EUROPEAN CENTRAL BANK
EUROSYSTEM

Exploring non-linearities using the micro-aggregated CompNet database:

The investment-leverage relationship

Heng Chen (BOC)
Marie-Hélène Felt (BOC)
Annalisa Ferrando (ECB)

CompNet Workshop
Madrid 26 March 2015

The views expressed are our own and should not interpreted to reflect those of the Bank of Canada or the ECB

Motivation

- Starting from the empirical evidence of the existence of non-linearities in investment regressions

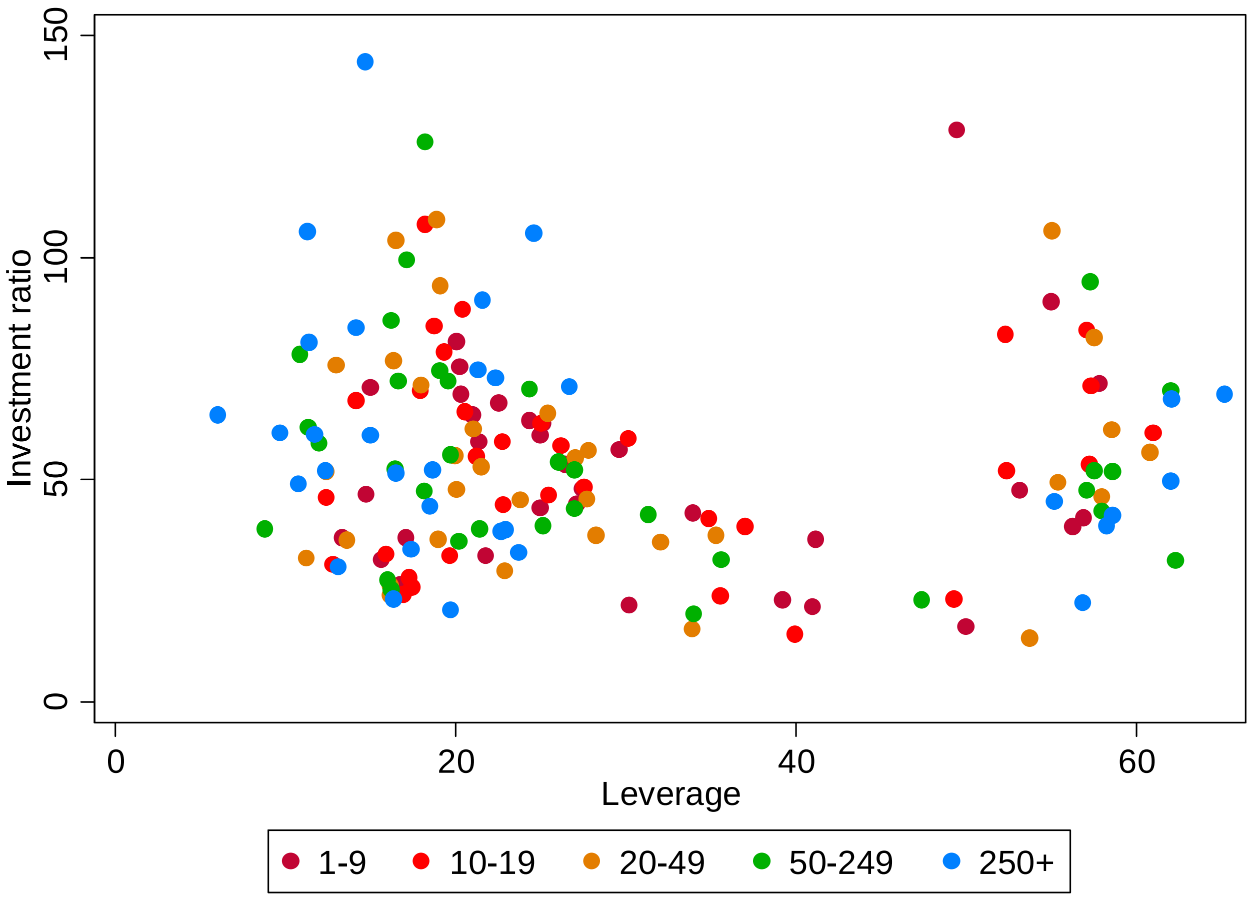
  ✓ Financial module's analysis

- How can we use the CompNet database to correctly estimate non-linearities?

