Non-Performing Assets and Lending Growth

Amartya Lahiri* Urvi Neelakantan†

*CAFRAL and University of British Columbia

†CAFRAL and FRB Richmond

March 2019

Questions

- Has bank lending followed standard profitability criteria?
- Have NPAs been driven by borrower productivity?
- How are loans, productivity and NPAs related?

This Presentation

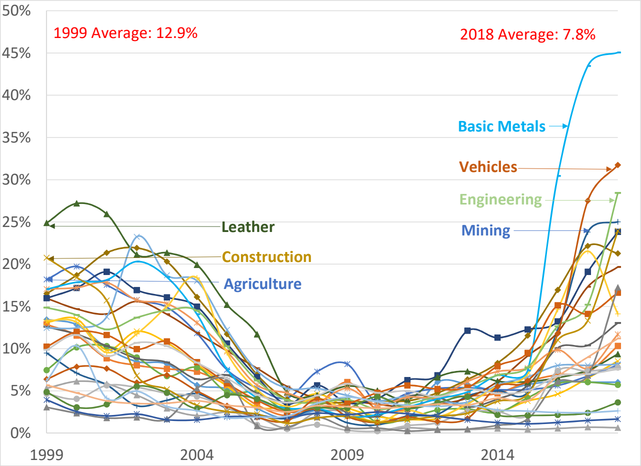
- Examine questions at the sectoral level
- Use variation in loans, NPAs and total factor productivity (TFP)
  - cross-sectional variation across 19 sectors
  - time-series variation between 1999-2018

Data Sources

- Bank Statistical Returns (BSR) data
  - amount outstanding, by sector, between 1999-2018
  - outstanding loans broken into four groups
  - groups 2-4 classified as NPA
- Indian KLEMS data
  - sectoral output, input and TFP data
  - covers 1981-2015
- PROWESS database for sectoral q

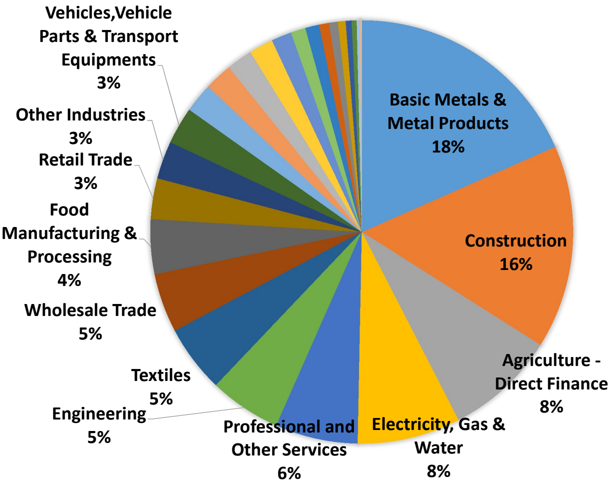
Pattern 1: Time trend of NPA

Figure: What fraction of each sector's loans are NPAs?



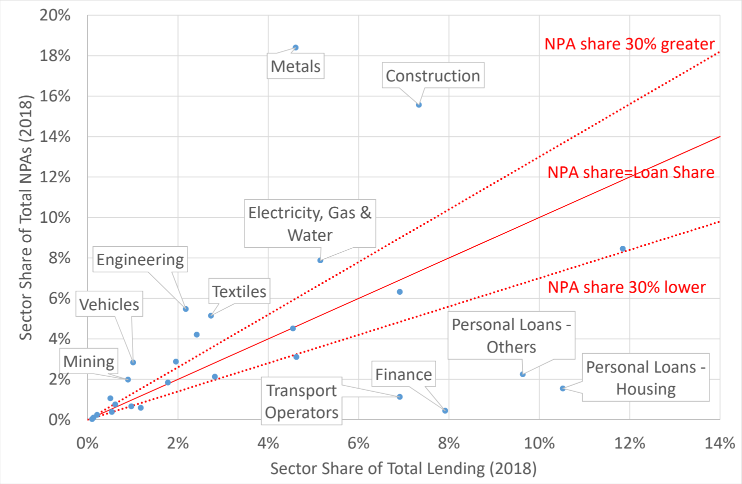
Pattern 2: Sectoral Share of NPAs in 2018

- Four sectors currently account for half of all NPAs: Metals, Construction, Agriculture and Electricity/Gas/Water



Pattern 3: Loan share v NPA share in 2018

Figure: Sector Share of Total Lending vs. Sector Share of Total NPAs



Pattern 4: Sectoral NPAs – Then and Now

What fraction of each sector's loans are NPAs (1999 vs. 2018)?

