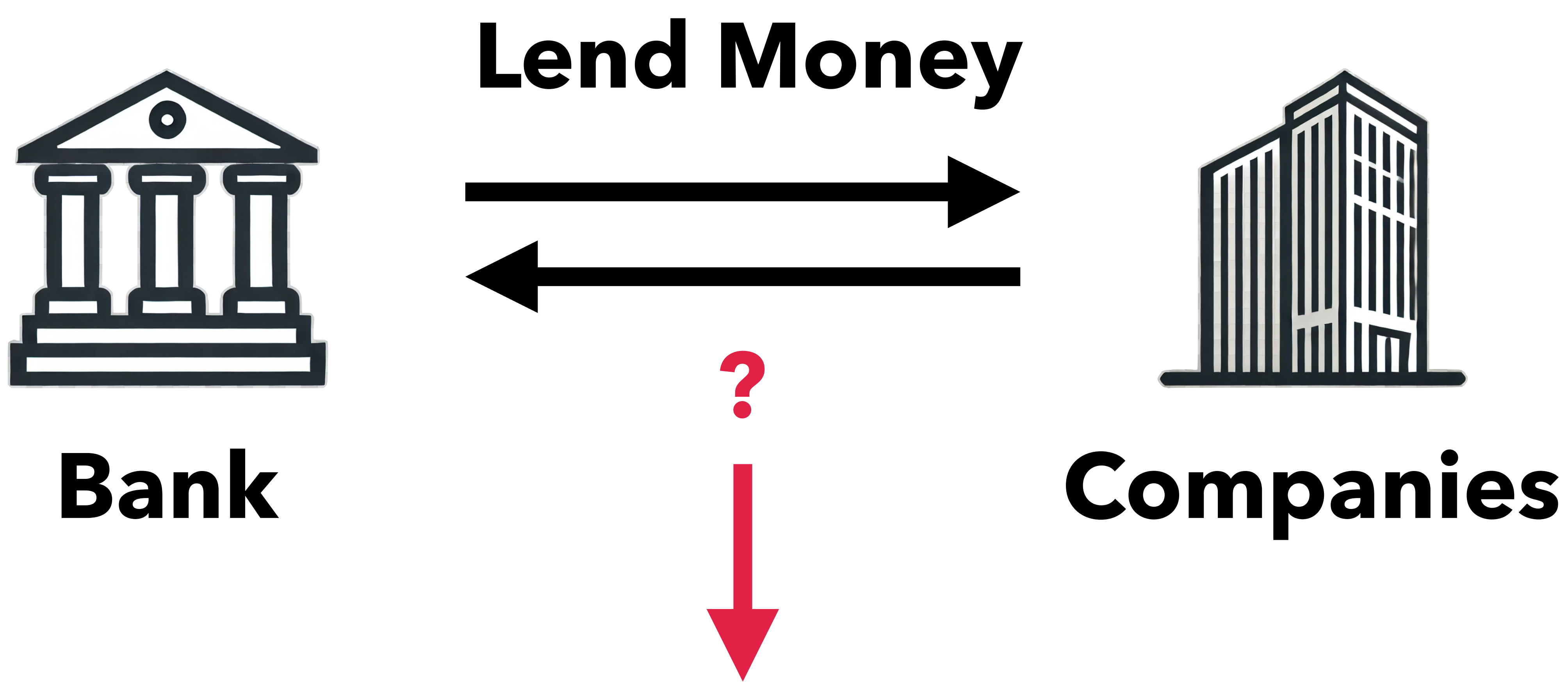
Time-aware Graph Attention Networks for Multiperiod Default Prediction

Cheng-Wei Lin, Yu-Pao Tu, Chuan-Ju Wang
Research Center for Information Technology Innovation, Academia Sinica

What is Default Prediction?



Would companies default on their obligations? When in the future?

Banks should know and manage default risks

Effective Default Prediction Requires Capturing both short-term and long-term Risk Dynamics

Company

2012-01-01

Company

2012-02-01

…

Company

2017-01-01

....