- Two aims of the project :
  - Methodological (novelty!)
  - Empirical (we reached step 1 out of 3 steps)

STILL WORK IN PROGRESS!!!

Evidence of nonlinear impacts of indebtedness on investment:



Characteristics of the CompNet database

Starting from balance-sheet/profit and loss data:

Aggregates of individual firm accounting statements by country, sector of economic activity and firm size.

Regression analysis on micro-aggregated data:

Cell-based regressions using mean values: do we explain individual behavior?

- Implied assumption of micro homogeneity or representative/ typical agent
- Aggregation bias if *heterogeneity* (Theil, 1954) or *nonlinearity.*
- Aggregation gain (Grunfeld and Griliches, 1960).

Presence of non-linearities

- Numerous studies indicate **non-linear** impact of some determinants of firms' investment

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 X_{it}^2 + \alpha_i + u_{it}$$

- Non-linearities at the firm level imply that the **aggregate relationship** differ from the **individual relationship evaluated at the average**:

$$\sum_i Y_{it} = n\beta_0 + \beta_1 \sum_i X_{it} + \beta_2 \sum_i X_{it}^2 + \sum_i \alpha_i + e_{it}$$

$$\bar{Y}_t = \gamma_0 + \gamma_1 \bar{X}_t + \gamma_2 \bar{X}_t^2 + \bar{\alpha} + e_{it}$$

Non-linearities and firm-level data

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 X_{it}^2 + \alpha_i + u_{it}$$

$$Y_i^D = \beta_1 X_i^D + \beta_2 X_i^{2D} + \varepsilon_i \text{ , with } X_i^{2D} = X_{it_2}^2 - X_{it_1}^2$$

- β is point identified if access to the individual panel micro data.

- What with aggregated data?

What with aggregated data?

$$\text{Min } E\left(Y_i^D - \beta_1 X_i^D - \beta_2 X_i^{2D}\right)^2$$

- $\beta_1 = \frac{E\left(X_i^D Y_i^D\right)E\left(X_i^{2D} X_i^{2D}\right) - E\left(X_i^D X_i^{2D}\right)E\left(Y_i^D X_i^{2D}\right)}{E\left(X_i^D X_i^D\right)E\left(X_i^{2D} X_i^{2D}\right) - \left[E\left(X_i^D X_i^{2D}\right)\right]^2}$

- $\beta_2 = \frac{E\left(X_i^{2D} Y_i^D\right)E\left(X_i^D X_i^D\right) - E\left(X_i^{2D} X_i^D\right)E\left(Y_i^D X_i^D\right)}{E\left(X_i^{2D} X_i^{2D}\right)E\left(X_i^D X_i^D\right) - \left[E\left(X_i^{2D} X_i^D\right)\right]^2}$

→ Points identified if all moments are known.

What with CompNet micro-aggregated data?

Some information is missing!

$$E\left(X_i^D X_i^D\right) = E\left(X_{it_2}^2\right) - 2E\left(X_{it_2} X_{it_1}\right) + E\left(X_{it_1}^2\right)$$

$$E\left(X_i^{2D} X_i^{2D}\right) = E\left(X_{it_2}^4\right) - 2E\left(X_{it_2}^2 X_{it_1}^2\right) + E\left(X_{it_1}^4\right)$$

$$E\left(Y_i^D X_i^D\right) = E\left(Y_{it_2} X_{it_2}\right) + E\left(Y_{it_2} X_{it_1}\right) - \; \ldots$$

➔ **Partial identification** is the best we can do.

Identification methodology (I)

Cambanis-Simons-Stout Inequality (1976):

(Fan, Sherman and Shum 2014)

Let S and T random variables with known marginal distribution functions F_S and F_T and finite variances. Then

$$\int_0^1 F_S^{-1}(1-u)F_T^{-1}(u)\mathrm{d}u \leq E(ST) \leq \int_0^1 F_S^{-1}(u)F_T^{-1}(1-u)\mathrm{d}u$$

The bounds are finite and, without additional information, sharp.

➔ We can find bounds for $E\left(Y_{it_2}X_{it_2}\right)$ using information on $F_{Y_{t_2}}$ and $F_{X_{t_2}}$.