Sector	NPA Share		Sector	NPA Share	
	1999	2018		1999	2018
Basic Metals & Metal Products	17%	45%	Professional and Other Services	13%	10%
Vehicles, Vehicle Parts & Transport Equipment	6%	32%	Leather & Leather Products	25%	9%
Engineering	15%	28%	All Others	13%	9%
Mining & Quarrying	9%	25%	Rubber & Plastic Products	12%	8%
Construction	21%	24%	Agriculture - Direct Finance	18%	8%
Paper, Paper Products & Printing	16%	24%	Petroleum, Coal Products & Nuclear Fuels	5%	8%
Textiles	17%	21%	Retail Trade	10%	8%
Food Manufacturing & Processing	16%	20%	Transport Operators	13%	6%
Electricity, Gas & Water	3%	17%	Agriculture - Indirect Finance	7%	6%
Other Industries	10%	17%	Personal Loans - Consumer Durables	5%	4%
Manufacture of Cement & Cement Products	13%	14%	Personal Loans - Others	6%	3%
Beverage & Tobacco	13%	13%	Personal Loans - Housing	4%	2%
Chemicals & Chemical Products	17%	12%	Finance	5%	1%
Wholesale Trade	6%	11%			

Profit-Based Lending

- Sectors with high TFP should have higher profits
- Profit maximizing lenders should lend more to sectors with higher profits

  $\Rightarrow$ Sectoral loans and TFP should be positively correlated
- High profit sectors should have lower NPAs

  $\Rightarrow$ Sectoral TFP and NPAs should be negatively correlated
- Does this show up in the data?

Aggregate Data

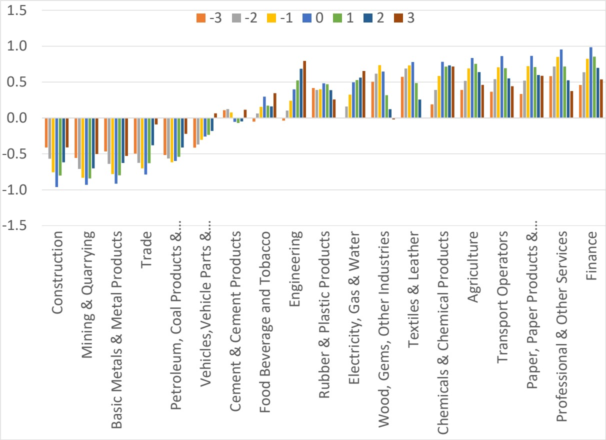
Correlations

	NPA Share	Loan Growth	TFP Growth
NPA Share	1		
Loan Growth	-0.1957	1	
TFP Growth	-0.5349**	0.5102*	1

Lending and Productivity: Sectoral Data

Expect Positive Correlations

Figure: Lead-Lag Correlation Between TFP Growth and Lending Growth Rates (1999-2015)

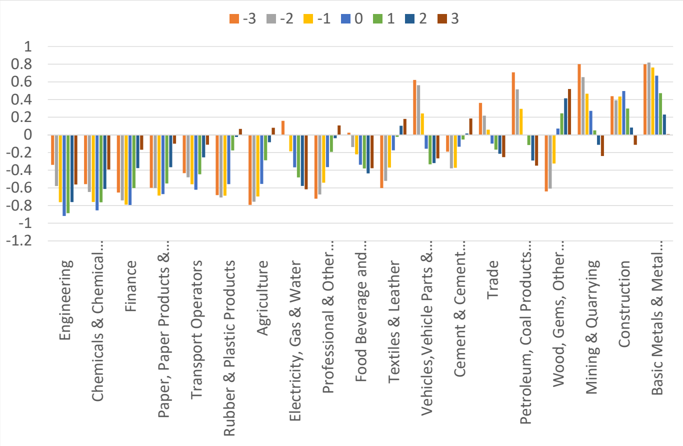


Note: TFP and lending indexed to 100 in 1999. Certain KLEMS sectors were combined to align with BSR sector codes. TFP Growth is anchor variable, Lending Growth Rate's lag and lead are taken.

