0.00)	0.260 (± 0.00)	0.260 (± 0.00)	0.260 (± 0.00)	0.630 (± 0.00)	0.260 (± 0.00)	0.552 (± 0.00)

Table 34: F1 scores for ELECTRIC under mechanism *FD-MNAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.588 (± 0.00)	0.915 (± 0.00)	0.963 (± 0.01)	0.885 (± 0.00)	0.929 (± 0.00)	0.861 (± 0.00)	0.903 (± 0.00)	0.938 (± 0.01)
0.10	0.468 (± 0.15)	0.879 (± 0.01)	0.944 (± 0.02)	0.862 (± 0.03)	0.844 (± 0.02)	0.878 (± 0.03)	0.870 (± 0.04)	0.916 (± 0.01)
0.20	0.491 (± 0.13)	0.850 (± 0.01)	0.938 (± 0.01)	0.808 (± 0.02)	0.813 (± 0.02)	0.867 (± 0.01)	0.859 (± 0.02)	0.906 (± 0.00)
0.30	0.496 (± 0.13)	0.800 (± 0.02)	0.922 (± 0.03)	0.744 (± 0.02)	0.793 (± 0.01)	0.864 (± 0.03)	0.861 (± 0.01)	0.877 (± 0.01)
0.40	0.506 (± 0.12)	0.772 (± 0.04)	0.906 (± 0.03)	0.701 (± 0.03)	0.751 (± 0.03)	0.850 (± 0.02)	0.839 (± 0.01)	0.864 (± 0.01)
0.50	0.438 (± 0.12)	0.743 (± 0.01)	0.877 (± 0.01)	0.648 (± 0.03)	0.707 (± 0.02)	0.842 (± 0.02)	0.817 (± 0.03)	0.837 (± 0.02)
0.60	0.331 (± 0.05)	0.688 (± 0.02)	0.836 (± 0.03)	0.594 (± 0.02)	0.663 (± 0.01)	0.807 (± 0.05)	0.775 (± 0.02)	0.806 (± 0.01)
0.70	0.461 (± 0.14)	0.626 (± 0.01)	0.834 (± 0.04)	0.514 (± 0.04)	0.590 (± 0.02)	0.776 (± 0.02)	0.761 (± 0.02)	0.760 (± 0.02)
0.80	0.435 (± 0.12)	0.570 (± 0.02)	0.742 (± 0.06)	0.463 (± 0.01)	0.490 (± 0.04)	0.743 (± 0.04)	0.700 (± 0.01)	0.707 (± 0.01)
0.90	0.275 (± 0.01)	0.484 (± 0.01)	0.560 (± 0.22)	0.330 (± 0.06)	0.426 (± 0.03)	0.663 (± 0.03)	0.500 (± 0.17)	0.620 (± 0.02)
0.99	0.342 (± 0.12)	0.377 (± 0.03)	0.260 (± 0.00)	0.260 (± 0.00)	0.347 (± 0.01)	0.629 (± 0.00)	0.274 (± 0.01)	0.537 (± 0.08)

