中央研究院
ACADEMIA SINICA

NUS National University of Singapore | Asian Institute of Digital Finance

Multiperiod Corporate Default Prediction Through Neural Parametric Family Learning

Wei-Lun Luo, Yu-Ming Lu, Jheng-Hong Yang, Jin-Chuan Duan, Chuan-Ju Wang

2022.4.28

Outline

- Introduction of Credit risk
  - Default risk
- Related works
  - Niche
- Methodology
- Results

Credit risk

Lend money to obligors

Meet its obligations
With the agreed terms

Lender

Banks

Obligor

People
Companies

Credit risk

Lend money to obligors

Meet i ? ations
With the d terms

Lender

Banks

Obligor

People
Companies

Credit risk

Lend money to obligors

Meet i ? ations
With th ? d terms

Credit risk

Lender

Banks

Obligor

People
Companies

Credit risk

Lend money to obligors

Meet i ? ations
With th ? d terms

Lender

Banks

Credit risk

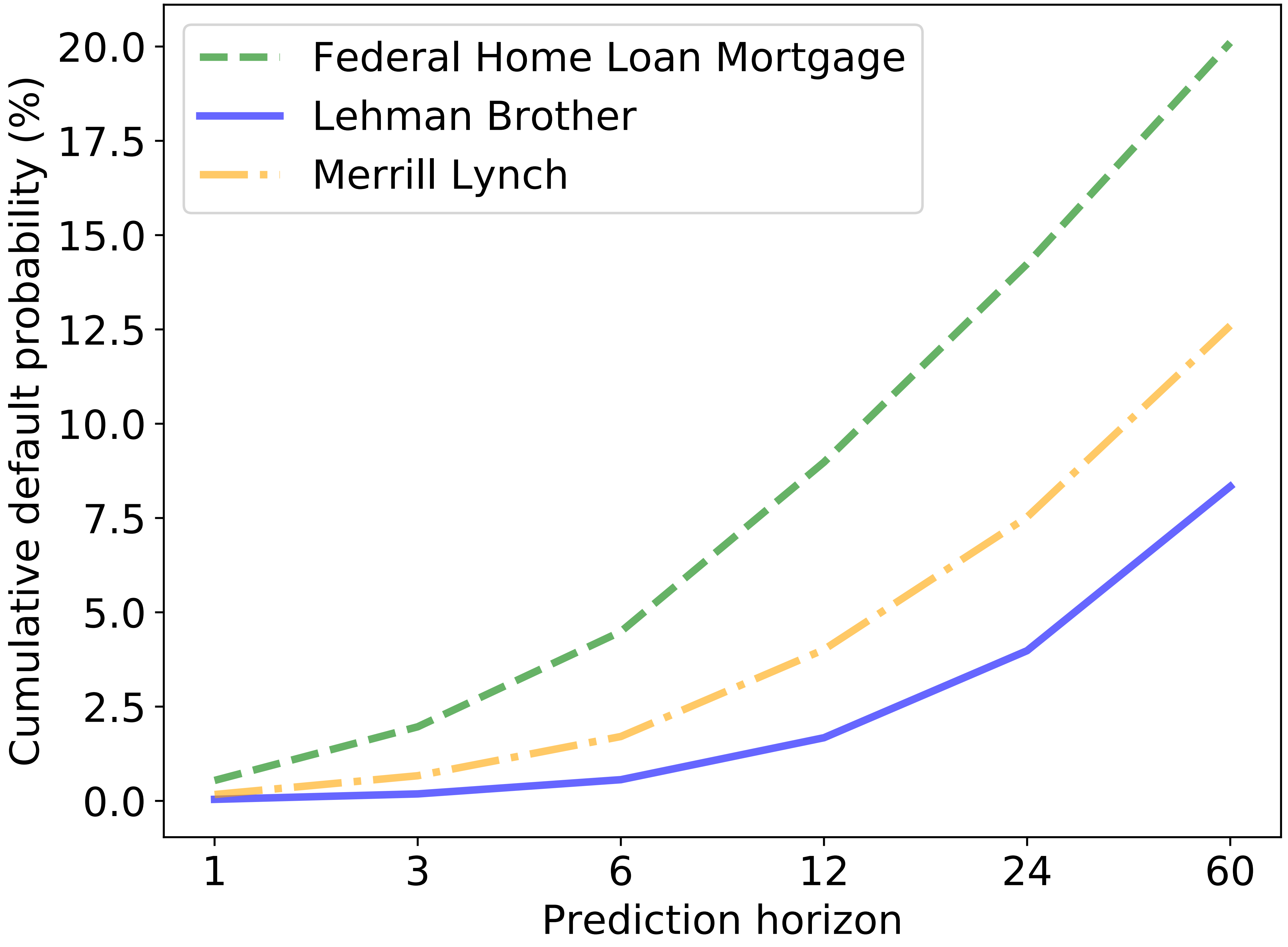
Default risk

Obligor

People
Companies

Default risk

A term structure of cumulative default probabilities (CDP)



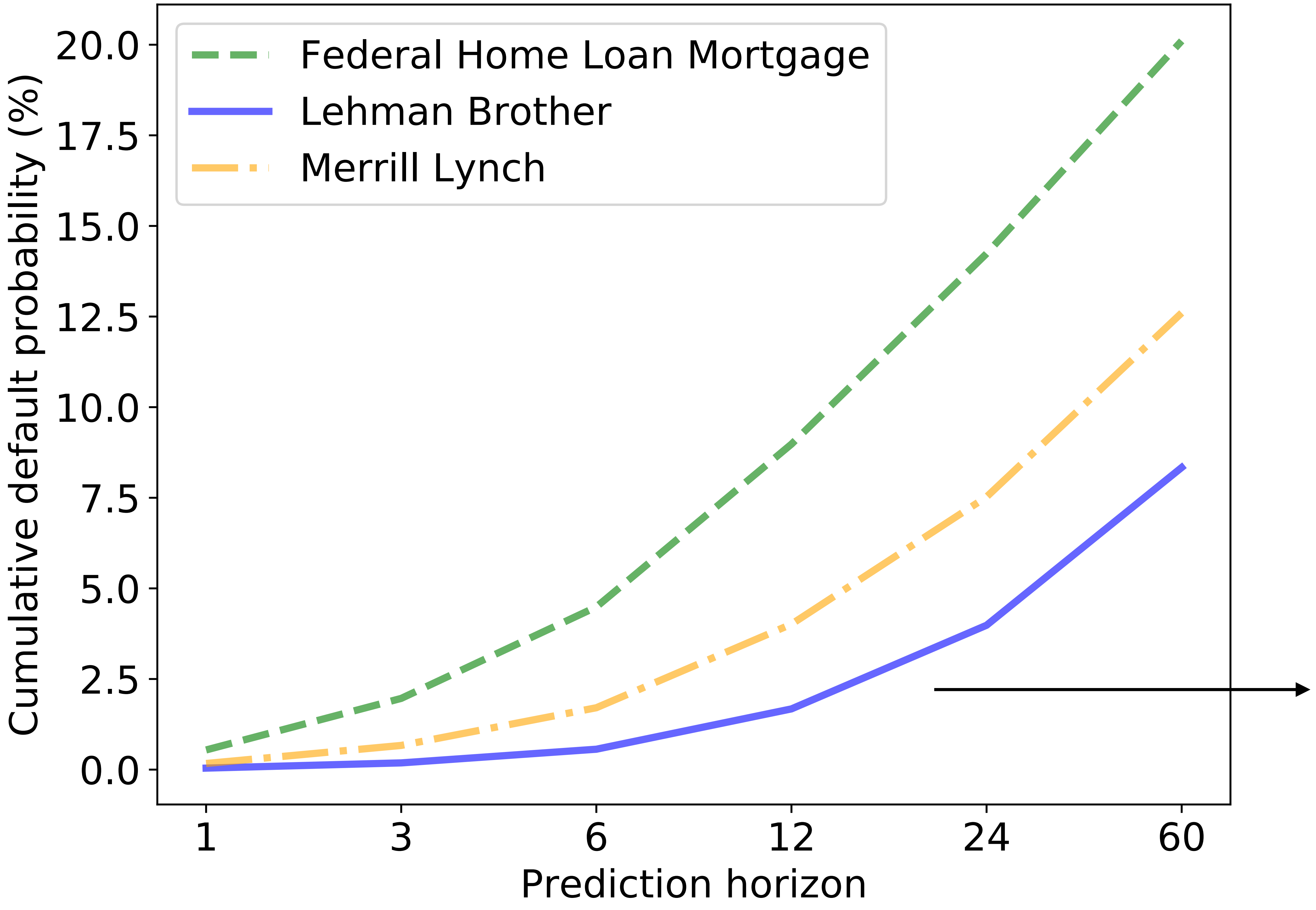
Debts structure

- Short-term
- Long-term

Dissimilar with each other

Default risk

A term structure of cumulative default probabilities (CDP)