Productivity and NPA: Sectoral Data

Expect Negative Correlations

Figure: Lead-Lag Correlation Between TFP Growth and Fraction of Sector Loans that are NPAs (1999-2015)



Note: TFP and lending indexed to 100 in 1999. Certain KLEMS sectors were combined to align with BSR sector codes. TFP Growth is Anchor variable, NPA share in total sector's lag and lead are taken

Summary of Sectoral Data Patterns

- Basic Metals, Construction and Mining are common aberrations
- Collectively account for 36 percent of NPAs
- Account for only 13 percent of loans outstanding
- Are these sectors statistical outliers?

Econometric Testing

- Formally examine relationship between loans, productivity and NPAs
- Which sectors are econometric outliers?

Model of Lending

- Baseline specification

$$\Delta L_{it} = F\left(\frac{NPA_{it-1}}{NPA_{t-1}}, q_{it-1}, \frac{D_{it-1}}{A_{it-1}}\right) + \epsilon_{it}$$

- $q = \frac{\text{market value}}{\text{book value}}$ is Tobin's q
- $\frac{D_{it-1}}{A_{it-1}}$ is debt-to-asset ratio
- Problem: NPA respond to q and debt-to-asset ratio

Two-Step Process

- Step 1: Estimate

$$\frac{NPA_{it}}{NPA_t} = G\left(\frac{L_{it-1}}{L_{t-1}}, q_{it-1}, \frac{D_{it-1}}{A_{it-1}}, CPI_{t-1}\right) + \eta_{it}$$

- Step 2: Estimate

$$\Delta L_{it} = F\left(\hat{\eta}_{it-1}, q_{it-1}, \frac{D_{it-1}}{A_{it-1}}, t\right) + \epsilon_{it}$$

- CPI is commodity price index

- $\hat{\eta}_{it}$ measures NPAs not due to other variables

Null Hypothesis

- Expect coefficient on loan share to be one
  - if NPAs are random then expected NPAs are zero
  - sectors with higher loan shares will have higher NPA shares
- Coefficient below one indicates banks doing better than randomly allocating loans
- Coefficient above one indicates coin toss would do better

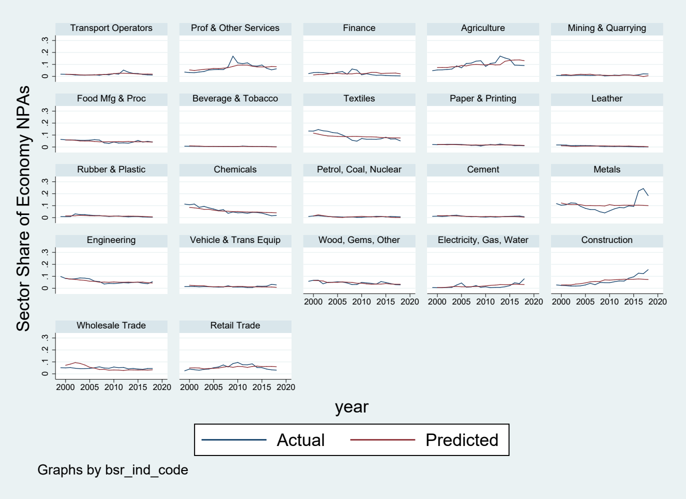
First-step Result: Sector Share of NPAs

	(1)	(2)	(3)	(4)
Sector Share	0.863^{***}	0.835^{***}	0.657^{***}	0.612^{***}
of Loans Lagged	(18.37)	(18.65)	(10.36)	(9.09)
Debt-to-Asset		0.0912^{***}		0.0331^{**}
Ratio Lagged		(7.68)		(1.98)
Tobin's Q		0.00542^{***}		-0.000457
Lagged		(3.69)		(-0.31)
Comm. Price		-0.0000185		0.00000538
Index Lagged		(-0.41)		(0.15)
Sector	No	No	Yes	Yes
Fixed Effects			(.)	(.)
Constant	0.00794^{***}	-0.0264^{***}	0.00588	-0.00346
	(3.57)	(-4.45)	(1.25)	(-0.47)
Observations	418	418	418	418
R^2	0.448	0.520	0.737	0.739