Identification methodology (II)

- $$\beta_2 = \frac{E(X_i^{2D} Y_i^D) E(X_i^D X_i^D) - E(X_i^{2D} X_i^D) E(Y_i^D X_i^D)}{E(X_i^{2D} X_i^{2D}) E(X_i^D X_i^D) - [E(X_i^{2D} X_i^D)]^2}$$

$$= \frac{M_1 M_2 - M_3 M_4}{M_5 M_2 - [M_3]^2} = \mathcal{H}(M_1, M_2, M_3, M_4, M_5)$$

- $M_k^L \leq M_k \leq M_k^U$ for $k = 1, \ldots, 5$

Technical challenges

Hence $\beta_2^L \leq \beta_2 \leq \beta_2^U$ where the upper bound β_2^U is defined as:

$$\beta_2^U = \sup_{M_1,M_2,M_3,M_4,M_5} \mathcal{H}(M_1, M_2, M_3, M_4, M_5)$$

$$\text{s.t.}\, M_k^L \leq M_k \leq M_k^U \text{ for } k = 1, \ldots, 5$$

M_k^L and M_k^U are not exactly known but **estimated**

1. F_X, F_Y must be estimated using percentiles of the distribution available in CompNet

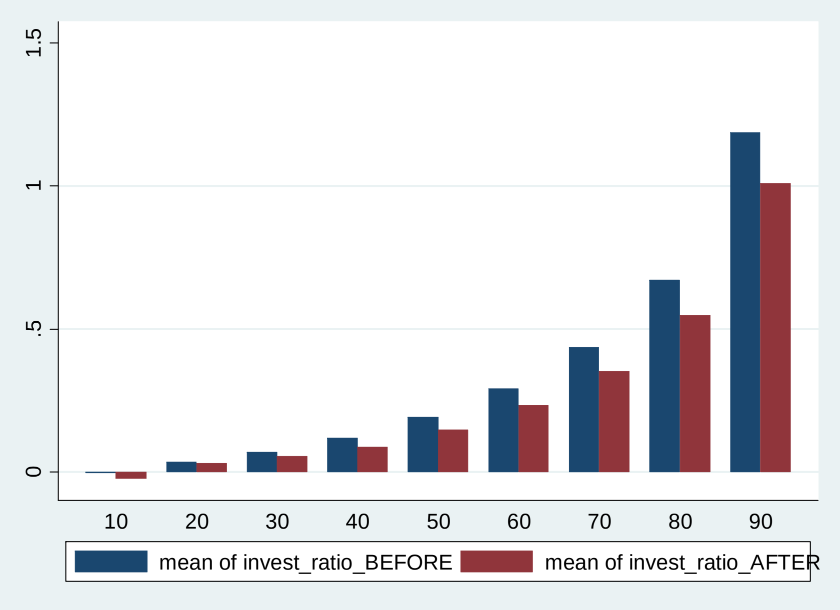
→ **interpolation** to get the complete distribution

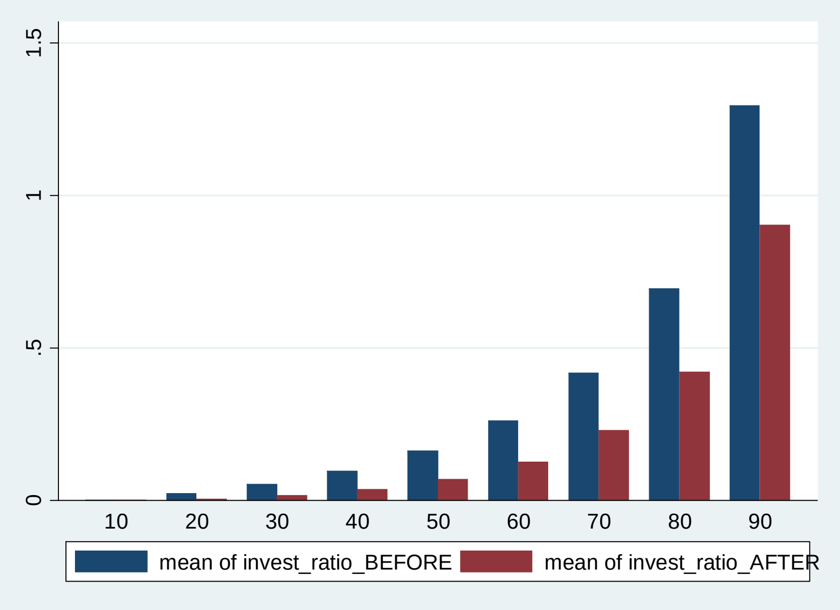
2. β_2^U might be imprecisely estimated near the estimated $\widehat{M}_k^L$ and $\widehat{M}_k^U$ bounds