Table 35: F1 scores for ELECTRIC under mechanism *CD-MNAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.588 (± 0.00)	0.915 (± 0.00)	0.963 (± 0.01)	0.885 (± 0.00)	0.929 (± 0.00)	0.861 (± 0.00)	0.903 (± 0.00)	0.938 (± 0.01)
0.10	0.486 (± 0.12)	0.888 (± 0.01)	0.962 (± 0.01)	0.869 (± 0.01)	0.874 (± 0.01)	0.908 (± 0.02)	0.885 (± 0.03)	0.922 (± 0.00)
0.20	0.476 (± 0.15)	0.851 (± 0.02)	0.931 (± 0.02)	0.815 (± 0.01)	0.802 (± 0.01)	0.879 (± 0.01)	0.879 (± 0.01)	0.902 (± 0.03)
0.30	0.478 (± 0.16)	0.819 (± 0.04)	0.922 (± 0.00)	0.789 (± 0.03)	0.789 (± 0.01)	0.872 (± 0.01)	0.880 (± 0.01)	0.890 (± 0.00)
0.40	0.431 (± 0.11)	0.807 (± 0.02)	0.902 (± 0.01)	0.775 (± 0.01)	0.762 (± 0.01)	0.835 (± 0.02)	0.865 (± 0.02)	0.869 (± 0.02)
0.50	0.450 (± 0.09)	0.758 (± 0.02)	0.867 (± 0.03)	0.722 (± 0.02)	0.748 (± 0.01)	0.835 (± 0.03)	0.827 (± 0.03)	0.850 (± 0.02)
0.60	0.436 (± 0.10)	0.706 (± 0.01)	0.853 (± 0.05)	0.663 (± 0.02)	0.608 (± 0.01)	0.847 (± 0.03)	0.780 (± 0.02)	0.836 (± 0.02)
0.70	0.337 (± 0.03)	0.604 (± 0.03)	0.812 (± 0.03)	0.585 (± 0.03)	0.538 (± 0.03)	0.770 (± 0.09)	0.729 (± 0.01)	0.765 (± 0.01)
0.80	0.411 (± 0.09)	0.594 (± 0.02)	0.824 (± 0.08)	0.540 (± 0.01)	0.486 (± 0.01)	0.703 (± 0.04)	0.671 (± 0.01)	0.742 (± 0.02)
0.90	0.392 (± 0.11)	0.531 (± 0.02)	0.735 (± 0.07)	0.473 (± 0.03)	0.449 (± 0.02)	0.686 (± 0.06)	0.600 (± 0.04)	0.683 (± 0.02)
0.99	0.304 (± 0.02)	0.329 (± 0.04)	0.264 (± 0.01)	0.303 (± 0.02)	0.294 (± 0.01)	0.629 (± 0.00)	0.312 (± 0.04)	0.561 (± 0.02)

Table 36: F1 scores for TADPOLE under mechanism *U-MCAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.804 (± 0.00)	0.648 (± 0.01)	0.790 (± 0.00)	0.806 (± 0.00)	0.832 (± 0.02)	0.786 (± 0.00)	0.792 (± 0.00)	0.809 (± 0.00)
0.10	0.789 (± 0.00)	0.590 (± 0.00)	0.795 (± 0.00)	0.801 (± 0.00)	0.832 (± 0.00)	0.809 (± 0.00)	0.821 (± 0.00)	0.820 (± 0.00)
0.20	0.808 (± 0.00)	0.590 (± 0.00)	0.803 (± 0.00)	0.823 (± 0.00)	0.836 (± 0.00)	0.779 (± 0.00)	0.802 (± 0.00)	0.799 (± 0.00)
0.30	0.814 (± 0.00)	0.567 (± 0.00)	0.791 (± 0.00)	0.806 (± 0.00)	0.825 (± 0.00)	0.757 (± 0.00)	0.803 (± 0.00)	0.802 (± 0.00)
0.40	0.804 (± 0.00)	0.610 (± 0.00)	0.831 (± 0.00)	0.800 (± 0.00)	0.820 (± 0.00)	0.794 (± 0.00)	0.799 (± 0.00)	0.805 (± 0.00)
0.50	0.752 (± 0.00)	0.581 (± 0.00)	0.813 (± 0.00)	0.809 (± 0.00)	0.830 (± 0.00)	0.799 (± 0.00)	0.810 (± 0.00)	0.814 (± 0.00)
0.60	0.756 (± 0.00)	0.575 (± 0.00)	0.808 (± 0.00)	0.785 (± 0.00)	0.797 (± 0.00)	0.722 (± 0.00)	0.791 (± 0.00)	0.799 (± 0.00)
0.70	0.610 (± 0.00)	0.552 (± 0.00)	0.795 (± 0.00)	0.740 (± 0.00)	0.772 (± 0.00)	0.729 (± 0.00)	0.762 (± 0.00)	0.802 (± 0.00)
0.80	0.669 (± 0.00)	0.552 (± 0.00)	0.804 (± 0.00)	0.757 (± 0.00)	0.728 (± 0.00)	0.669 (± 0.00)	0.775 (± 0.00)	0.764 (± 0.00)
0.90	0.759 (± 0.00)	0.590 (± 0.00)	0.241 (± 0.00)	0.758 (± 0.00)	0.408 (± 0.00)	0.608 (± 0.00)	0.767 (± 0.00)	0.763 (± 0.00)
0.99	0.707 (± 0.00)	0.523 (± 0.00)	0.241 (± 0.00)	0.241 (± 0.00)	0.241 (± 0.00)	0.353 (± 0.00)	0.241 (± 0.00)	0.700 (± 0.00)