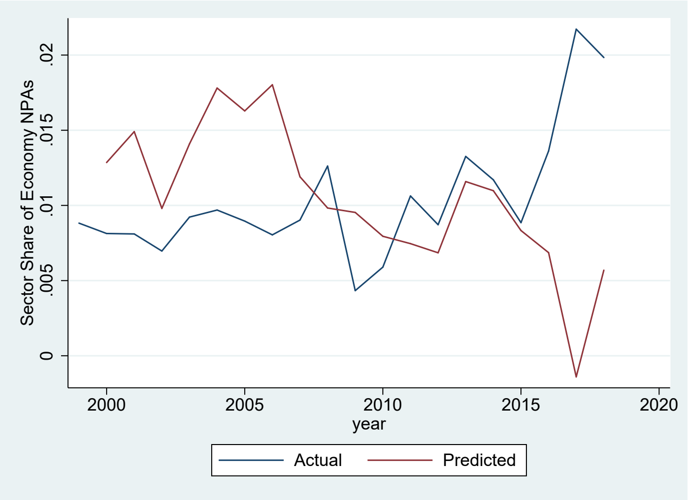
t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

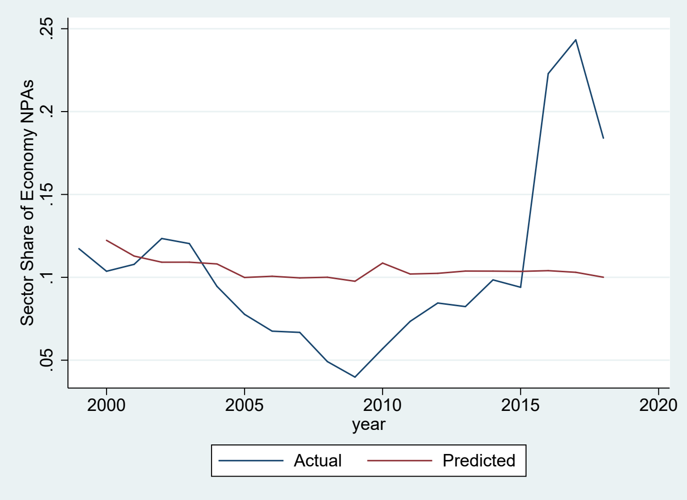
Actual and Predicted Share of NPAs by Sector



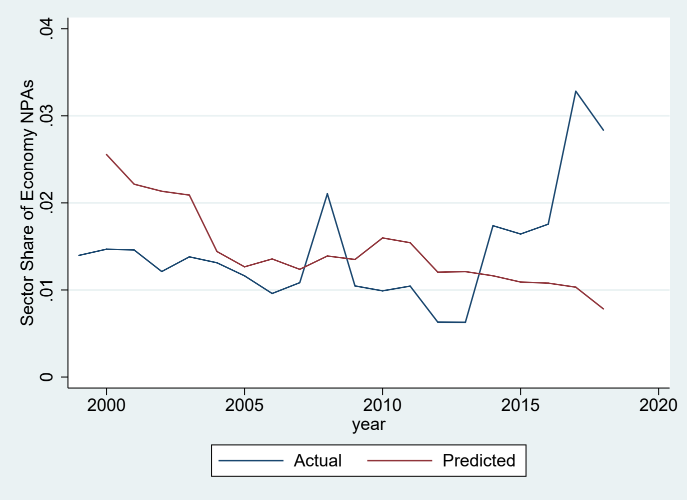
Actual and Predicted Share of NPAs: Mining