→ **precision-correction** to control for sampling errors

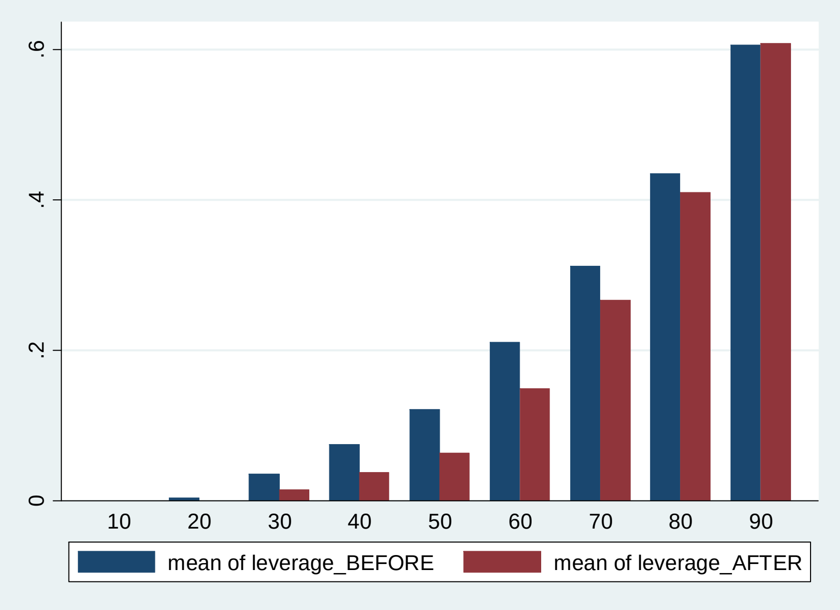
Investment ratio and leverage by percentiles

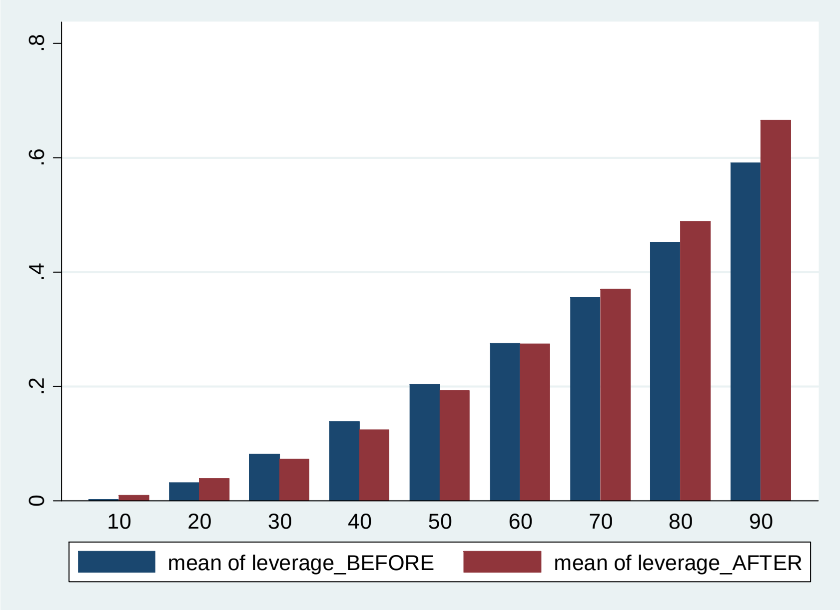
Non-stressed countries INVESTMENT Stressed countries





Non-stressed countries (excluding DE) LEVERAGE Stressed countries (excluding SI)





A new methodology: Multi-step implementation (I)

CompNet data provides aggregates at the country-sector-size cell level over several years (T>2).