Table 37: F1 scores for TADPOLE under mechanism *S-MCAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.804 (± 0.00)	0.648 (± 0.01)	0.790 (± 0.00)	0.806 (± 0.00)	0.832 (± 0.02)	0.786 (± 0.00)	0.792 (± 0.00)	0.831 (± 0.04)
0.10	0.554 (± 0.00)	0.542 (± 0.00)	0.805 (± 0.00)	0.803 (± 0.00)	0.815 (± 0.00)	0.751 (± 0.00)	0.804 (± 0.00)	0.810 (± 0.00)
0.20	0.497 (± 0.00)	0.486 (± 0.00)	0.818 (± 0.00)	0.818 (± 0.00)	0.811 (± 0.00)	0.737 (± 0.00)	0.814 (± 0.00)	0.794 (± 0.00)
0.30	0.523 (± 0.00)	0.502 (± 0.00)	0.775 (± 0.00)	0.799 (± 0.00)	0.825 (± 0.00)	0.777 (± 0.00)	0.818 (± 0.00)	0.775 (± 0.00)
0.40	0.482 (± 0.00)	0.581 (± 0.00)	0.800 (± 0.00)	0.797 (± 0.00)	0.794 (± 0.00)	0.719 (± 0.00)	0.784 (± 0.00)	0.790 (± 0.00)
0.50	0.501 (± 0.00)	0.523 (± 0.00)	0.757 (± 0.00)	0.777 (± 0.00)	0.769 (± 0.00)	0.739 (± 0.00)	0.798 (± 0.00)	0.795 (± 0.00)
0.60	0.539 (± 0.00)	0.498 (± 0.00)	0.802 (± 0.00)	0.769 (± 0.00)	0.734 (± 0.00)	0.693 (± 0.00)	0.804 (± 0.00)	0.816 (± 0.00)
0.70	0.480 (± 0.00)	0.453 (± 0.00)	0.748 (± 0.00)	0.719 (± 0.00)	0.738 (± 0.00)	0.642 (± 0.00)	0.752 (± 0.00)	0.795 (± 0.00)
0.80	0.502 (± 0.00)	0.422 (± 0.00)	0.689 (± 0.00)	0.736 (± 0.00)	0.703 (± 0.00)	0.555 (± 0.00)	0.730 (± 0.00)	0.812 (± 0.00)
0.90	0.377 (± 0.00)	0.280 (± 0.00)	0.503 (± 0.00)	0.680 (± 0.00)	0.650 (± 0.00)	0.420 (± 0.00)	0.739 (± 0.00)	0.742 (± 0.00)
0.99	0.272 (± 0.00)	0.249 (± 0.00)	0.241 (± 0.00)	0.384 (± 0.00)	0.241 (± 0.00)	0.241 (± 0.00)	0.241 (± 0.00)	0.370 (± 0.00)