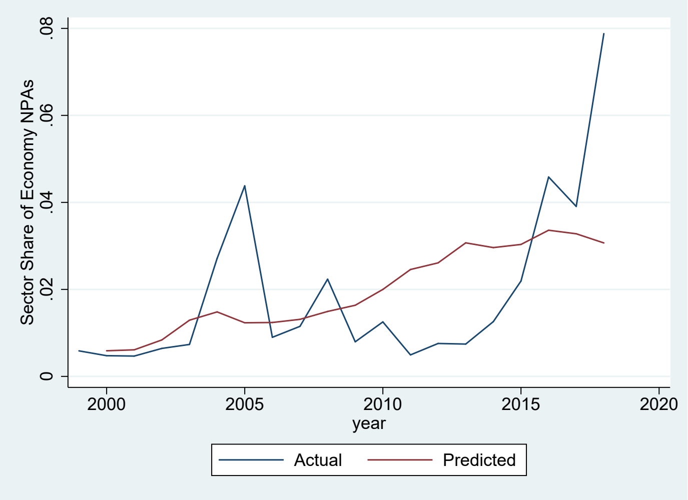
Actual and Predicted Share of NPAs: Metals



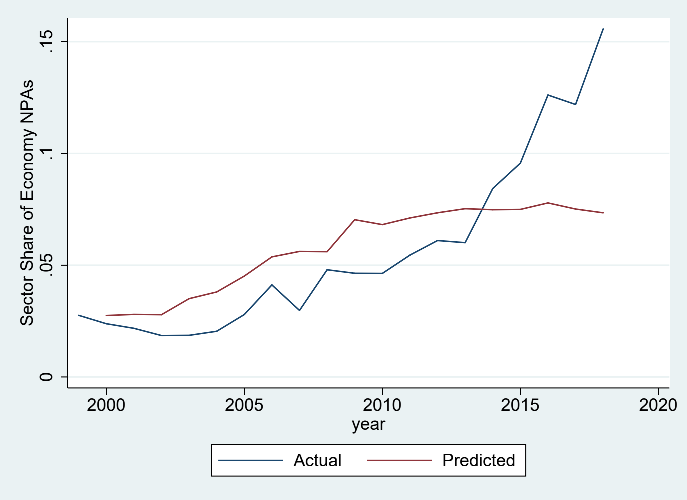
Actual and Predicted Share of NPAs: Vehicles



Actual and Predicted Share of NPAs: Electricity, Gas, Water



Actual and Predicted Share of NPAs: Construction



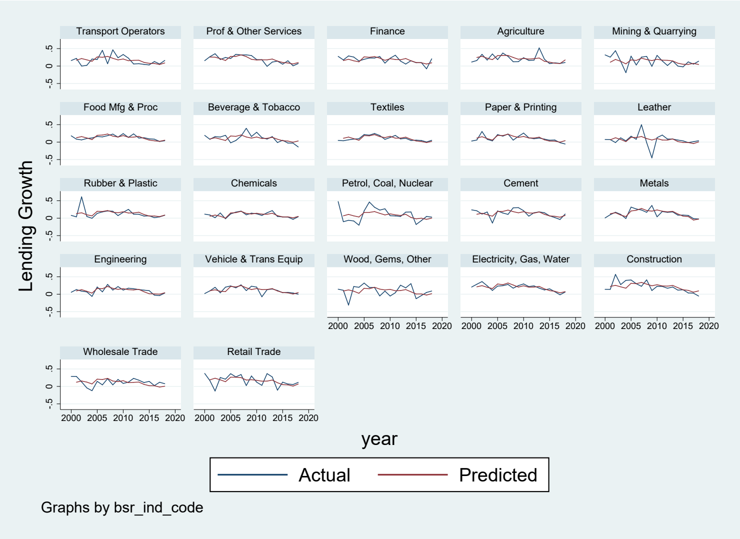
Second-step Result

	Lending Growth
Residual Lagged	-0.560*
	(-1.89)
Tobin's Q Lagged	0.00742
	(0.84)
Debt-to-Asset Ratio Lagged	-0.250***
	(-2.63)
Year Fixed Effects	Yes
Sector Fixed Effects	Yes
Observations	396
R^2	0.371

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Actual and Predicted Lending Growth



Conclusion

- NPAs have risen secularly since 2011, and spiked over the past three years
- A few sectors have been outliers, construction and metals being two such
- NPAs of outliers have exceeded the amount that fundamentals can explain
- Banks, on average, did well: sectoral NPA shares $<$ sectoral loan shares
- Problem appears to be specific misallocation, not general inefficiency

Sector Share of NPAs and Sector Share of Risky Firms

	Sector Share of NPAs	Debt-to-Asset Ratio	Sector Share of Risky Firms
Sector Share of NPAs	1.0000		
Debt-to-Asset Ratio	0.3562***	1.0000	
Sector Share of Risky Firms	0.5035***	0.1757***	1.0000

Risky firms refers to firms in the top quartile of fitted probability of default.