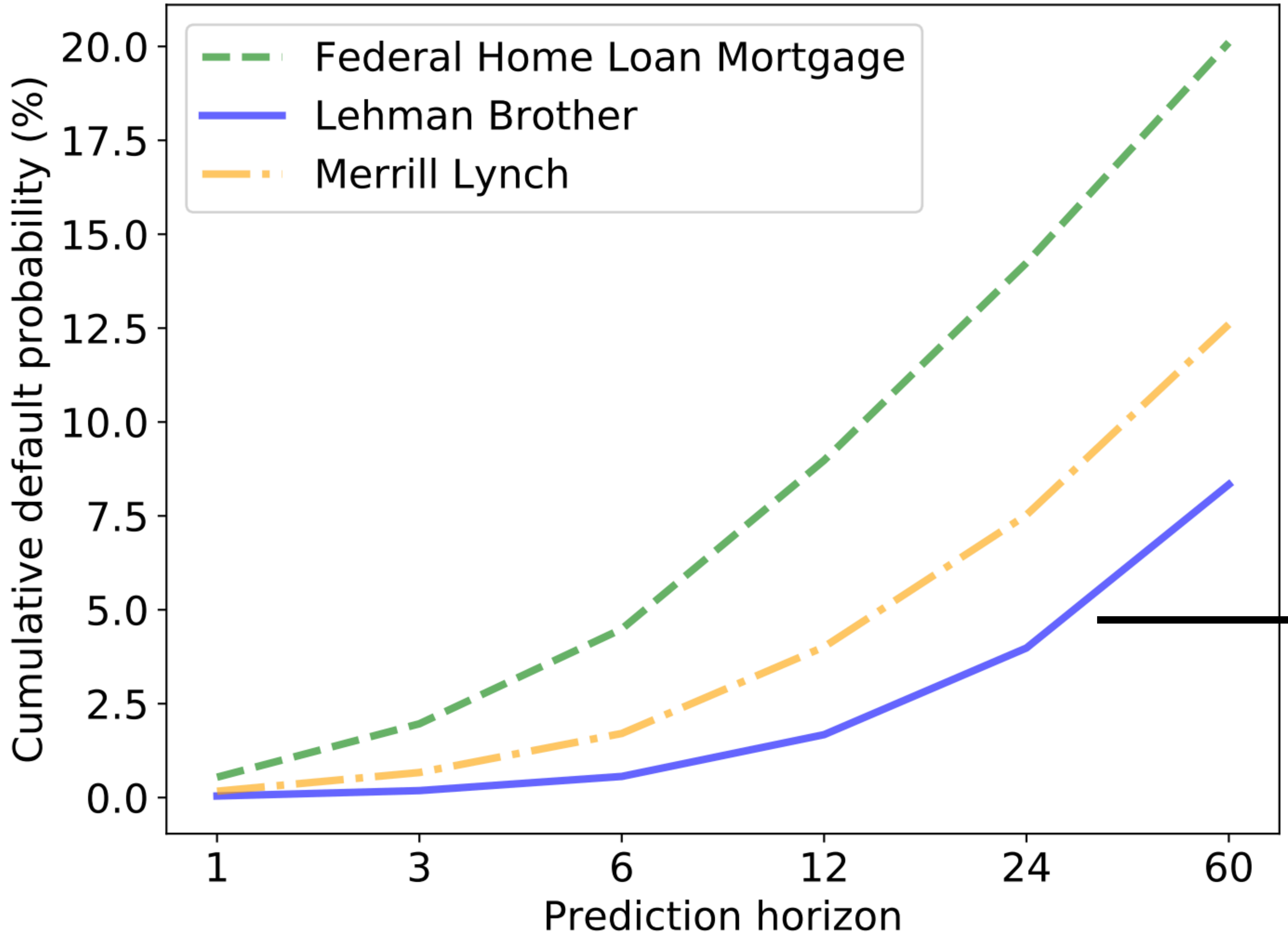
Company

Infinite Future

Predict multiple timestamps at a time: **Multiperiod**

Cumulative default probabilities

A Term Structure of Cumulative Default Probabilities



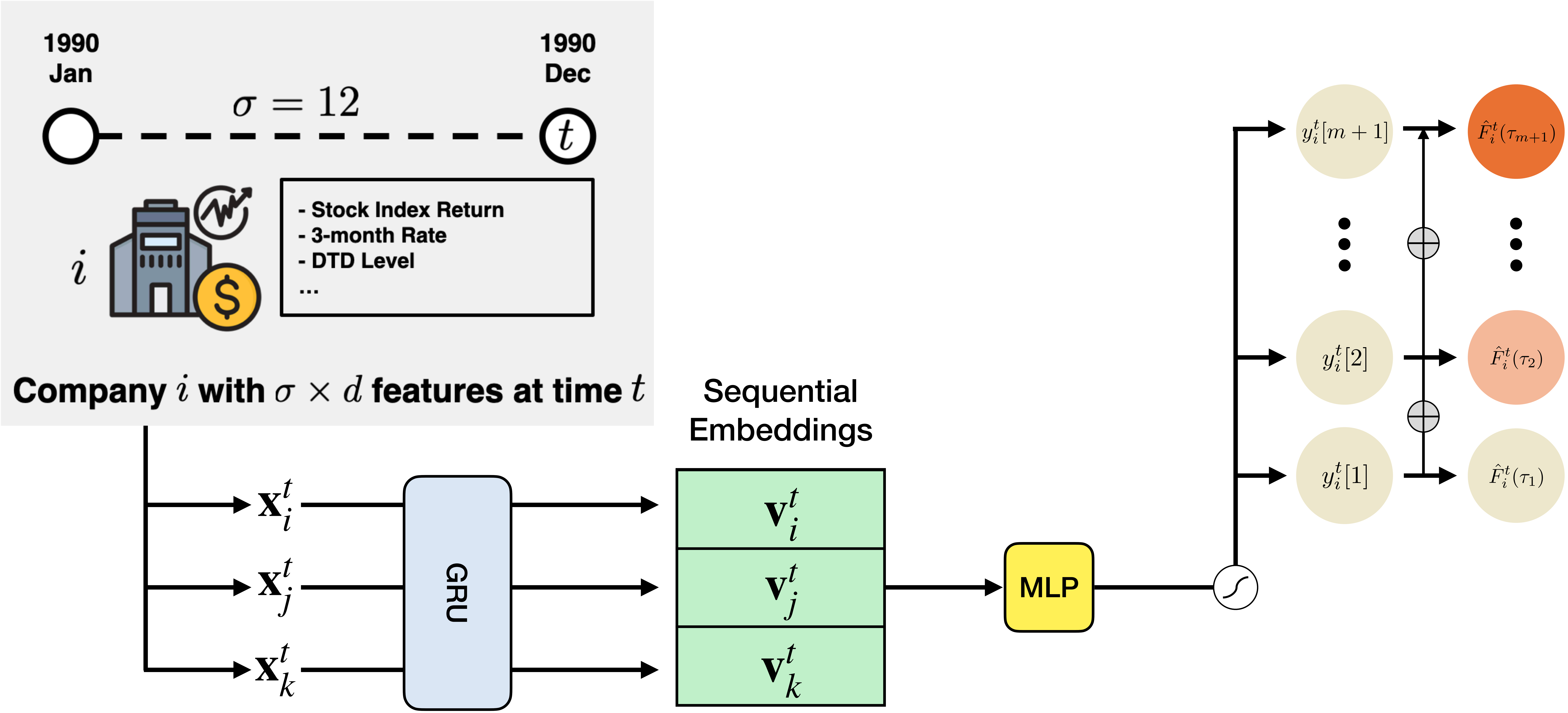
Term Structure

- Short-term: 1, 3, 6, 12 months
- Long-term: 24, 36, 48, 60 months

Monotonically increase

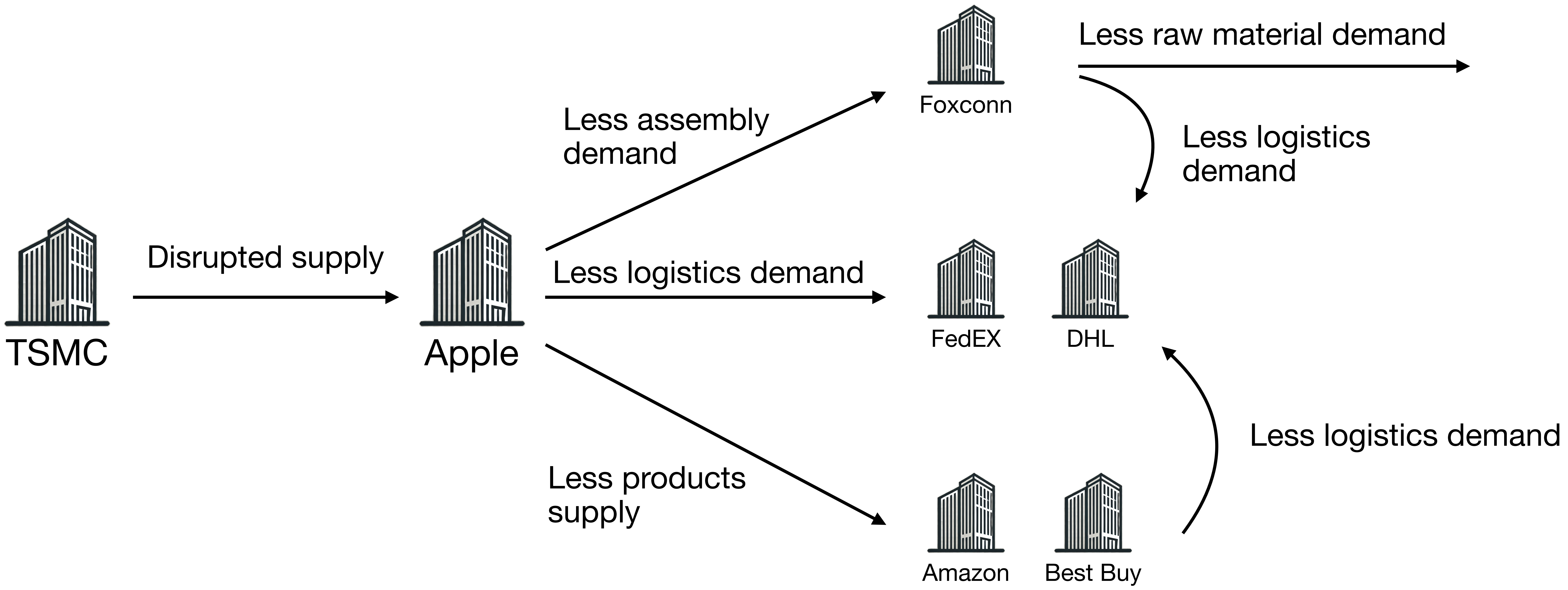
Source: Multiperiod Corporate Default Prediction Through Neural Parametric Family Learning