Debts structure

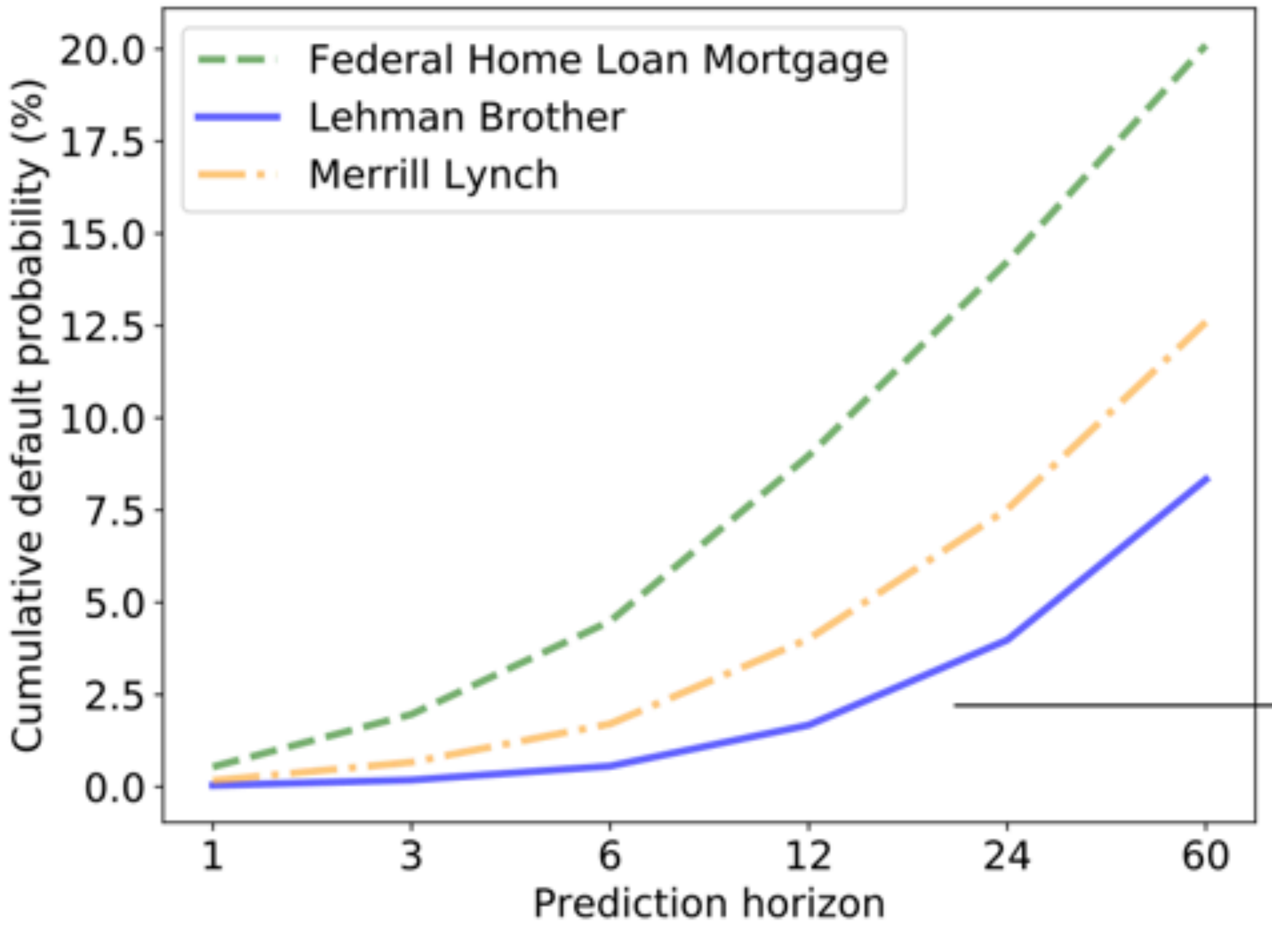
- Short-term
- Long-term

Dissimilar with each other

Monotonically Increasing

Default risk

A term structure of cumulative default probabilities (CDP)



Debts structure

- Short-term
- Long-term

Dissimilar with each other

Monotonically Increasing

Multiperiod default prediction

Multiperiod corporate default prediction

Related works

Default risk analysis

Machine learning

Risk classification (e.g. 3-months, 6 months)

Risk rankings

Hyeongjun Kim, Hoon Cho, and Doojin Ryu. Corpo- rate default predictions using machine learning: Liter- ature review. Sustainability, 12(16), 2020
Zan Huang, Hsinchun Chen, Chia-Jung Hsu, Wun- Hwa Chen, and Soushan Wu. Credit Rating Analysis with Support Vector Machines and Neural Networks: A Market Comparative Study. Decision Support Systems, 37(4):543–558, 2004.

Related works

Default risk analysis

Machine learning

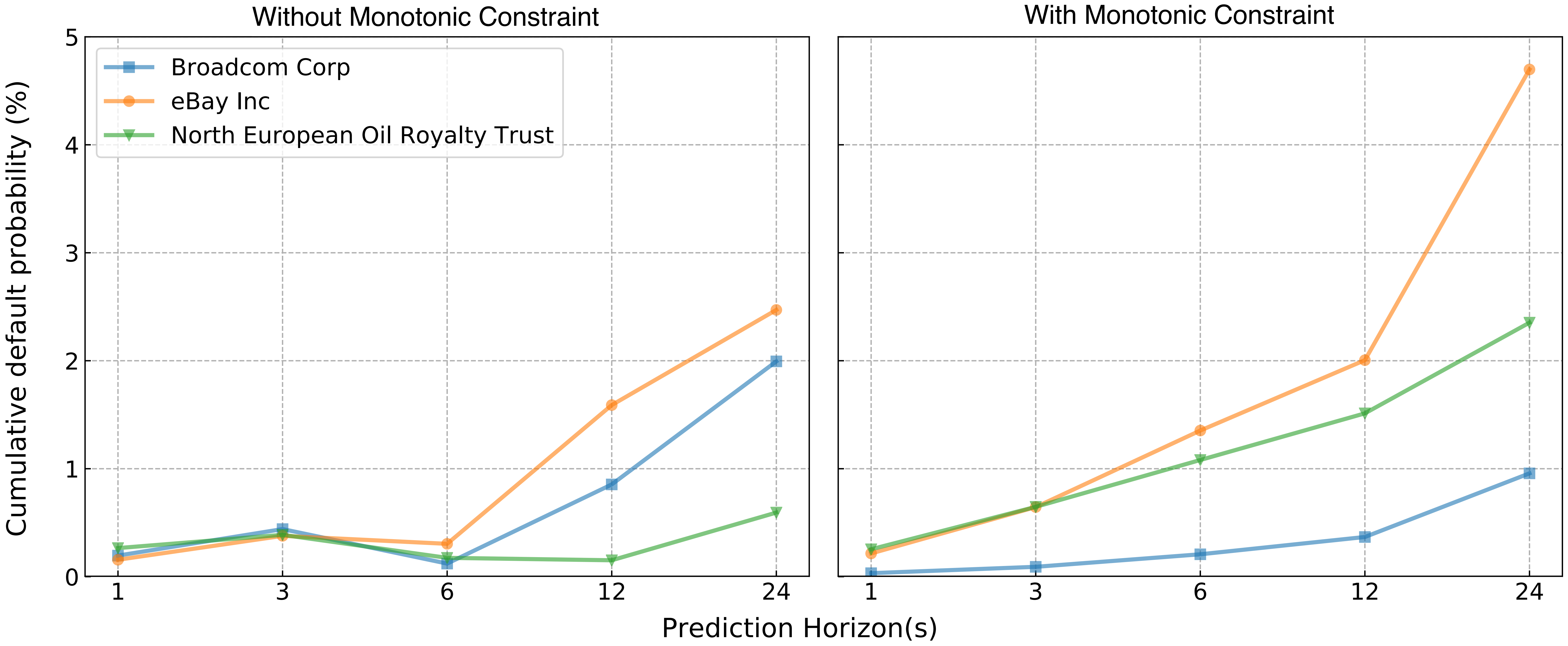
Risk classification (e.g. 3-months, 6 months)

Risk rankings

A consistent term structure

Real world applications

E.g. multi-period



Hyeongjun Kim, Hoon Cho, and Doojin Ryu. Corpo- rate default predictions using machine learning: Liter- ature review. Sustainability, 12(16), 2020

Zan Huang, Hsinchun Chen, Chia-Jung Hsu, Wun- Hwa Chen, and Soushan Wu. Credit Rating Analysis with Support Vector Machines and Neural Networks: A Market Comparative Study. Decision Support Systems, 37(4):543–558, 2004.