Probability of Default Series: Methodology

A. Sample Construction

- Data regarding listed firms characteristics and defaults is collected for the period 2002 – 2018 as follows:
  - **A1. Firm characteristics data** : Accounting and stock market data gives a daily panel of all listed firms
  - **A2. Firm level default data** : Sample of firms that have defaulted on their debt payment obligations in each quarter
- A 'quarterly default dummy' is created, which is set to 0 if for a given quarter a firm appears in the characteristic dataset but not the default dataset.If a firm appears in both the datasets, the variable is set to 1.

B. Variable Description

- **B1. Dependent Variable**: YDD (Yearly Default Dummy)
  - If for firm i in year t, the sum of 'quarterly default dummy' from Q1 to Q4 is greater than 0, then YDDit is 1, else 0.
  - After the first occurrence of YDD being 1 for a company, all subsequent observations for the company are dropped.

Probability of Default Series: Methodology

- **B2. Independent Variables**
  The independent variables are calculated using the dataset described in A1. For obtaining an annual measure of each of these variables, the value is taken as of the 1st of Jan for company i in year t in the yearly panel.
  - **Distance to Default (DTD)** : Using the KMV Merton distance-to-default methodology
  - **Interest Expense to PBDITA (Int2PB)** : $\frac{Interest\ \ Expense}{PBDITA}$
  - **Book-to- Market Ratio (Bk2Mkt)** : $\frac{Total\ \ Assets}{Market\ \ Capitalization\ + Total\ \ Liabilities}$
  - **Liquidity (Lqdty)** : $\frac{Cash\ \ and\ \ Cash\ \ Equivalents}{Total\ \ Assets}$
  - **Profitability (Pft)** : $\frac{Net\ \ Income}{Total\ \ Assets}$

Probability of Default Series: Methodology

C. Logistic Regression

- For obtaining the coefficients required for calculating the fitted PDs for year t, a rolling regression of the following form is estimated on the yearly panel over the last four years i.e. over window r ranging from t-4 to t-1:

$$YDD_r = \alpha + \beta_1 DTD_r + \beta_2 Int2PB_r + \beta_3 Bk2Mkt_r + \beta_4 Lqdty_r +$$

$$\beta_5 Pft_r + \epsilon_r ....(1)$$

- Thus, for the period 2002 to 2018, we get 13 coefficients (one for each year, starting in 2006 and ending in 2018).

Probability of Default Series: Methodology

D. Daily Fitted PDs

Base data is the daily panel of listed firms as created in A1. The daily fitted PD are calculated as follows:

- For each firm i for each day j in a given year t, the Sum is calculated using the coefficients obtained in equation (1) as follows:

$$Sum_{itj} = \alpha + \beta_1 DTD_{itj} + \beta_2 Int2PB_{itj} + \beta_3 Bk2Mkt_{itj} + \beta_4 Lqdty_{itj} + \beta_5 Pft_{itj} \text{....(2)}$$

- For each day j, the fitted PD (FittedPD) is calculated by using Sumitj in a logistic function as follows:

$$FittedPD_j = \frac{e^{Sum_{itj}}}{1 + e^{Sum_{itj}}}$$

E. Annual Fitted PDs

For each firm i in a given year t, we obtain the annual fitted PD value by taking an average across the daily values in that year.

Measure Of Riskiness : Methodology

A.Classification Of Firms Based on Risk

- Based on the distribution of fitted PD variable, a threshold value is defined to classify each of the firms as high risk or low risk. For a firm i in a year t, if the fitted PD value is greater than the threshold, then the variable 'high risk dummy' (HRDit) takes the value 1, otherwise 0.
- Two different 'high risk dummy' variables are generated using the threshold as 75th and 80th percentile value of fitted PD respectively.

B.Calculating Industry Risk

To generate the risk measure at an industry level, we define a new variable which is calculated as follows:

- For an industry m in year t,

$$RM_{mt} = \frac{No\ of\ firms\ in\ industry\ m\ with\ HRD_{it} = 1}{Total\ no.\ of\ firms\ with\ HRD_{it} = 1}$$