Table 38: F1 scores for TADPOLE under mechanism *LD-MCAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.804 (± 0.00)	0.648 (± 0.01)	0.790 (± 0.00)	0.806 (± 0.00)	0.832 (± 0.02)	0.786 (± 0.00)	0.792 (± 0.00)	0.831 (± 0.04)
0.10	0.786 (± 0.00)	0.550 (± 0.00)	0.765 (± 0.00)	0.760 (± 0.00)	0.793 (± 0.00)	0.789 (± 0.00)	0.785 (± 0.00)	0.815 (± 0.00)
0.20	0.785 (± 0.00)	0.462 (± 0.00)	0.758 (± 0.00)	0.777 (± 0.00)	0.786 (± 0.00)	0.763 (± 0.00)	0.804 (± 0.00)	0.806 (± 0.00)
0.30	0.654 (± 0.00)	0.517 (± 0.00)	0.766 (± 0.00)	0.788 (± 0.00)	0.784 (± 0.00)	0.779 (± 0.00)	0.782 (± 0.00)	0.800 (± 0.00)
0.40	0.685 (± 0.00)	0.550 (± 0.00)	0.780 (± 0.00)	0.764 (± 0.00)	0.780 (± 0.00)	0.779 (± 0.00)	0.774 (± 0.00)	0.780 (± 0.00)
0.50	0.778 (± 0.00)	0.558 (± 0.00)	0.700 (± 0.00)	0.728 (± 0.00)	0.776 (± 0.00)	0.746 (± 0.00)	0.731 (± 0.00)	0.785 (± 0.00)
0.60	0.783 (± 0.00)	0.508 (± 0.00)	0.731 (± 0.00)	0.708 (± 0.00)	0.729 (± 0.00)	0.760 (± 0.00)	0.714 (± 0.00)	0.745 (± 0.00)
0.70	0.725 (± 0.00)	0.545 (± 0.00)	0.684 (± 0.00)	0.638 (± 0.00)	0.663 (± 0.00)	0.704 (± 0.00)	0.710 (± 0.00)	0.722 (± 0.00)
0.80	0.656 (± 0.00)	0.442 (± 0.00)	0.576 (± 0.00)	0.391 (± 0.00)	0.442 (± 0.00)	0.543 (± 0.00)	0.419 (± 0.00)	0.615 (± 0.00)
0.90	0.704 (± 0.00)	0.419 (± 0.00)	0.241 (± 0.00)	0.348 (± 0.00)	0.361 (± 0.00)	0.337 (± 0.00)	0.292 (± 0.00)	0.409 (± 0.00)
0.99	0.687 (± 0.00)	0.402 (± 0.00)	0.241 (± 0.00)	0.348 (± 0.00)	0.361 (± 0.00)	0.337 (± 0.00)	0.292 (± 0.00)	0.409 (± 0.00)