Related works

Default risk analysis

Machine learning

Risk classification (e.g. 3-months, 6 months)

Risk rankings

A consistent term structure

Real world applications

E.g. multi-period

Statistical

A term structure of CDP

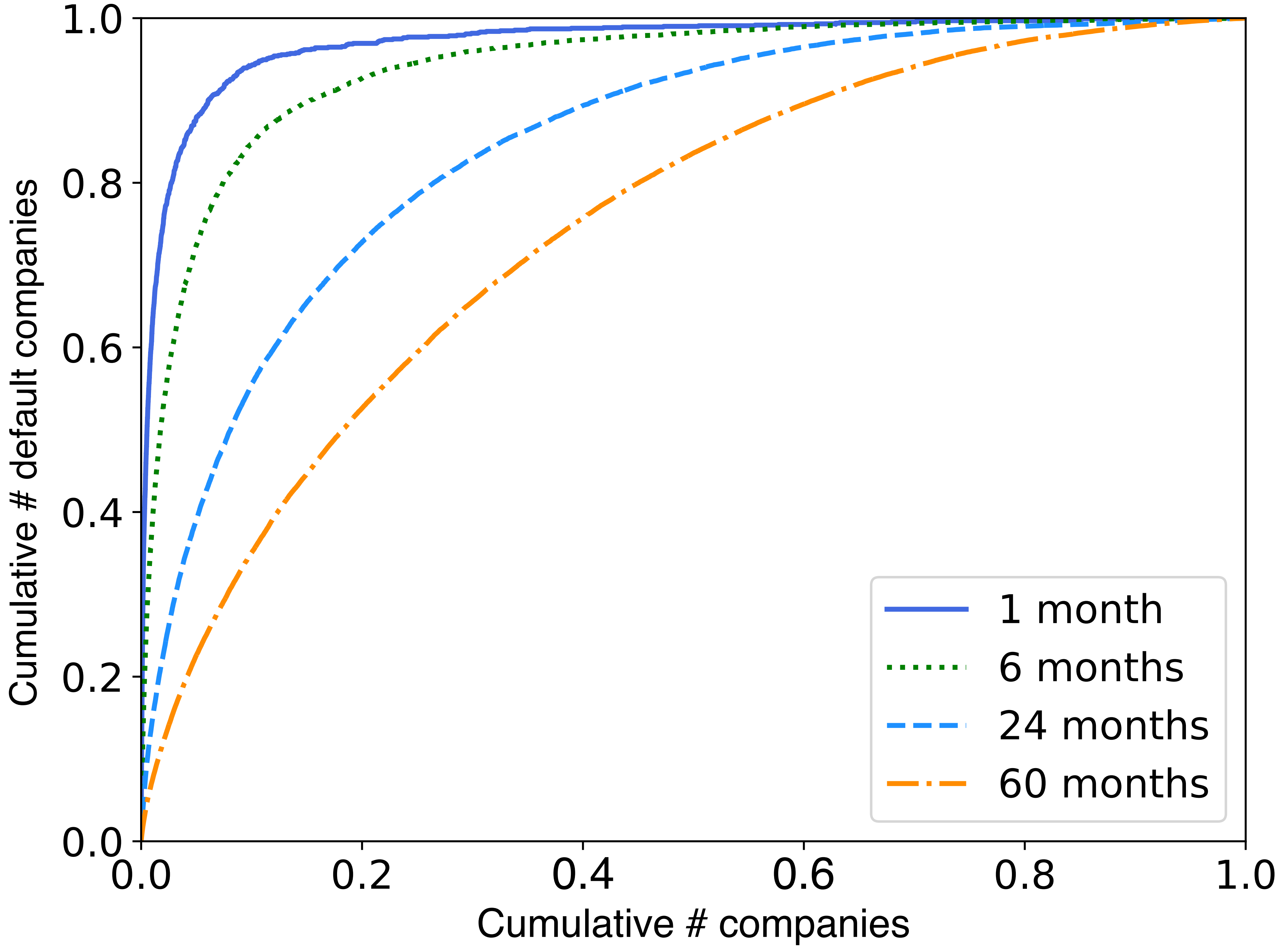
Number of default occurrences

FIM

Jin-Chuan Duan, Jie Sun, and Tao Wang. Multiperiod Corporate Default PredictionA forward Intensity Ap- proach. Journal of Econometrics, 170(1)(1):191–209, 2012.