Previous Approaches Rely Exclusively on Individual Company Data



Source: Multiperiod Corporate Default Prediction Through Neural Parametric Family Learning

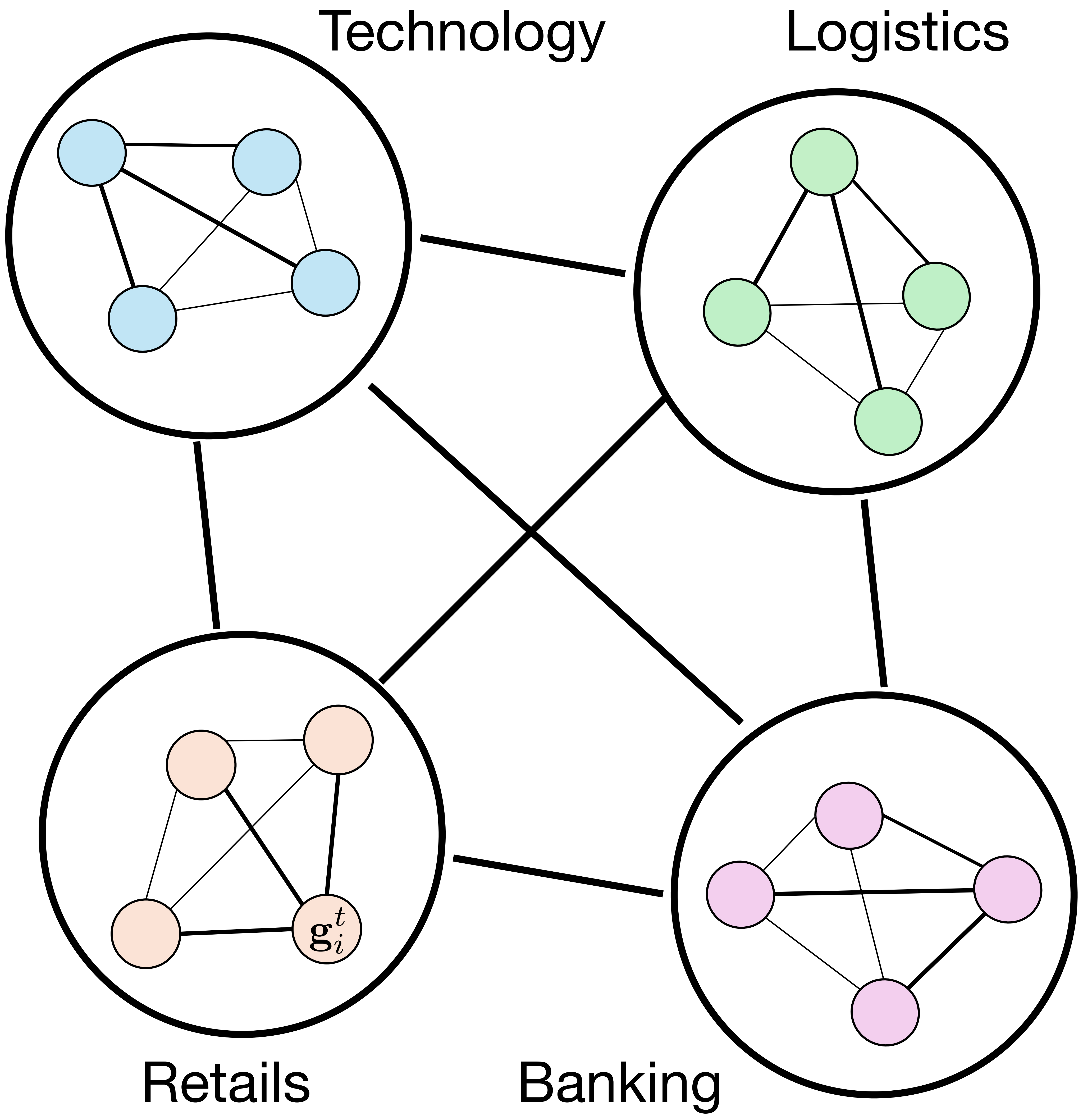
Inter-Company Relationships Should Also be Considered



However, mapping intricate inter-company relationships is a complex and challenging task

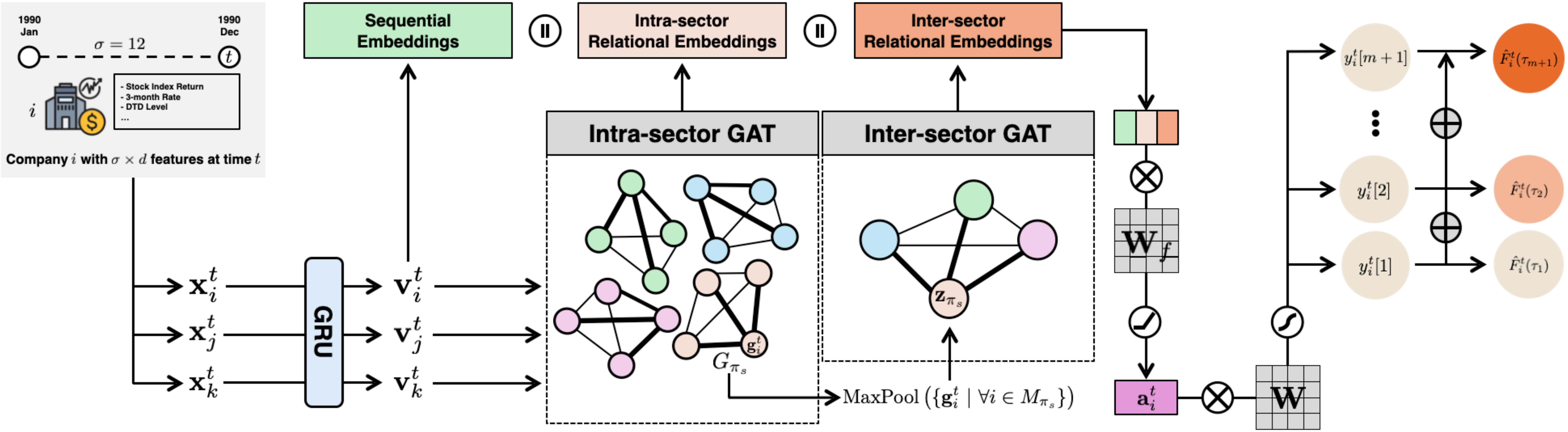
Fully-Connected Graphs Simplify Complex Inter-Company Relationships