Step 1: obtain bounds for each cell and each t=2,…,T

$$\beta_2^U = \sup_{M_1,M_2,M_3,M_4,M_5} \mathcal{H}(M_1, M_2, M_3, M_4, M_5)$$

$$\text{s.t.}\, M_k^L \leq M_k \leq M_k^U \text{ for } k = 1, \ldots, 5$$

Multi-step implementation (II)

Step 2: for each cell, intersect bounds obtained at each t

$$Y_{it_2}^{D} = \beta_1 X_{it_2}^{D} + \beta_2 X_{it_2}^{2D} + \varepsilon_{it_2} \rightarrow (\beta_2^L)^2 \leq \beta_2 \leq \left(\beta_2^U\right)^2$$
$$Y_{it_3}^{D} = \beta_1 X_{it_3}^{D} + \beta_2 X_{it_3}^{2D} + \varepsilon_{it_3} \rightarrow (\beta_2^L)^3 \leq \beta_2 \leq \left(\beta_2^U\right)^3$$
$$\ldots$$

$$\rightarrow \max_{\tau=2,\ldots,T} (\beta_2^L)^\tau \leq \beta_2 \leq \min_{\tau=2,\ldots,T} \left(\beta_2^U\right)^\tau$$

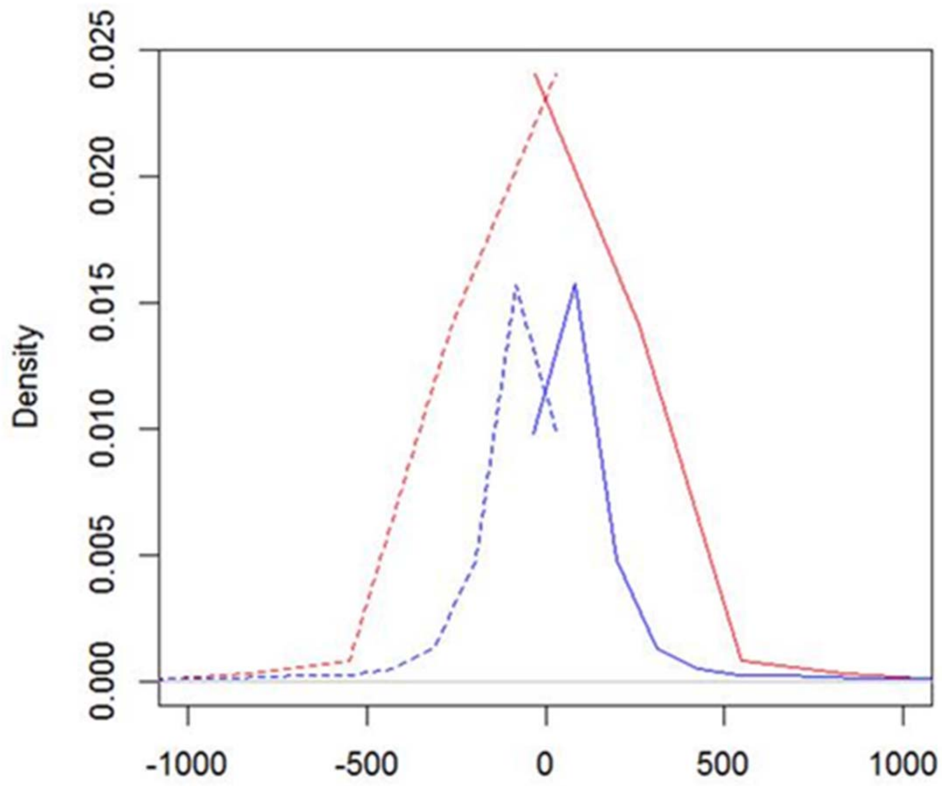
Step 3: intersect bounds obtained for each cell

$$(\beta_2^L)^c \leq \beta_2 \leq \left(\beta_2^U\right)^c, \; c = 1, \ldots, C$$

$$\rightarrow \max_{c=1,\ldots,C} (\beta_2^L)^c \leq \beta_2 \leq \min_{c=1,\ldots,C} \left(\beta_2^U\right)^c$$

Preliminary results on the bounds (I)