Related works

FIM



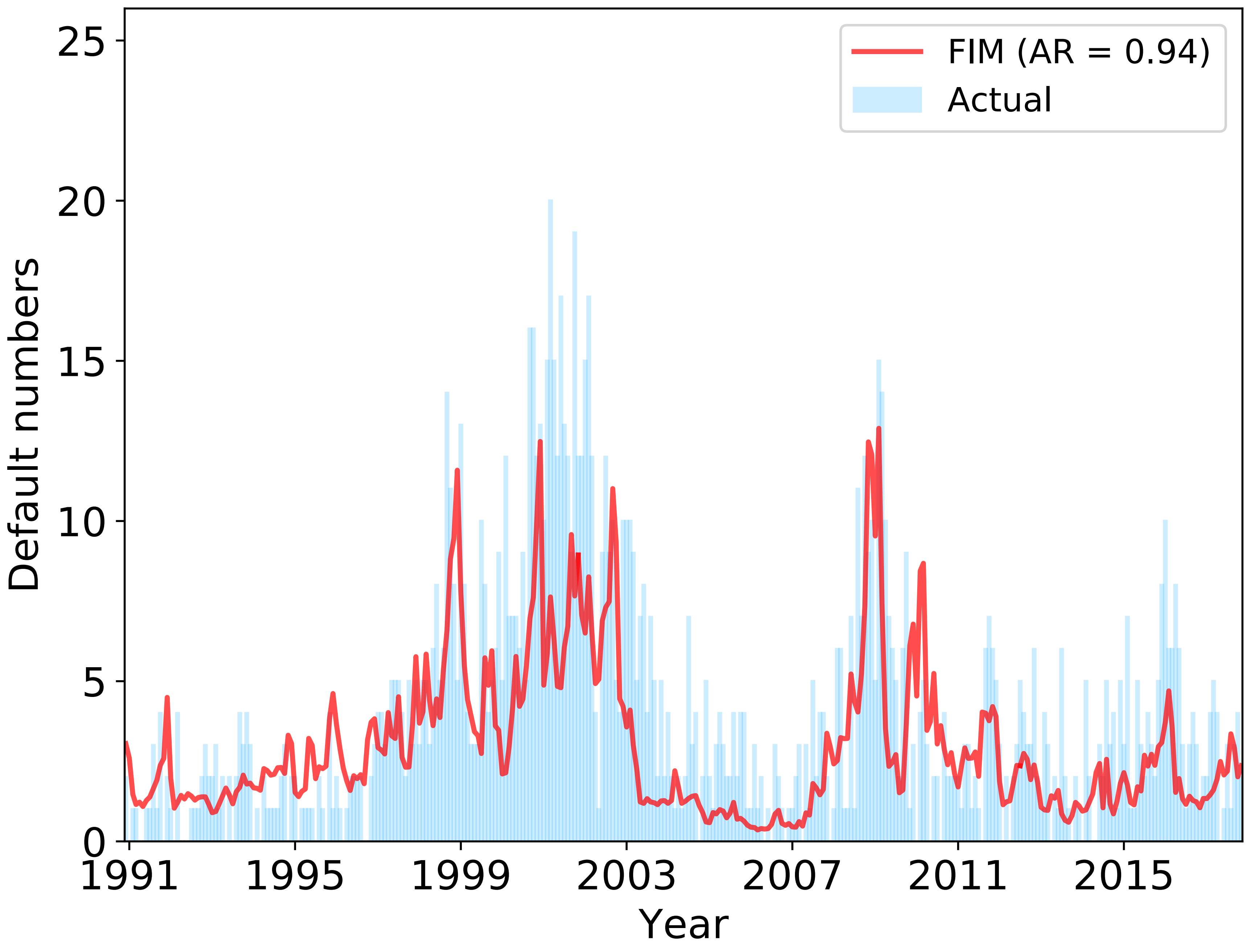
CAP Curve

performance deteriorates rapidly

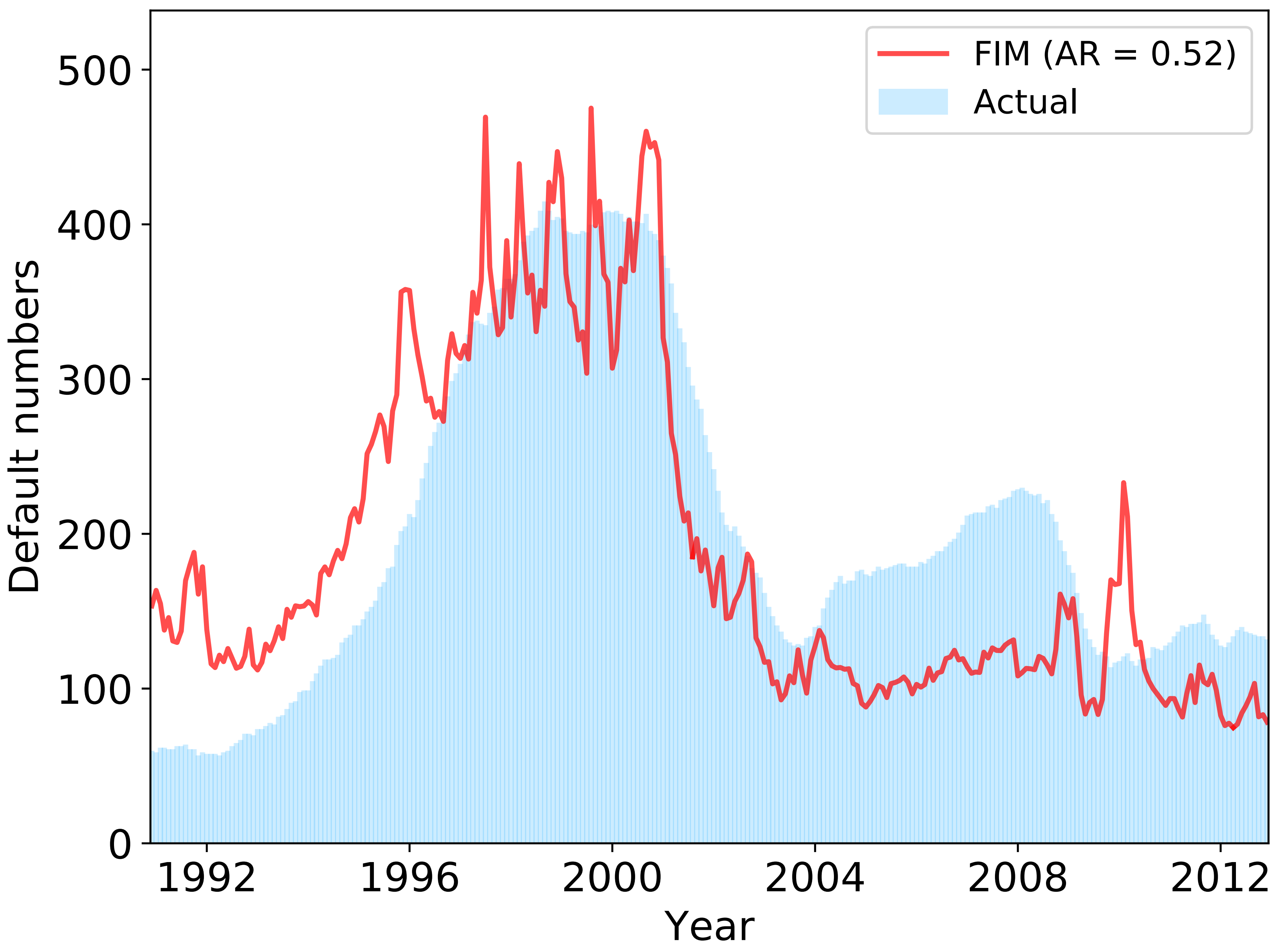
Jin-Chuan Duan, Jie Sun, and Tao Wang. Multiperiod Corporate Default PredictionA forward Intensity Ap- proach. Journal of Econometrics, 170(1)(1):191–209, 2012.

Related works

FIM



Prediction horizon: 1 month



Prediction horizon: 60 months

Jin-Chuan Duan, Jie Sun, and Tao Wang. Multiperiod Corporate Default PredictionA forward Intensity Ap- proach. Journal of Econometrics, 170(1)(1):191–209, 2012.

Related works

Default risk analysis

Machine learning

Risk classification (e.g. 3-months, 6 months)

Risk rankings

A consistent term structure

Real world applications

E.g. multi-period

Statistical

FIM

A term structure of CDP

Number of default occurrences

Rigorous assumption

E.g. same parametric family
for both long-term and short-term

Related works

Niche of each type of approaches

Machine learning

Risk classification (e.g. 3-months, 6 months)

Risk rankings

Rigorous assumption
E.g. same parametric family
for both long-term and short-term

GAP

A consistent term structure
Real world applications
E.g. multi-period

Statistical

FIM

A term structure of CDP

Number of default occurrences

Related works

Leverage big data

Machine learning

Risk classification (e.g. 3-months, 6 months)

Risk rankings

Rigorous assumption
E.g. same parametric family
for both long-term and short-term

GAP

A consistent term structure
Real world applications
E.g. multi-period

Statistical

A term structure of CDP

Number of default occurrences

FIM

Related works

Special model

Machine learning

Risk classification (e.g. 3-months, 6 months)

Risk rankings

Rigorous assumption
E.g. same parametric family
for both long-term and short-term