- **How each companies is represented?**
  - Sequential embedding from its own data
- **How to decide the weight for each edge?**
  - Graph Attention Network (GATs)
- **How to represent each sector?**
  - Aggregate every companies within it by MaxPooling



TAGAT - Overall Framework

Company-level seq. Modeling	Intra-sector Modeling	Inter-sector Modeling	Emb. Integration and Default Prediction



Intra-Sector Relation Modeling

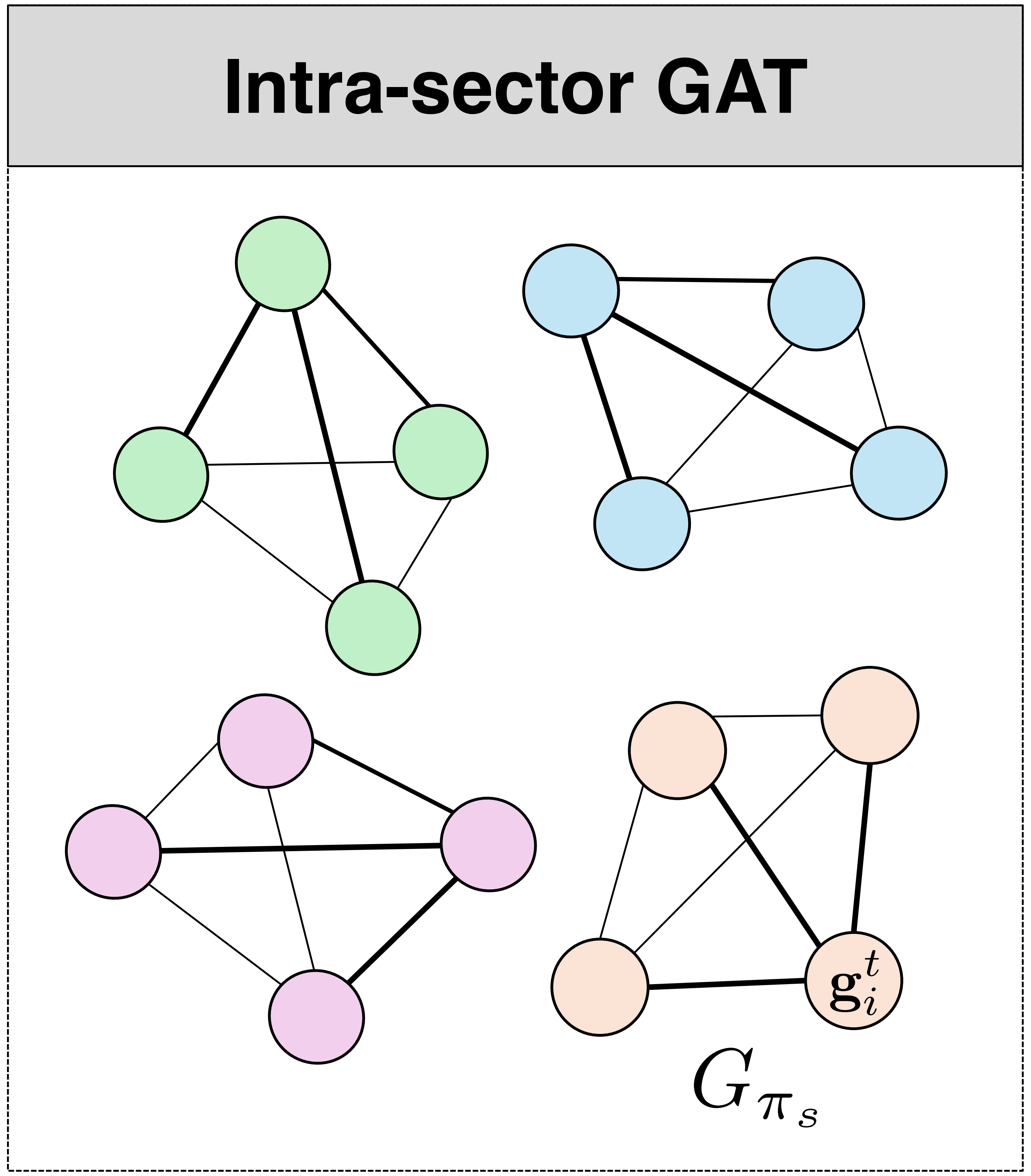
- Each company i at time t is represented as:

$$\mathbf{v}_i^t = \mathrm{GRU}(\mathbf{x}_\mathrm{i}^\mathrm{t})$$

- Each company's influence at time t from other companies within sector π_s is formulated using a GAT as:

$$\mathbf{g}_i^t = \mathrm{GAT}(\{\mathbf{g}_\mathrm{j}^\mathrm{t} \mid \forall \mathrm{j} \in \mathrm{M}_{\pi_\mathrm{s}}\}$$