Table 39: F1 scores for TADPOLE under mechanism *FD-MNAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.804 (± 0.00)	0.648 (± 0.01)	0.790 (± 0.00)	0.806 (± 0.00)	0.832 (± 0.02)	0.786 (± 0.00)	0.792 (± 0.00)	0.831 (± 0.04)
0.10	0.546 (± 0.07)	0.643 (± 0.01)	0.801 (± 0.01)	0.797 (± 0.01)	0.822 (± 0.02)	0.830 (± 0.04)	0.838 (± 0.03)	0.846 (± 0.04)
0.20	0.531 (± 0.11)	0.624 (± 0.05)	0.793 (± 0.04)	0.836 (± 0.01)	0.810 (± 0.01)	0.832 (± 0.02)	0.827 (± 0.01)	0.796 (± 0.00)
0.30	0.573 (± 0.12)	0.580 (± 0.04)	0.804 (± 0.05)	0.811 (± 0.03)	0.806 (± 0.04)	0.829 (± 0.04)	0.831 (± 0.02)	0.828 (± 0.03)
0.40	0.562 (± 0.09)	0.615 (± 0.03)	0.751 (± 0.03)	0.803 (± 0.04)	0.793 (± 0.04)	0.811 (± 0.02)	0.802 (± 0.03)	0.781 (± 0.02)
0.50	0.673 (± 0.04)	0.646 (± 0.07)	0.793 (± 0.02)	0.789 (± 0.02)	0.796 (± 0.05)	0.780 (± 0.01)	0.815 (± 0.03)	0.784 (± 0.03)
0.60	0.529 (± 0.09)	0.633 (± 0.06)	0.722 (± 0.07)	0.805 (± 0.04)	0.785 (± 0.05)	0.758 (± 0.02)	0.810 (± 0.03)	0.795 (± 0.03)
0.70	0.634 (± 0.05)	0.571 (± 0.04)	0.804 (± 0.03)	0.795 (± 0.04)	0.746 (± 0.06)	0.720 (± 0.06)	0.795 (± 0.05)	0.780 (± 0.03)
0.80	0.378 (± 0.10)	0.590 (± 0.06)	0.612 (± 0.14)	0.785 (± 0.02)	0.692 (± 0.05)	0.708 (± 0.02)	0.797 (± 0.03)	0.765 (± 0.00)
0.90	0.309 (± 0.10)	0.597 (± 0.01)	0.241 (± 0.00)	0.771 (± 0.03)	0.663 (± 0.05)	0.719 (± 0.01)	0.787 (± 0.02)	0.777 (± 0.06)
0.99	0.241 (± 0.00)	0.600 (± 0.05)	0.241 (± 0.00)	0.736 (± 0.03)	0.241 (± 0.00)	0.584 (± 0.05)	0.241 (± 0.00)	0.794 (± 0.04)

Table 40: F1 scores for TADPOLE under mechanism *CD-MNAR* and varying μ

μ	GOODIE	GSPN	FairAC	FP	GNNmi	GCNmf	PCFI	GNNmim
0.00	0.804 (± 0.00)	0.648 (± 0.01)	0.790 (± 0.00)	0.806 (± 0.00)	0.832 (± 0.02)	0.786 (± 0.00)	0.792 (± 0.00)	0.809 (± 0.00)
0.10	0.553 (± 0.06)	0.534 (± 0.09)	0.793 (± 0.05)	0.813 (± 0.03)	0.829 (± 0.04)	0.792 (± 0.03)	0.806 (± 0.03)	0.803 (± 0.01)
0.20	0.485 (± 0.06)	0.515 (± 0.04)	0.804 (± 0.03)	0.812 (± 0.03)	0.832 (± 0.03)	0.810 (± 0.02)	0.806 (± 0.02)	0.815 (± 0.02)
0.30	0.441 (± 0.02)	0.584 (± 0.06)	0.805 (± 0.03)	0.785 (± 0.02)	0.811 (± 0.03)	0.786 (± 0.02)	0.812 (± 0.02)	0.827 (± 0.03)
0.40	0.502 (± 0.07)	0.671 (± 0.03)	0.828 (± 0.01)	0.818 (± 0.03)	0.808 (± 0.02)	0.793 (± 0.02)	0.814 (± 0.03)	0.830 (± 0.01)
0.50	0.448 (± 0.02)	0.621 (± 0.04)	0.784 (± 0.02)	0.804 (± 0.04)	0.799 (± 0.04)	0.756 (± 0.03)	0.803 (± 0.05)	0.828 (± 0.04)
0.60	0.457 (± 0.01)	0.529 (± 0.03)	0.791 (± 0.01)	0.781 (± 0.03)	0.803 (± 0.03)	0.710 (± 0.07)	0.797 (± 0.03)	0.792 (± 0.03)
0.70	0.485 (± 0.07)	0.590 (± 0.09)	0.639 (± 0.29)	0.797 (± 0.05)	0.787 (± 0.04)	0.710 (± 0.05)	0.822 (± 0.02)	0.818 (± 0.01)
0.80	0.376 (± 0.10)	0.605 (± 0.04)	0.434 (± 0.27)	0.785 (± 0.05)	0.767 (± 0.09)	0.744 (± 0.01)	0.798 (± 0.02)	0.800 (± 0.02)
0.90	0.362 (± 0.09)	0.563 (± 0.03)	0.241 (± 0.00)	0.788 (± 0.01)	0.730 (± 0.08)	0.689 (± 0.05)	0.776 (± 0.06)	0.803 (± 0.05)
0.99	0.324 (± 0.12)	0.547 (± 0.08)	0.241 (± 0.00)	0.255 (± 0.02)	0.241 (± 0.00)	0.348 (± 0.05)	0.241 (± 0.00)	0.652 (± 0.04)