GAP

A consistent term structure
Real world applications
E.g. multi-period

Design a special model

Statistical

A term structure of CDP

Number of default occurrences

FIM

Methodology

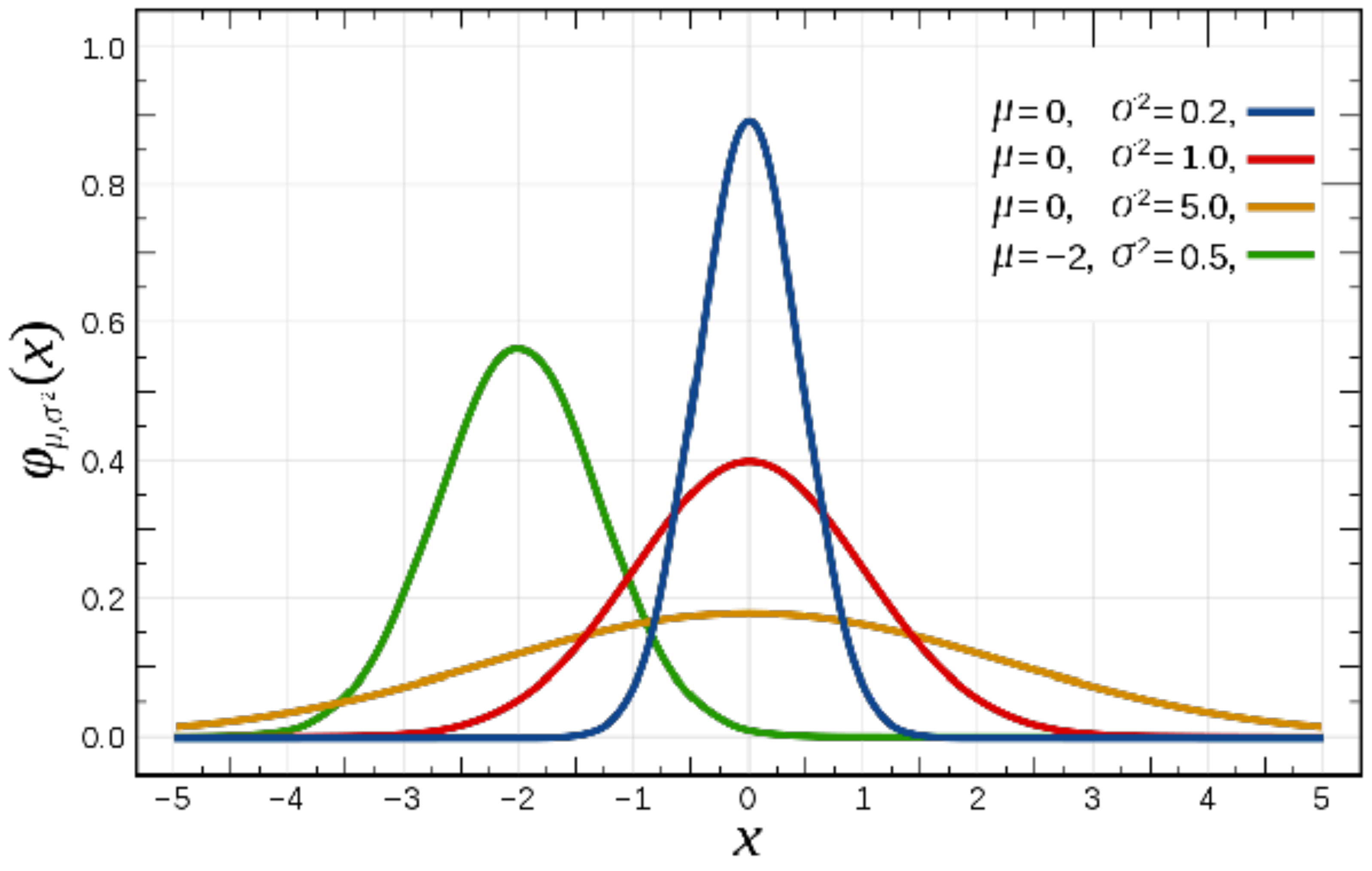
Neural Parametric Family Learning

Parameters

Mean
Standard Deviation

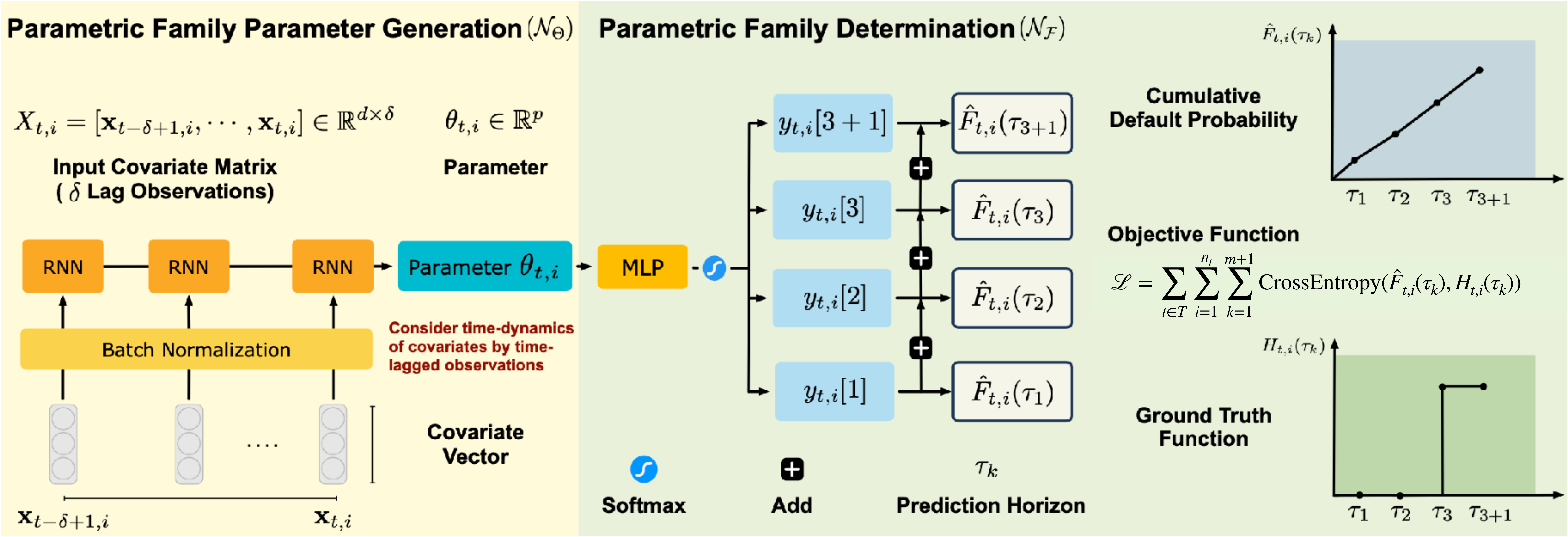
Parametric Family

Normal distribution
Pdf
Cdf



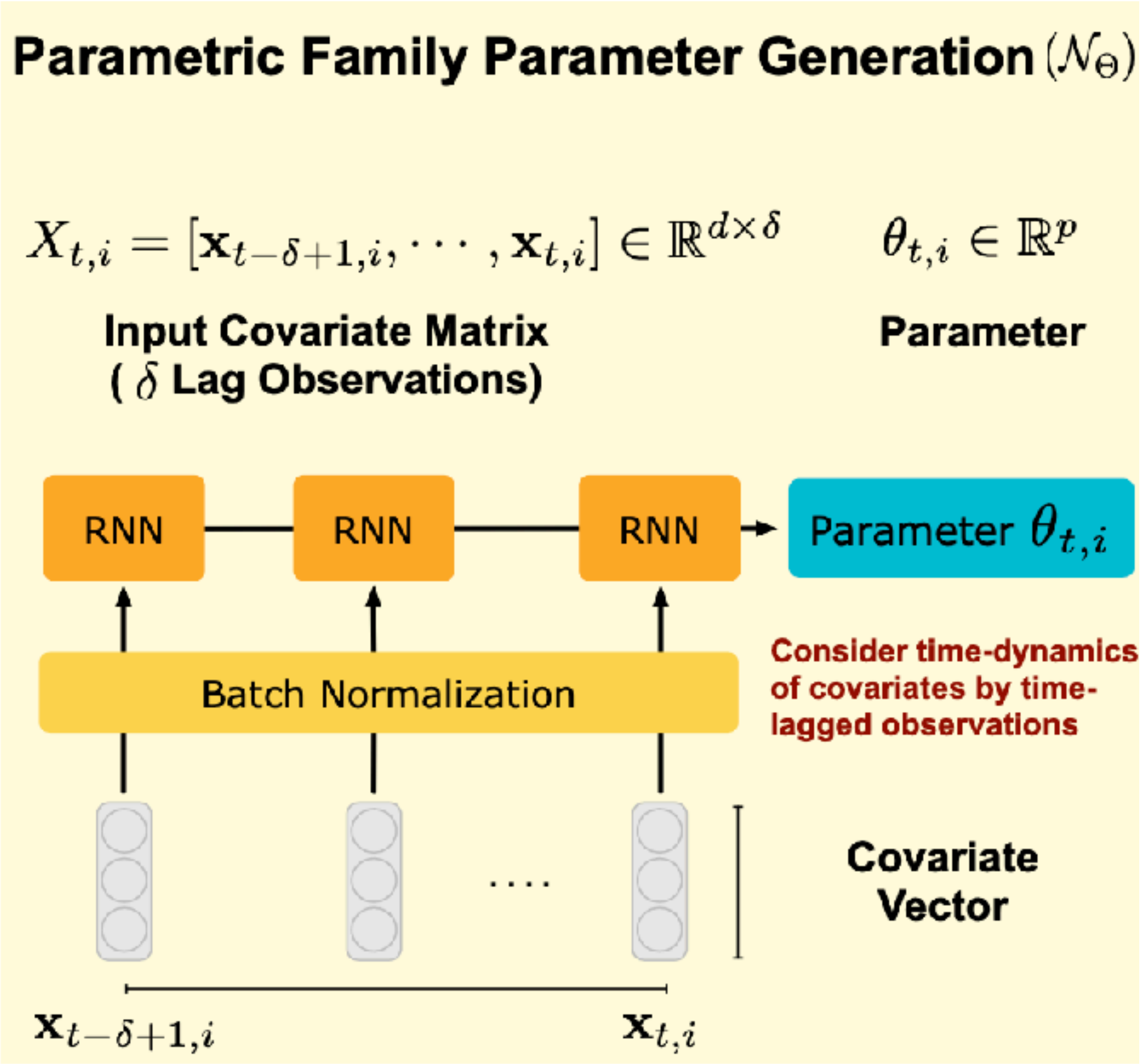
Methodology

Two phase



Methodology

The first phase



T — **The time point set**

I — **The company set**

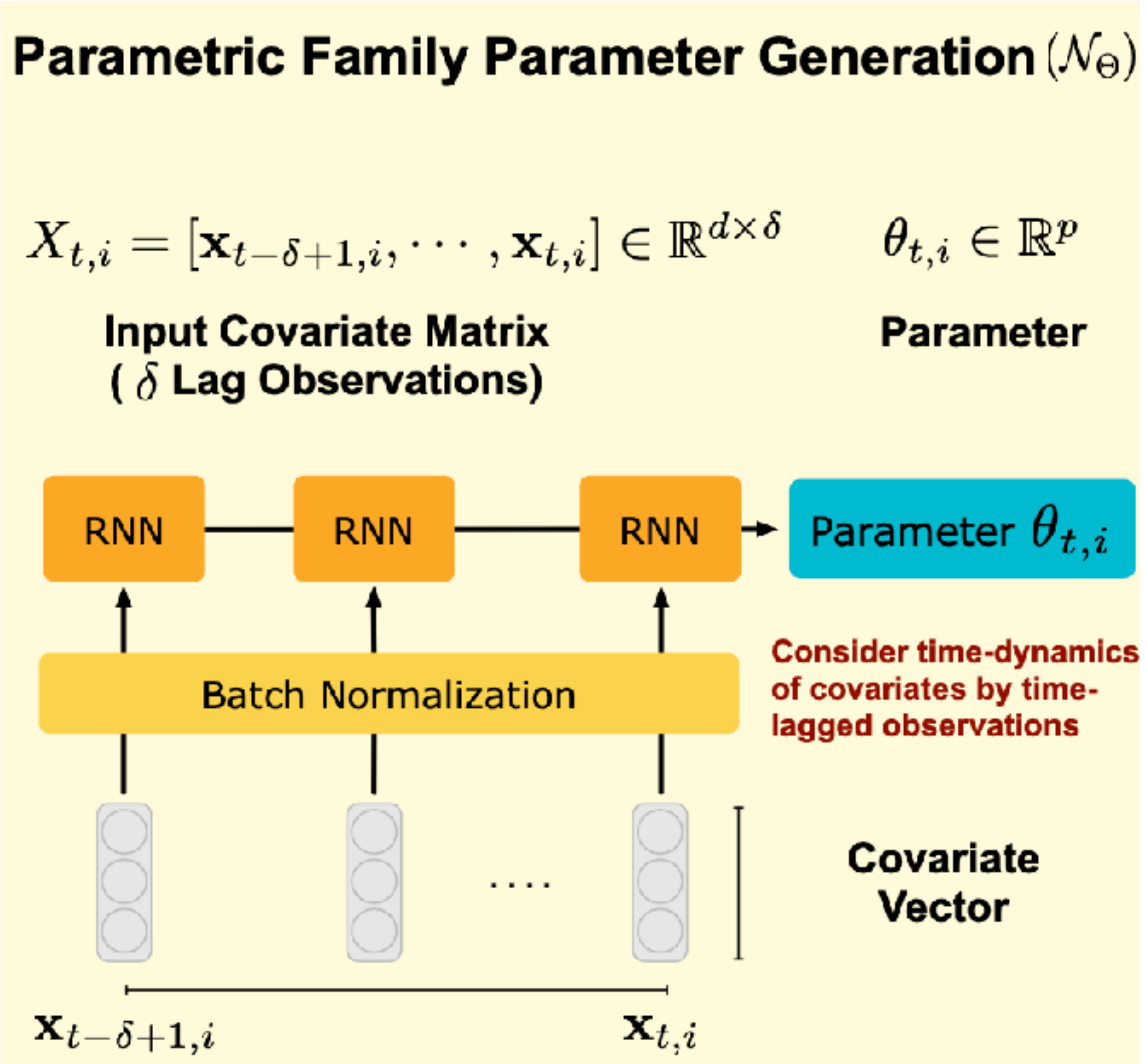
Input

$x_{t,i}$ d-dimension vector for the i-th company at time point t

$X_{t,i}$ A set of δ lag observations $x_{t,i}$

Methodology

The first phase