$$= \sum_{j \in M_{\pi_s}} \alpha_{ij} \mathbf{W}_1 \mathbf{v}_j^t$$



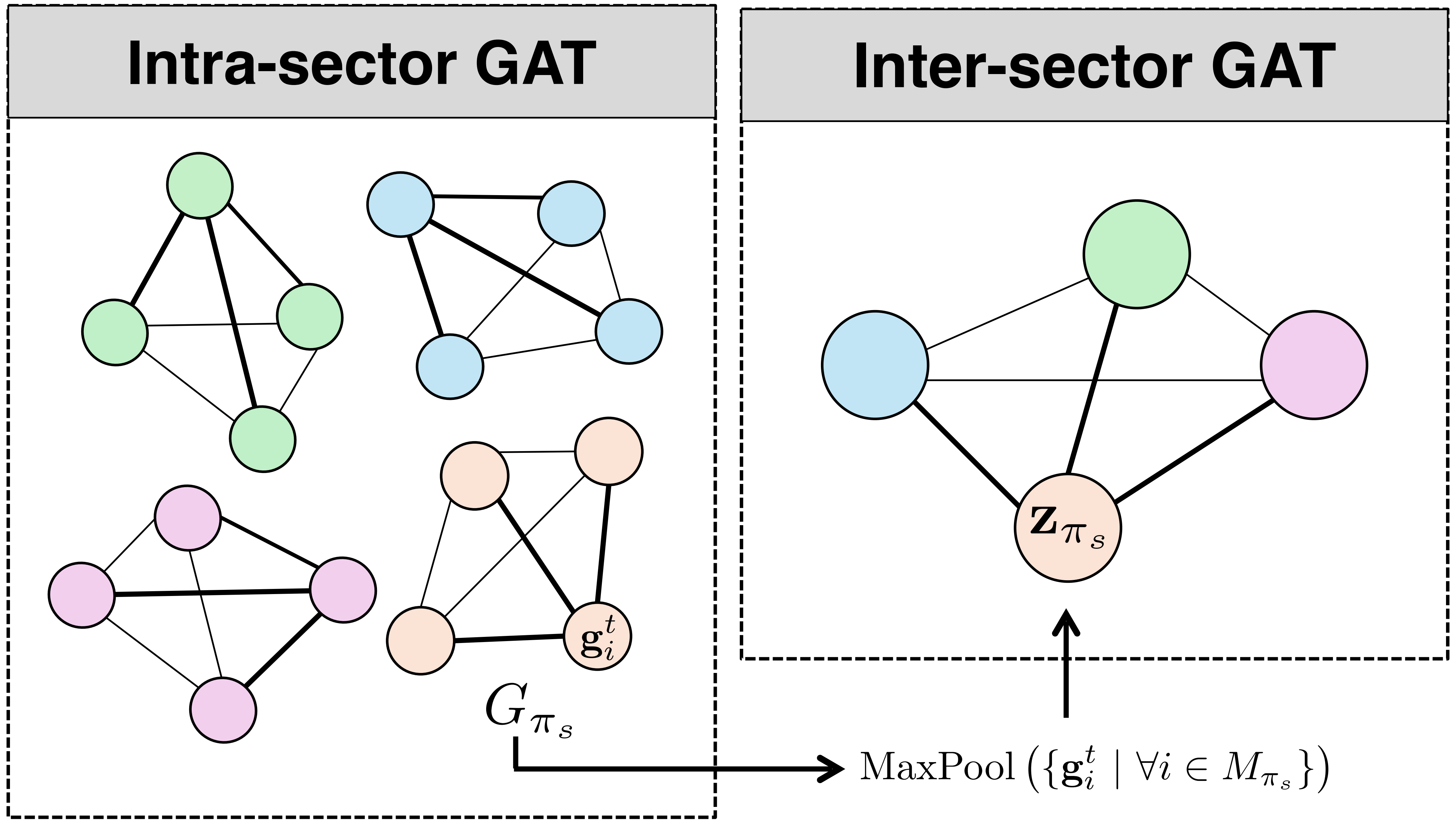
Inter-Sector Relation Modeling

- For each sector, we apply element-wise max-pooling on intra-sector relation embeddings $\mathbf{g}_i^t$ to obtain its representation

$$\mathbf{z}_{\pi_s}^t = \mathrm{MaxPool}(\{\mathbf{g}_\mathrm{i}^\mathrm{t} \mid \forall \mathrm{i} \in \mathrm{M}_{\pi_\mathrm{s}}\})$$

- Each sector's influence at time t from other companies formulated using a GAT as:

$$\tilde{\mathbf{z}}_{\pi_s}^t = \mathrm{GAT}(Z_\pi^t) = \sum_{k=1,\cdots,S} \beta_{\pi_s\pi_k} \mathbf{W}_2 \mathbf{z}_{\pi_k}^t$$



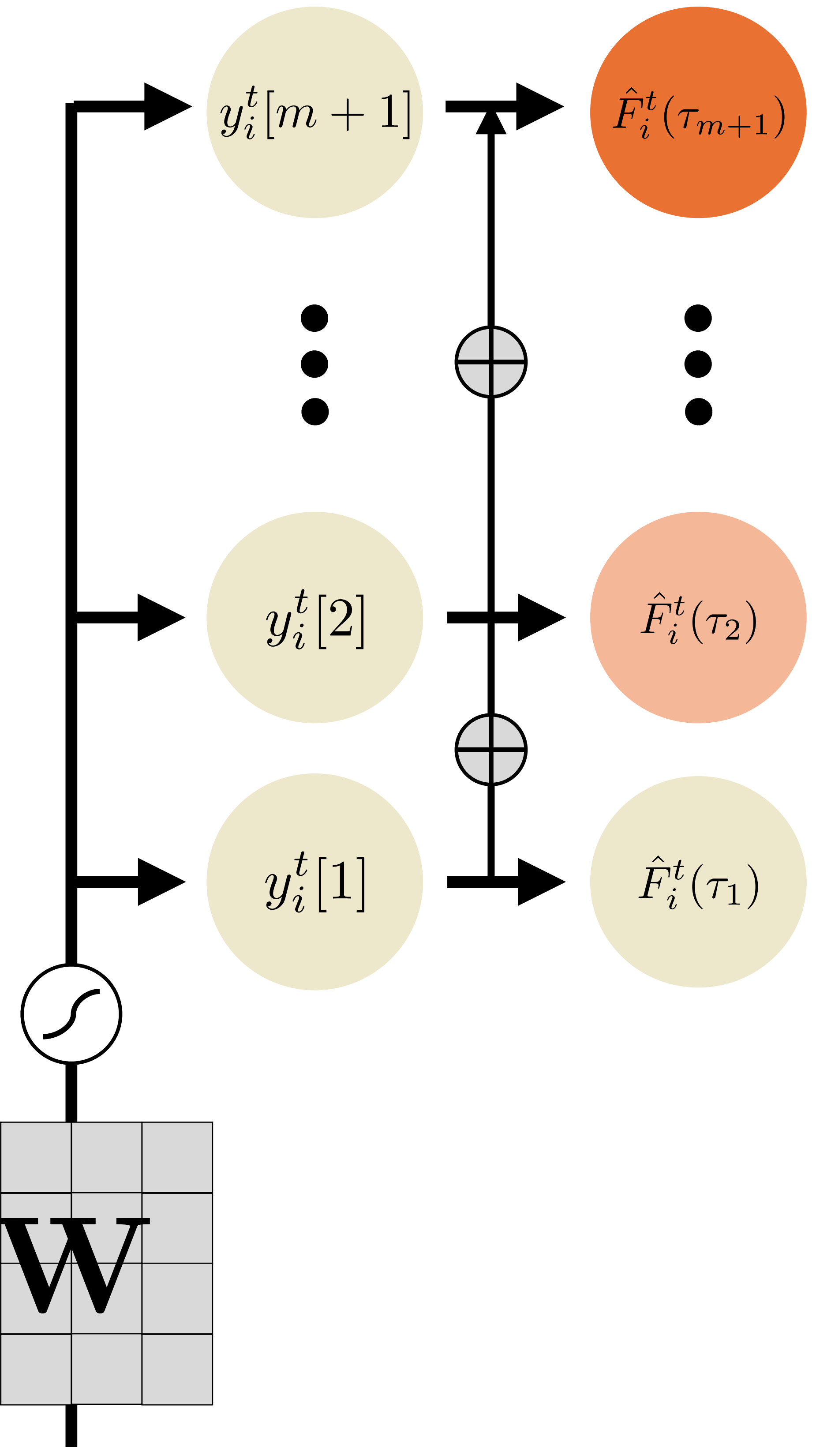
Embedding Integration and Multiperiod Default Prediction