G COMPLETE RESULT TABLES – R2 REGIME

This appendix complements the analysis of Research Question 3 (Section 4). It reports the complete set of results for the R2 regime, where training and test data are subject to different missingness mechanisms. We include both numerical tables (F1-score mean \pm std over 5 runs) and extended visualizations across all models and datasets.

G.1 NUMERICAL RESULTS

Table 41 reports the full F1-scores for all models, datasets, and shift configurations considered in the R2 regime.

Table 41: F1 (mean \pm std over 5 runs). Setup: **R2** missingness distribution shift, where training data are subject to either *FD-MNAR* or *CD-MNAR*, while test data have either no missingness, 25% or 50% of *U-MCAR*

Task	Train mech.	μ Test	GOODIE	GSPN	FairAC	GCNmf	PCFI	FP	GNNmi	GNNmim
SYNTHETIC	<i>FD-MNAR</i>	0	0.50 (\pm 0.15)	0.68 (\pm 0.01)	0.69 (\pm 0.05)	0.81 (\pm 0.01)	0.79 (\pm 0.02)	0.80 (\pm 0.01)	0.80 (\pm 0.01)	0.82 (\pm 0.01)
	<i>FD-MNAR</i>	0.25	0.47 (\pm 0.13)	0.64 (\pm 0.03)	0.69 (\pm 0.04)	0.74 (\pm 0.03)	0.75 (\pm 0.03)	0.76 (\pm 0.03)	0.75 (\pm 0.03)	0.77 (\pm 0.03)
	<i>FD-MNAR</i>	0.50	0.47 (\pm 0.13)	0.64 (\pm 0.02)	0.65 (\pm 0.04)	0.71 (\pm 0.03)	0.73 (\pm 0.02)	0.71 (\pm 0.02)	0.71 (\pm 0.02)	0.73 (\pm 0.02)
	<i>CD-MNAR</i>	0	0.71 (\pm 0.07)	0.70 (\pm 0.03)	0.70 (\pm 0.05)	0.80 (\pm 0.04)	0.81 (\pm 0.02)	0.80 (\pm 0.02)	0.78 (\pm 0.02)	0.85 (\pm 0.04)
	<i>CD-MNAR</i>	0.25	0.66 (\pm 0.05)	0.68 (\pm 0.05)	0.68 (\pm 0.03)	0.75 (\pm 0.06)	0.78 (\pm 0.04)	0.77 (\pm 0.04)	0.77 (\pm 0.02)	0.80 (\pm 0.03)
	<i>CD-MNAR</i>	0.50	0.56 (\pm 0.10)	0.64 (\pm 0.04)	0.65 (\pm 0.01)	0.73 (\pm 0.02)	0.72 (\pm 0.03)	0.72 (\pm 0.05)	0.72 (\pm 0.01)	0.75 (\pm 0.03)
AIR	<i>FD-MNAR</i>	0	0.50 (\pm 0.14)	0.33 (\pm 0.04)	0.66 (\pm 0.07)	0.83 (\pm 0.05)	0.88 (\pm 0.01)	0.86 (\pm 0.03)	0.86 (\pm 0.03)	0.87 (\pm 0.02)
	<i>FD-MNAR</i>	0.25	0.51 (\pm 0.12)	0.42 (\pm 0.04)	0.65 (\pm 0.08)	0.68 (\pm 0.05)	0.83 (\pm 0.05)	0.81 (\pm 0.02)	0.81 (\pm 0.01)	0.85 (\pm 0.01)
	<i>FD-MNAR</i>	0.50	0.52 (\pm 0.11)	0.55 (\pm 0.03)	0.70 (\pm 0.03)	0.71 (\pm 0.03)	0.80 (\pm 0.07)	0.79 (\pm 0.06)	0.79 (\pm 0.05)	0.80 (\pm 0.05)