T — The time point set

I — The company set

Input

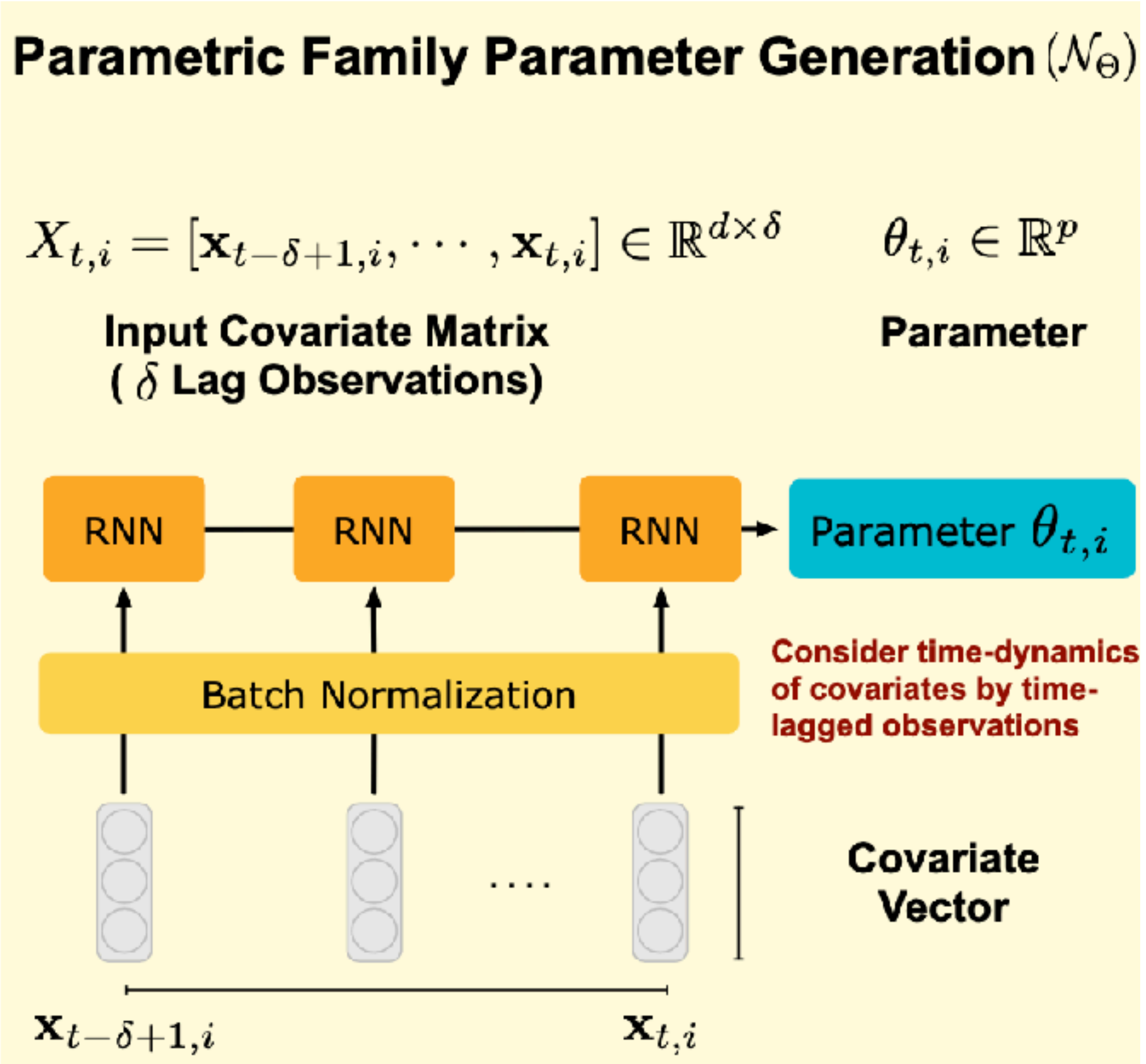
$x_{t,i}$ d-dimension vector for the i-th company at time point t

$X_{t,i}$ A set of δ lag observations $x_{t,i}$

t = 2008/9, δ = 3 $\quad X_{2008/9,i} = [x_{2008/7,i}, x_{2008/8,i}, x_{2008/9,i}]$

Methodology

The first phase



T — **The time point set**

I — **The company set**

Input

$x_{t,i}$ d-dimension vector for the i-th company at time point t

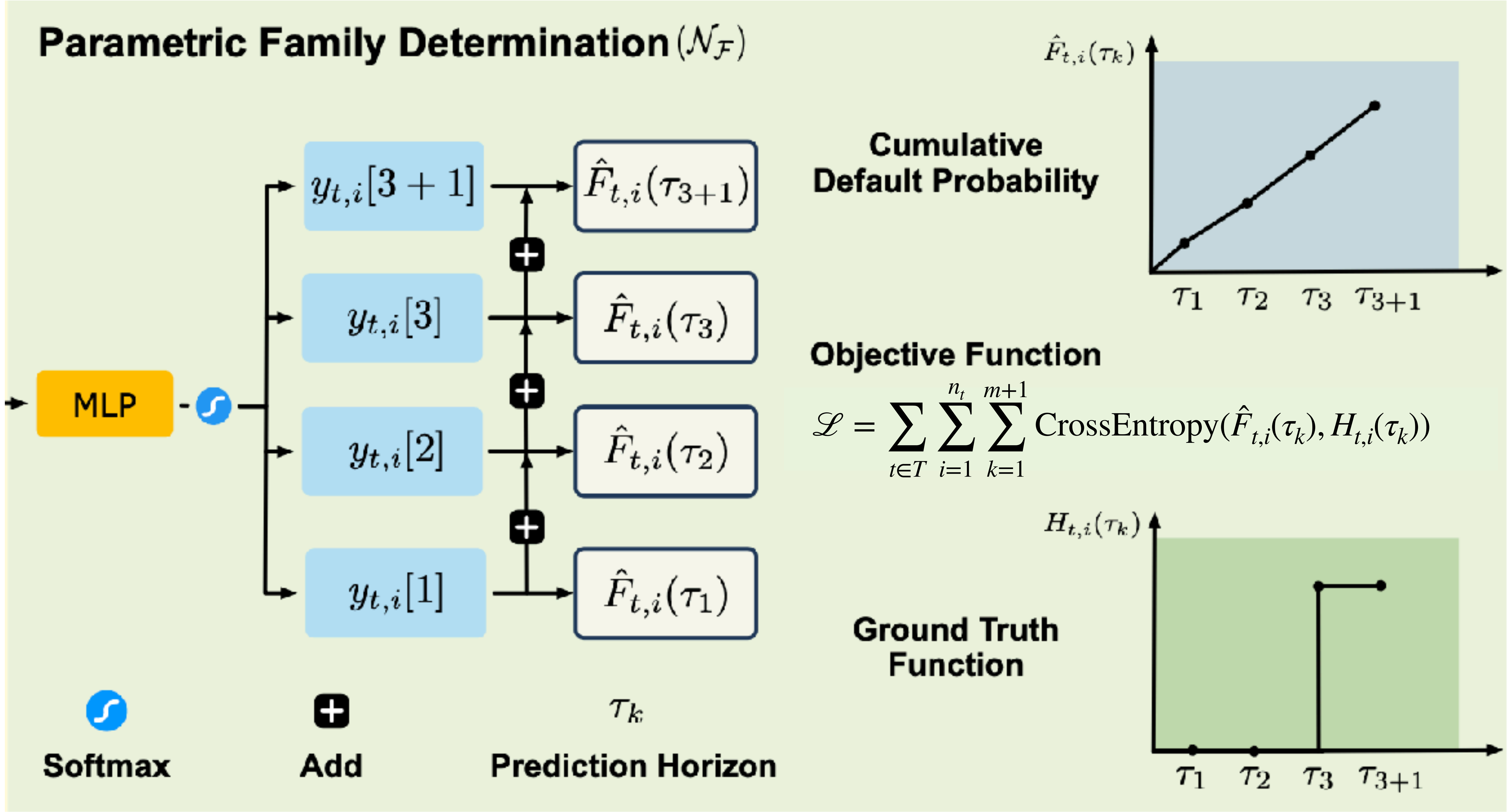
$X_{t,i}$ A set of δ lag observations $x_{t,i}$

Output

$\theta_{t,i}$ p-dimension vector for the i-th company at time point t

Methodology

The second phase



$$\hat{F}(\tau_\ell) = \sum_{k=1}^{\ell} y[k], \quad \text{for } \ell = 1, 2, \cdots, m+1$$