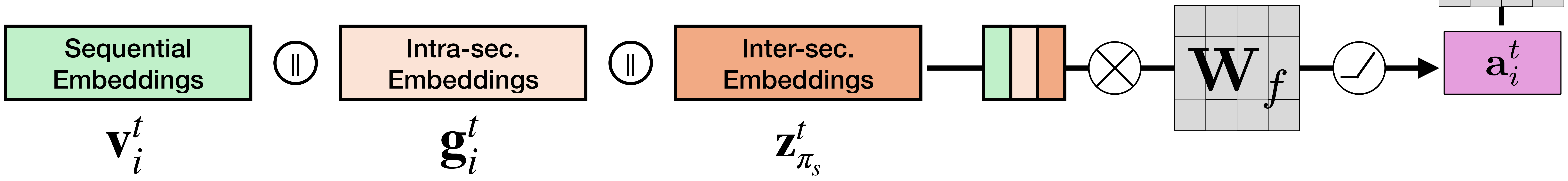
- The three embeddings are concatenated and passed through an MLP for fusion, creating a holistic company representation

$$\mathbf{a}_i^t = \mathrm{ReLU}\left([\mathbf{v}_i^t \,||\, \mathbf{g}_i^t \,||\, \tilde{\mathbf{z}}_{\pi_s}^t]\, \mathbf{W}_f + \mathbf{b}_f\right)$$

- Finally, the fused embedding passes through an MLP for multi-period default prediction

$$\mathbf{y}_i^t = \phi(\mathbf{W}\mathbf{a}_i^t + \mathbf{b}), \text{ where } \mathbf{y}_i^t \in \mathbb{R}^{(m+1)}$$





Dataset

- **Dates**: January 1990 - December 2017
- **Data**: 1.5 M monthly samples of US public companies
- **Features**: 14 covariates
- **Events**: 0 (alive), 1 (default), 2 (other exit)
- **Maximum Prediction Horizons**: 60 months
- **Data Partition**:

1.5 Million Monthly Data (Sort by Datetime)

Training (1991-2010)	Validation (2011)	Testing (2012)

Evaluation Metrics

- **Accuracy Ratio (AR)**
  - Measures a model's ability to discriminate the risk ranking among companies' default probabilities

$$\mathrm{AR} = 2 \times \mathrm{AUC} - 1$$

1	Perfect prediction
0	Random guessing
-1	Opposite prediction

- **Root Mean Square Normalized Error (RMSNE)**
  - Measures how accurately a model predicts the number of defaults over a specific prediction horizon

$$\mathrm{RMSNE} = \sqrt{\frac{1}{T}\sum_{i=1}^{T}\left(\frac{\hat{D}_i - D_i}{D_i}\right)^2}$$