	<i>CD-MNAR</i>	0	0.56 (\pm 0.16)	0.35 (\pm 0.02)	0.65 (\pm 0.08)	0.60 (\pm 0.2)	0.88 (\pm 0.01)	0.71 (\pm 0.07)	0.86 (\pm 0.06)	0.85 (\pm 0.00)
	<i>CD-MNAR</i>	0.25	0.56 (\pm 0.16)	0.45 (\pm 0.5)	0.70 (\pm 0.05)	0.70 (\pm 0.05)	0.84 (\pm 0.05)	0.75 (\pm 0.05)	0.84 (\pm 0.04)	0.84 (\pm 0.06)
	<i>CD-MNAR</i>	0.50	0.62 (\pm 0.07)	0.47 (\pm 0.04)	0.68 (\pm 0.07)	0.70 (\pm 0.02)	0.80 (\pm 0.05)	0.72 (\pm 0.03)	0.76 (\pm 0.05)	0.76 (\pm 0.02)
ELECTRIC	<i>FD-MNAR</i>	0	0.45 (\pm 0.11)	0.67 (\pm 0.11)	0.92 (\pm 0.02)	0.88 (\pm 0.12)	0.69 (\pm 0.00)	0.76 (\pm 0.03)	0.80 (\pm 0.02)	0.92 (\pm 0.01)
	<i>FD-MNAR</i>	0.25	0.53 (\pm 0.10)	0.68 (\pm 0.06)	0.89 (\pm 0.00)	0.80 (\pm 0.02)	0.73 (\pm 0.03)	0.69 (\pm 0.03)	0.74 (\pm 0.02)	0.87 (\pm 0.01)
	<i>FD-MNAR</i>	0.50	0.50 (\pm 0.10)	0.68 (\pm 0.01)	0.90 (\pm 0.02)	0.83 (\pm 0.01)	0.75 (\pm 0.03)	0.62 (\pm 0.02)	0.66 (\pm 0.03)	0.82 (\pm 0.02)
	<i>CD-MNAR</i>	0	0.52 (\pm 0.10)	0.78 (\pm 0.04)	0.92 (\pm 0.02)	0.86 (\pm 0.01)	0.88 (\pm 0.01)	0.83 (\pm 0.05)	0.81 (\pm 0.01)	0.94 (\pm 0.00)
	<i>CD-MNAR</i>	0.25	0.50 (\pm 0.10)	0.78 (\pm 0.01)	0.88 (\pm 0.01)	0.86 (\pm 0.02)	0.85 (\pm 0.02)	0.74 (\pm 0.04)	0.73 (\pm 0.03)	0.85 (\pm 0.03)
	<i>CD-MNAR</i>	0.50	0.49 (\pm 0.12)	0.70 (\pm 0.02)	0.87 (\pm 0.02)	0.82 (\pm 0.03)	0.81 (\pm 0.00)	0.66 (\pm 0.01)	0.70 (\pm 0.03)	0.83 (\pm 0.02)
TADPOLE	<i>FD-MNAR</i>	0	0.52 (\pm 0.07)	0.53 (\pm 0.00)	0.75 (\pm 0.03)	0.74 (\pm 0.05)	0.79 (\pm 0.00)	0.77 (\pm 0.00)	0.76 (\pm 0.01)	0.83 (\pm 0.02)
	<i>FD-MNAR</i>	0.25	0.48 (\pm 0.03)	0.48 (\pm 0.02)	0.77 (\pm 0.01)	0.73 (\pm 0.01)	0.82 (\pm 0.02)	0.78 (\pm 0.03)	0.76 (\pm 0.03)	0.81 (\pm 0.01)
	<i>FD-MNAR</i>	0.50	0.48 (\pm 0.04)	0.53 (\pm 0.02)	0.79 (\pm 0.02)	0.71 (\pm 0.04)	0.78 (\pm 0.02)	0.74 (\pm 0.02)	0.73 (\pm 0.03)	0.82 (\pm 0.03)
	<i>CD-MNAR</i>	0	0.60 (\pm 0.02)	0.26 (\pm 0.02)	0.79 (\pm 0.05)	0.75 (\pm 0.04)	0.80 (\pm 0.04)	0.80 (\pm 0.03)	0.79 (\pm 0.05)	0.79 (\pm 0.06)
	<i>CD-MNAR</i>	0.25	0.47 (\pm 0.09)	0.52 (\pm 0.02)	0.82 (\pm 0.05)	0.78 (\pm 0.01)	0.80 (\pm 0.04)	0.80 (\pm 0.04)	0.77 (\pm 0.04)	0.75 (\pm 0.03)
	<i>CD-MNAR</i>	0.50	0.49 (\pm 0.07)	0.62 (\pm 0.05)	0.81 (\pm 0.03)	0.75 (\pm 0.00)	0.79 (\pm 0.01)	0.82 (\pm 0.02)	0.76 (\pm 0.03)	0.74 (\pm 0.02)