Marginal default probability

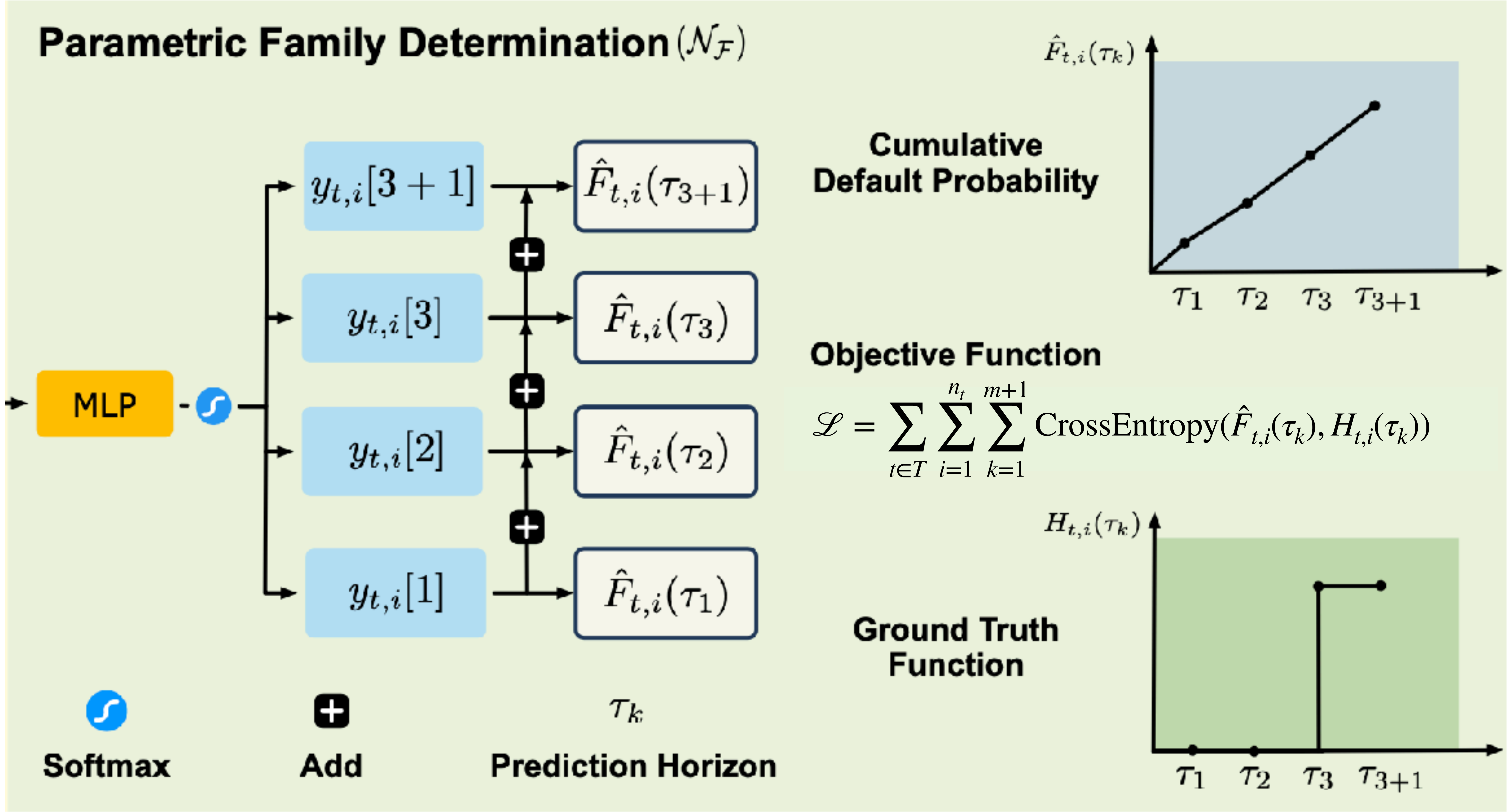
The company can not survive forever

$$H_{t,i}(s) = \begin{cases} 1, & \text{if } s \geq \zeta_i \\ 0, & \text{if } s < \zeta_i. \end{cases}$$

The time of the default

Methodology

The second phase



Solve GAP

A consistent term structure

Real world applications

E.g. multi-period

$$\hat{F}(\tau_\ell) = \sum_{k=1}^{\ell} y[k], \quad \text{for } \ell = 1, 2, \cdots, m+1$$

Marginal default probability

The company can not survive forever

$$H_{t,i}(s) = \begin{cases} 1, & \text{if } s \geq \zeta_i \\ 0, & \text{if } s < \zeta_i. \end{cases}$$

The time of the default

Dataset

NUS Credit Research Initiative (CRI)

Dates: January 1990 - December 2017

Data: 1.5 M monthly samples of US public companies

Covariates: 14, 2 common and 10 firm-specific covariates

Events: 0 (alive), 1 (default), 2 (other exit)

Prediction horizons: 60 months

Experiment

Two types

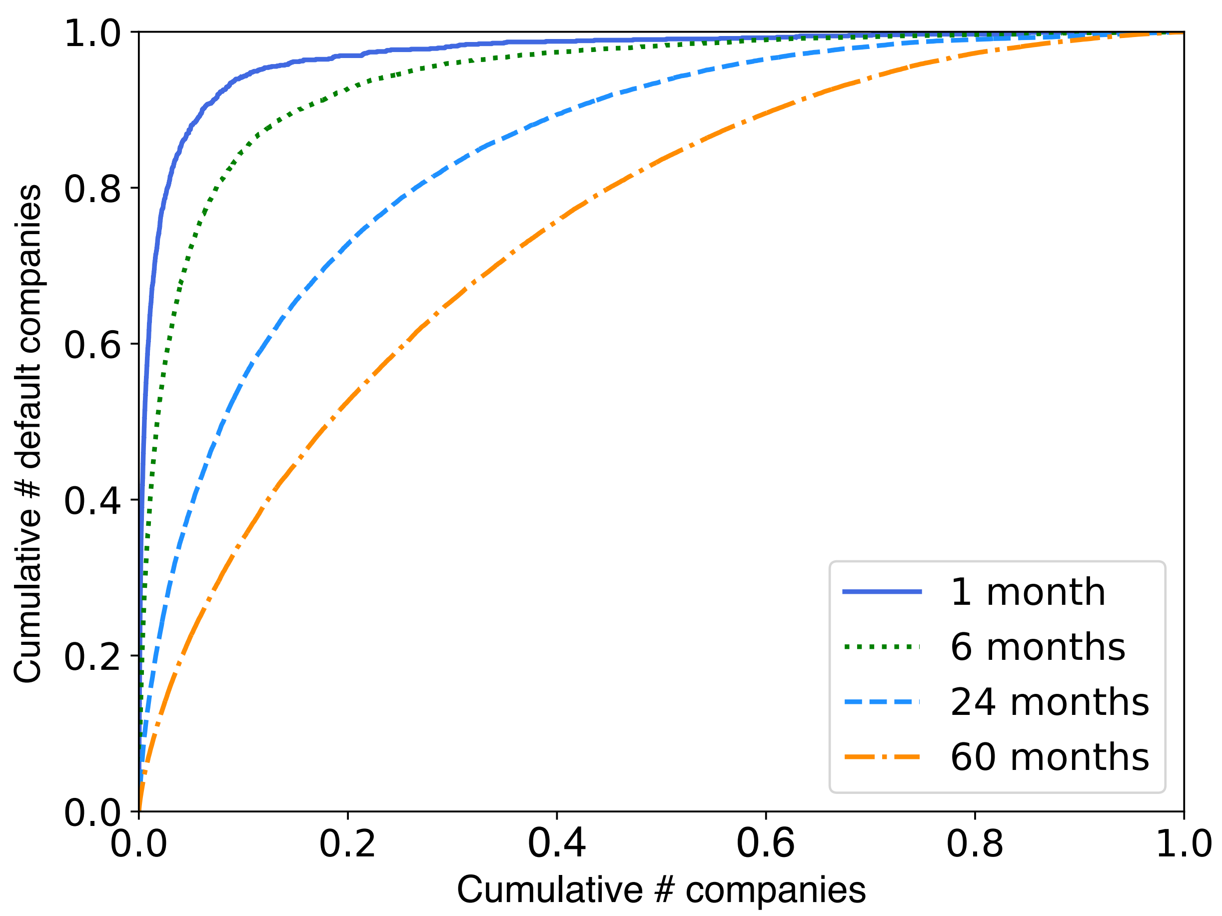
Cross-sectional: randomly split data into 13 folds