$\tilde{D}_i$: the estimated default numbers

D : actual default numbers

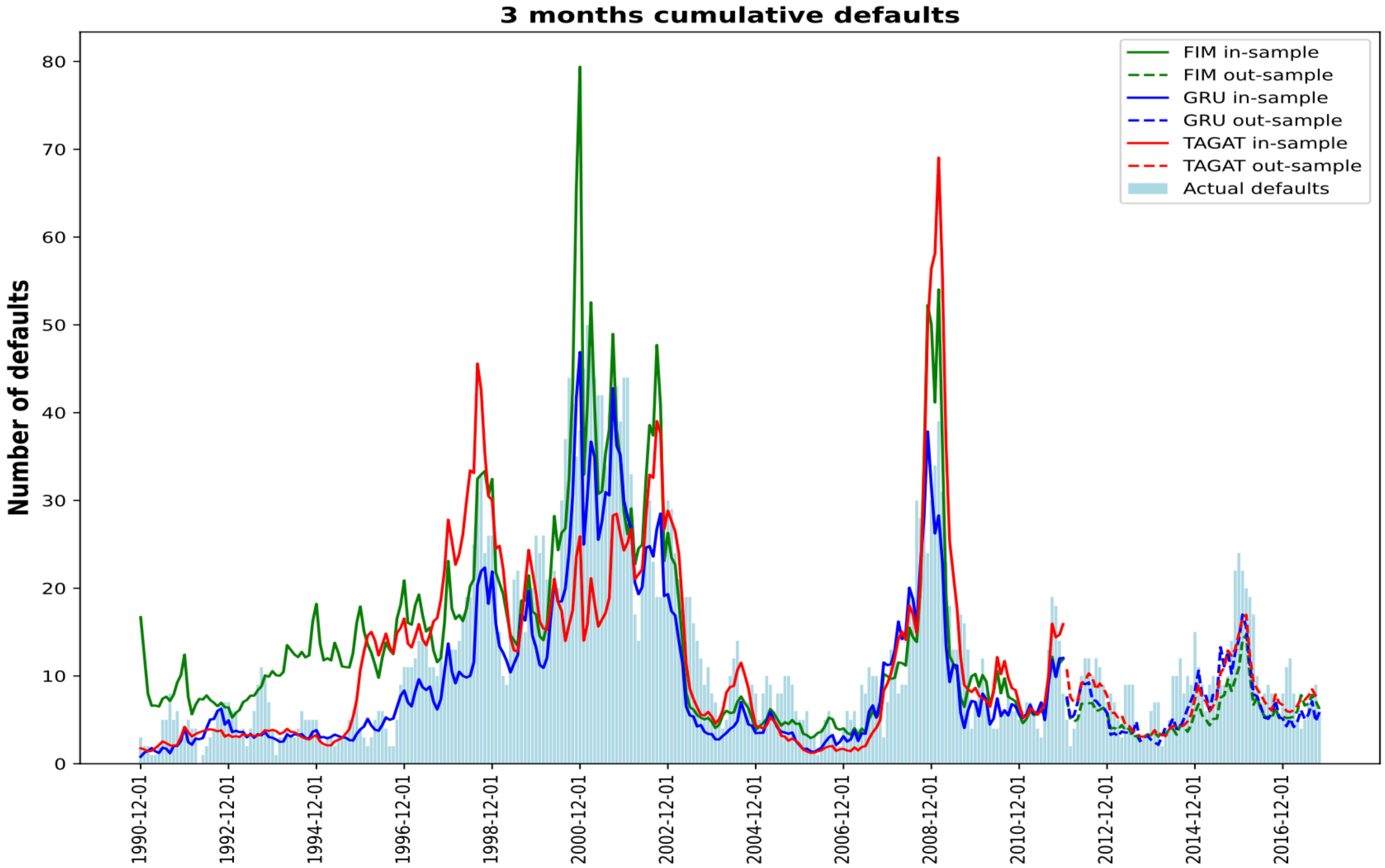
Main Result

Table 2: Main results

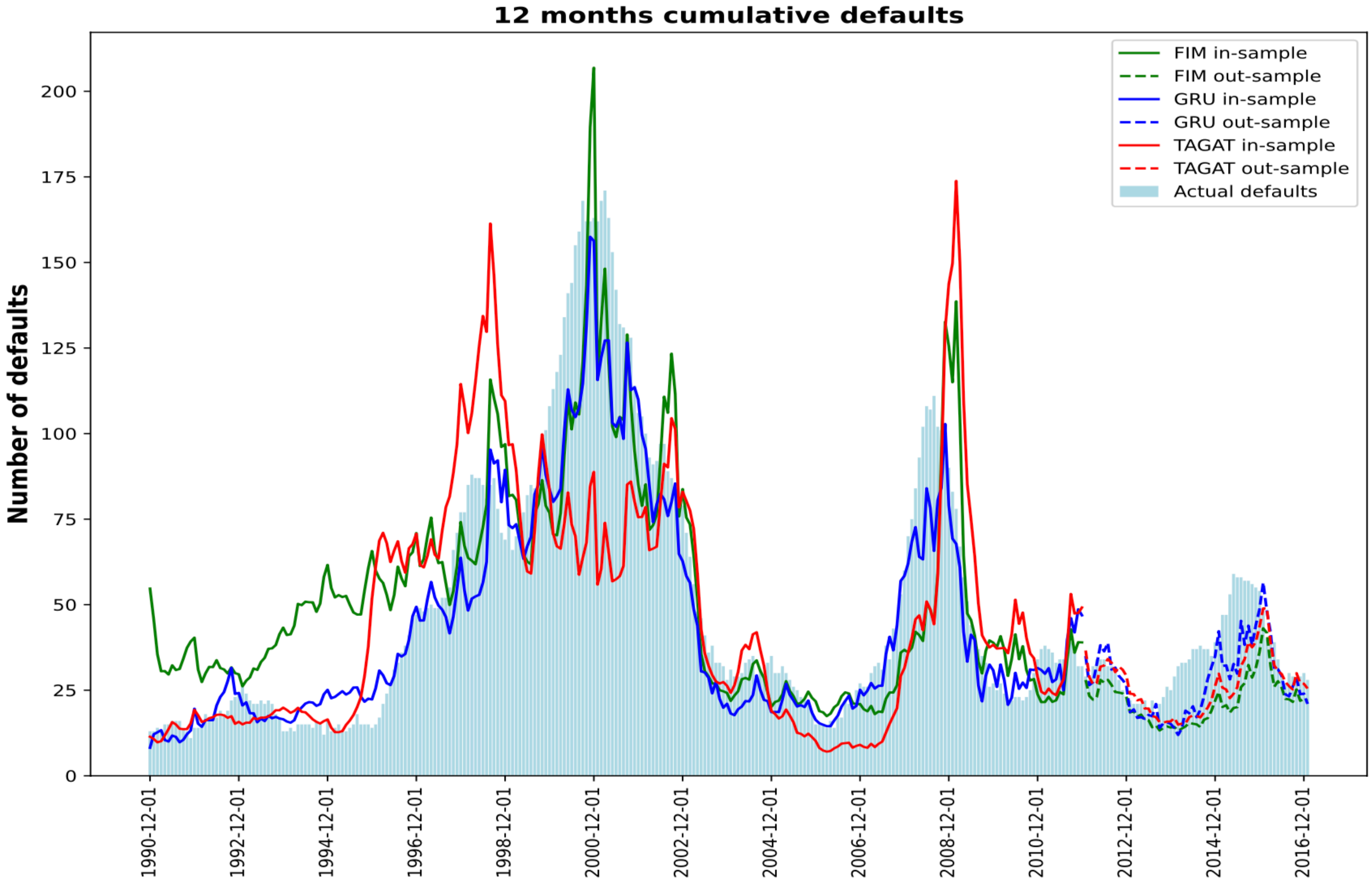
Horizons (months)	1	3	6	12	24	36	48	60
Panel A: Accuracy ratio (AR) (%)								
FIM	**95.88**	**95.01**	93.16	89.11	81.38	75.21	74.11	72.46
GRU	95.00	93.91	92.72	88.24	78.95	71.32	66.16	63.76
TAGAT	95.78	95.00	**94.30**	**91.12**	**84.11**	**78.31**	**76.07**	**73.87**
Panel B: Root mean square normalized error (RMSNE)								
FIM	**0.5408**	0.4617	0.3730	0.3898	0.4222	0.4176	0.3497	0.2245
GRU	0.5722	**0.4259**	**0.2856**	**0.2843**	0.4222	0.4827	0.3349	0.1568
TAGAT	0.9071	0.5134	0.3293	0.3190	**0.3805**	**0.3625**	**0.1802**	**0.1114**

- In terms of AR, TAGAT surpass both baselines across all prediction horizons, except for 1 and 3 months AR
- In terms of RMSNE, TAGAT exhibits notable enhancements in long-term prediction horizons, highlighting its strong capability of capturing more complex and long-term signals