G.2 EXTENDED VISUALIZATIONS

In addition to Figure 3 in the main paper, Figures 6 and 7 report the full results for all models under both training mechanisms.

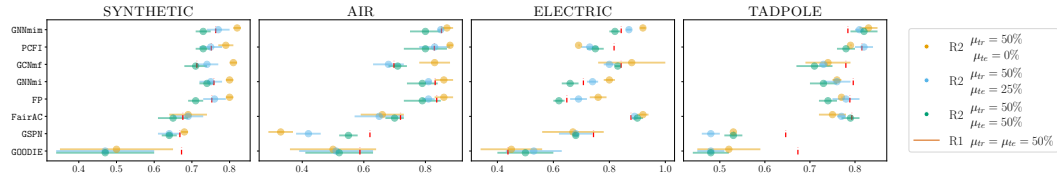


Figure 6: Full results for all models trained with *FD-MNAR* at $\mu_{tr} = 50\%$, tested on *U-MCAR* with $\mu_{te} \in \{0\%, 25\%, 50\%\}$. Each panel corresponds to one dataset; each row to one model. Reported values are mean \pm std over 5 runs.

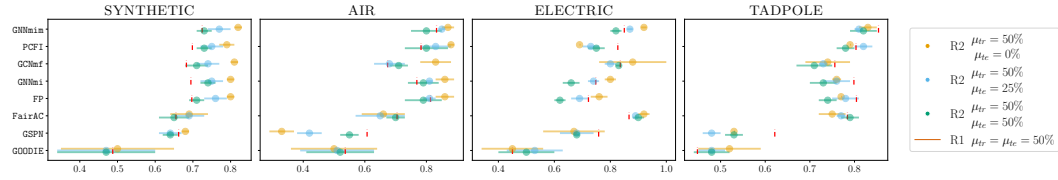


Figure 7: Full results for all models trained with *CD-MNAR* at $\mu_{tr} = 50\%$, tested on *U-MCAR* with $\mu_{te} \in \{0\%, 25\%, 50\%\}$. Same layout as Figure 6