Cross-time →

	Train	Test
1	1990 - 1999	2000
2	1991 - 2000	2001
⋮		
13	2002 - 2011	2012

Metrics

CAP curve



Order

Accuracy ratio (AR, %)

$$AR = \frac{\text{Area above CAP curve}}{\text{Area under CAP curve}}$$

Value

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

$\hat{D}_i$ **Estimated default occurrences (monthly)**

D_i **Default occurrences (monthly)**

Results

Table 1: Results of cross-sectional experiments

Horizons (months)	1	3	6	12	24	36	48	60
Panel A	Accuracy ratio (AR) (%)							
FIM	94.57	92.37	88.74	81.45	70.85	63.46	58.33	53.37
MLP ($\delta = 1$)	94.48	92.85	90.43	85.10	75.63	68.08	62.87	58.26
MLP ($\delta = 6$)	94.29	92.76	90.47	85.73	76.88	69.73	64.55	60.07
MLP ($\delta = 12$)	93.99	92.64	90.55	86.05	77.67	70.81	65.93	61.45
LSTM ($\delta = 1$)	**94.78**	93.17	90.87	86.11	77.47	70.69	65.70	61.09
LSTM ($\delta = 6$)	94.63	93.29	91.23	87.05	79.00	72.63	67.55	62.96
LSTM ($\delta = 12$)	94.68	**93.48**	**91.77**	**87.91**	**80.79**	**74.76**	**69.91**	**65.32**
GRU ($\delta = 1$)	94.66	93.03	90.77	85.94	77.21	70.34	65.39	60.79
GRU ($\delta = 6$)	94.41	92.97	90.84	86.54	78.26	71.60	66.45	61.91
GRU ($\delta = 12$)	94.26	92.94	91.12	86.98	79.22	72.77	67.80	63.27
Improvement (%)	0.22	1.20	3.41	7.93	14.03	17.81	19.85	22.39
Panel B	Root mean square normalized error (RMSNE)							
FIM	0.74	0.64	0.62	0.84	1.23	1.18	1.06	0.96
MLP ($\delta = 1$)	0.63	0.58	0.62	0.88	1.03	1.30	1.24	1.11
MLP ($\delta = 6$)	0.64	0.58	0.61	0.86	1.23	1.32	1.26	1.12
MLP ($\delta = 12$)	0.63	**0.57**	**0.60**	0.83	1.21	1.27	1.17	1.03
LSTM ($\delta = 1$)	0.62	0.60	0.64	0.89	1.26	1.30	1.23	1.11
LSTM ($\delta = 6$)	0.64	0.61	0.62	0.86	1.23	1.25	1.19	1.07
LSTM ($\delta = 12$)	0.64	0.62	0.61	**0.81**	**1.11**	**1.12**	**1.03**	**0.90**
GRU ($\delta = 1$)	**0.61**	0.61	0.65	0.91	1.25	1.32	1.23	1.11
GRU ($\delta = 6$)	0.64	0.63	0.64	0.87	1.24	1.29	1.22	1.11
GRU ($\delta = 12$)	0.64	0.64	0.64	0.83	1.13	1.18	1.10	0.98
Improvement (%)	17.57	10.94	3.23	3.57	9.76	5.08	2.83	6.25

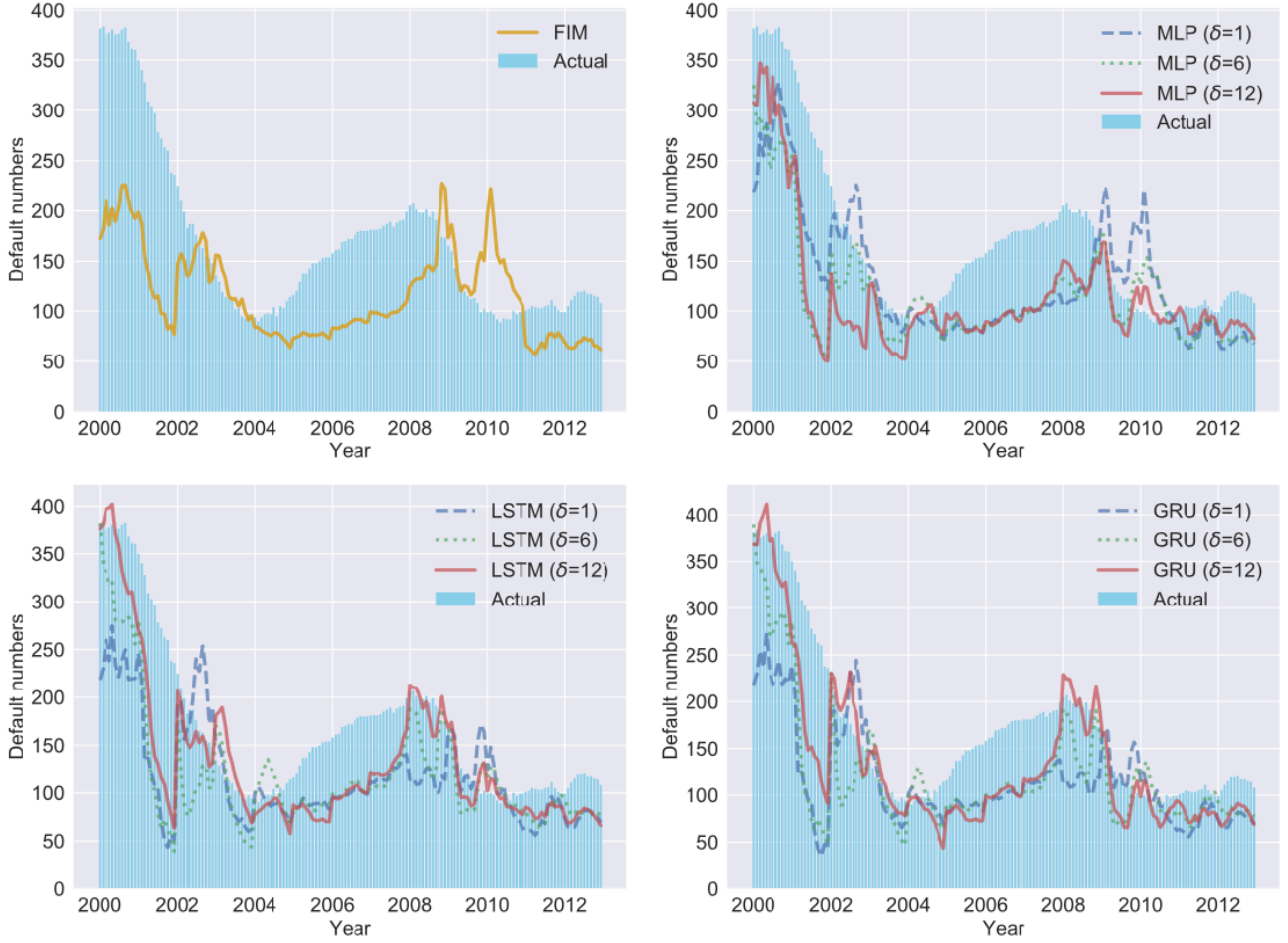
Results

Table 2: Results of cross-time experiments

Horizons (months)	1	3	6	12	24	36	48	60
Panel A	Accuracy ratio (AR) (%)							
FIM	**94.08**	91.86	87.74	81.88	74.86	69.20	64.40	59.61
MLP ($\delta = 1$)	93.69	91.76	89.26	84.92	78.06	72.16	67.30	62.63
MLP ($\delta = 6$)	93.30	91.52	89.10	85.05	78.44	72.48	67.45	62.72
MLP ($\delta = 12$)	92.77	91.11	88.78	85.05	78.61	72.81	67.92	63.21
LSTM ($\delta = 1$)	93.67	**92.03**	**89.54**	85.45	78.67	72.88	67.89	63.38
LSTM ($\delta = 6$)	93.46	91.84	89.41	85.43	78.70	72.87	67.90	63.26
LSTM ($\delta = 12$)	92.81	91.27	88.96	85.27	78.57	72.79	67.70	62.77
GRU ($\delta = 1$)	93.54	91.87	89.53	**85.53**	78.63	72.79	**68.05**	**63.49**
GRU ($\delta = 6$)	93.48	91.91	89.51	85.45	78.65	72.83	67.86	63.25
GRU ($\delta = 12$)	93.03	91.45	89.26	85.34	**78.76**	**72.89**	67.98	63.35
Improvement (%)	0	0.19	2.05	4.46	5.21	5.33	5.67	6.51
Panel B	Root mean square normalized error (RMSNE)							
FIM	1.09	0.77	0.51	0.47	0.40	0.36	0.39	0.39
MLP ($\delta = 1$)	0.83	**0.60**	0.43	0.44	0.38	0.34	0.35	0.34
MLP ($\delta = 6$)	0.73	**0.60**	0.40	0.40	0.34	0.33	0.35	0.33
MLP ($\delta = 12$)	**0.72**	0.62	0.40	0.37	0.34	0.31	0.32	0.32
LSTM ($\delta = 1$)	1.00	0.67	0.43	0.40	0.37	0.34	0.35	0.35
LSTM ($\delta = 6$)	0.82	0.64	0.41	0.38	0.32	0.32	0.33	0.31
LSTM ($\delta = 12$)	0.97	0.61	**0.34**	0.33	0.28	**0.26**	0.27	**0.25**
GRU ($\delta = 1$)	1.08	0.69	0.41	0.39	0.36	0.34	0.34	0.33
GRU ($\delta = 6$)	0.86	**0.60**	0.39	0.36	0.32	0.32	0.32	0.31
GRU ($\delta = 12$)	1.14	**0.60**	**0.34**	**0.28**	**0.26**	**0.26**	**0.26**	0.26
Improvement (%)	33.95	22.08	33.33	40.43	35.00	27.78	33.33	35.90

Results

48-month prediction horizon



Conclusion

Multiperiod default prediction

Real world practical scenarios

A term structure of monotonically increasing CDP
Default occurrences

Outperform the SOTA

Thank you