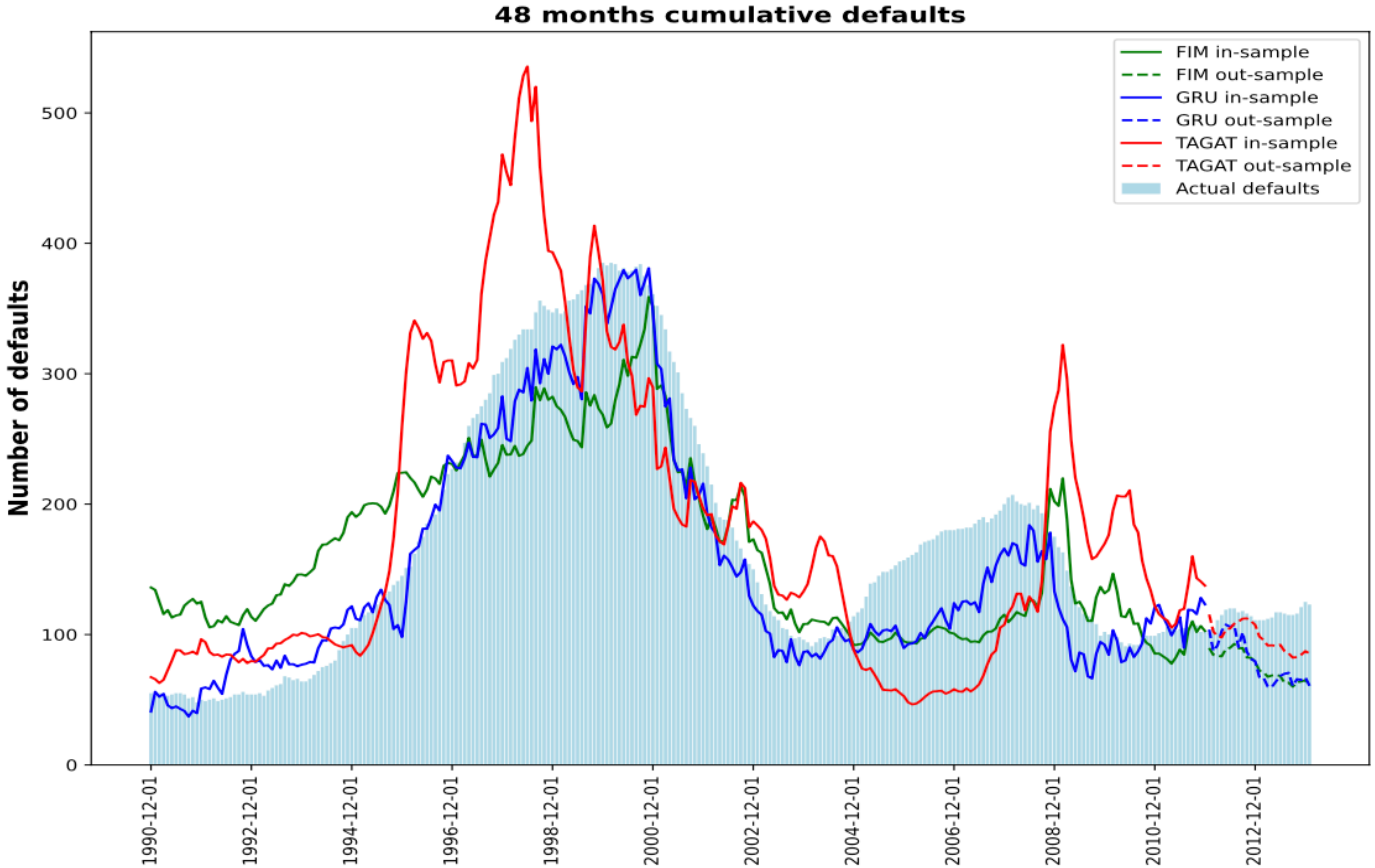
Default Distribution - 3 months



Default Distribution - 12 months



Default Distribution - 48 months



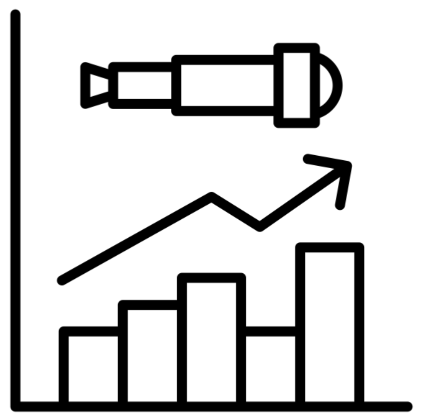
Ablation Studies

Table 3: Ablation studies on the GAT module

Horizons (months)	1	3	6	12	24	36	48	60
Panel A: Accuracy ratio (AR) (%)								
TAGAT (full model)	95.78	95.00	94.30	91.12	84.11	78.31	76.07	**73.87**
TAGAT (w/o GATs)	**96.11**	**95.25**	**94.30**	**91.20**	**84.85**	**79.54**	**76.67**	73.23
Panel B: Root mean square normalized error (RMSNE)								
TAGAT (full model)	**0.9071**	**0.5134**	**0.3293**	**0.3190**	**0.3805**	**0.3625**	**0.1802**	0.1114
TAGAT (w/o GATs)	1.1616	0.9160	0.4437	0.3722	0.3962	0.3772	0.2142	**0.0698**

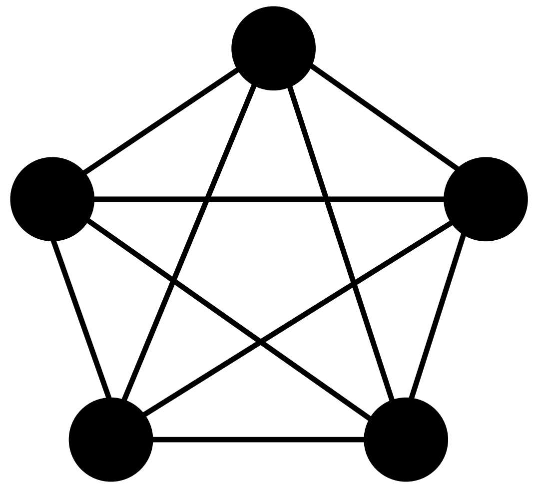
- The inclusion of GAT components marginally affects performance in terms of AR but significantly influences RMSNE
- The ability to predict more accurate numbers of defaults is more relevant for financial institutions to gain a whole picture of its current financial risk

Conclusion



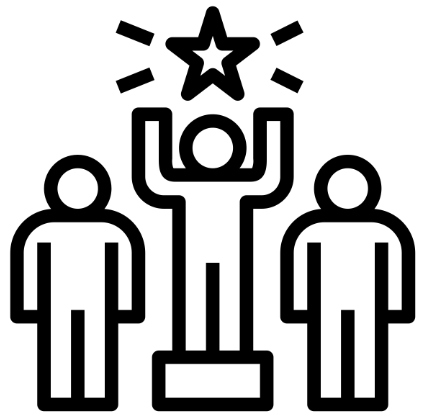
Time-aware Representation

Effectively capture the sequential characteristics of individual companies



Intra-/Inter-sector Relation Modeling

Effectively capture the company's relations with other companies within the same sector, and the impacts arising from broader sector-level dynamics



Better Long-term Default Prediction

Our experiments demonstrate that our TAGAT model excels at making accurate predictions for more challenging long-term horizons