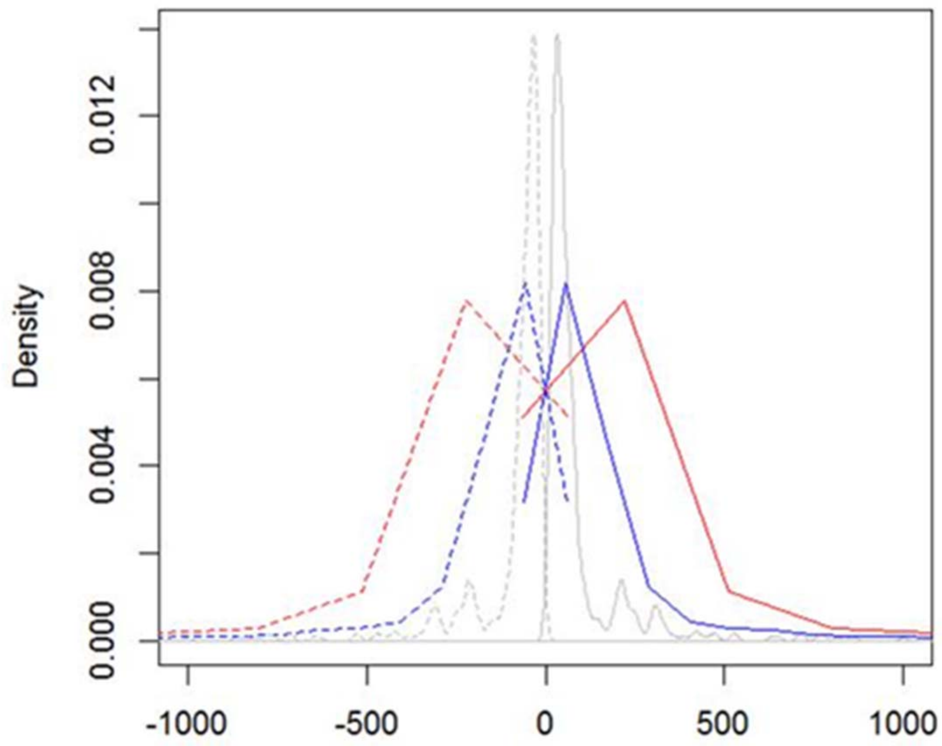
Step 1: obtain bounds for each cell and each t=2,…,T



Densities of both the upper (solid) and lower (dashed) bounds both before 2008 (in blue) and after 2008 (in red):

Preliminary results on the bounds (II)

Step 1: obtain bounds for each cell and each t=2,…,T



Densities of both the upper (solid) and lower (dashed) bounds for different countries: Belgium in blue, Germany in red and Italy in grey

Challenges and next steps

1. Estimating the bounds:

✓ done

Confirmation of our intuition for the method:

- by taking the intersection of the bounds across time and cells, we can sharpen the bounds.

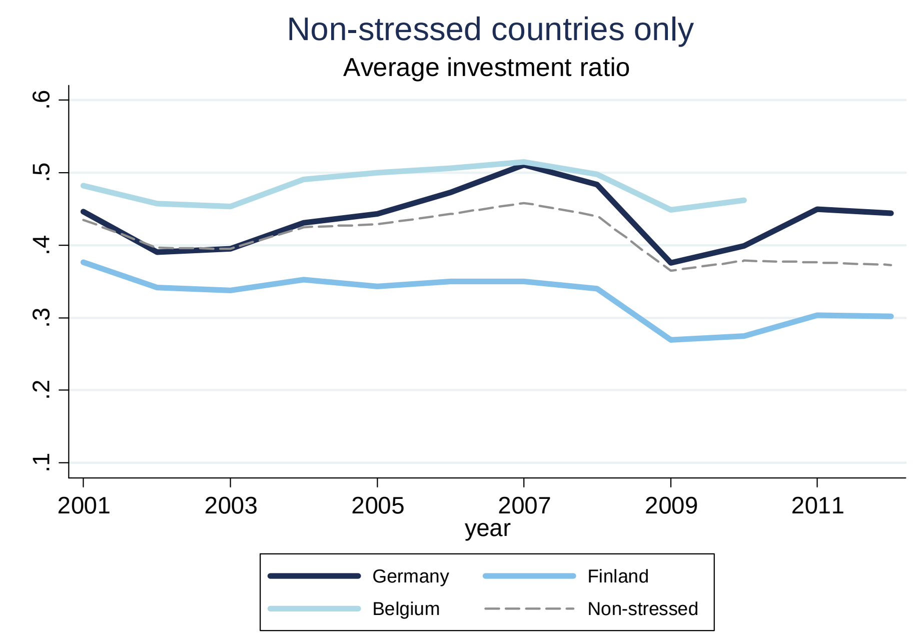
However, <u>for now</u>, the estimated bounds contain zero (the estimated lower bounds are always negative and the estimated upper bounds are always positive).

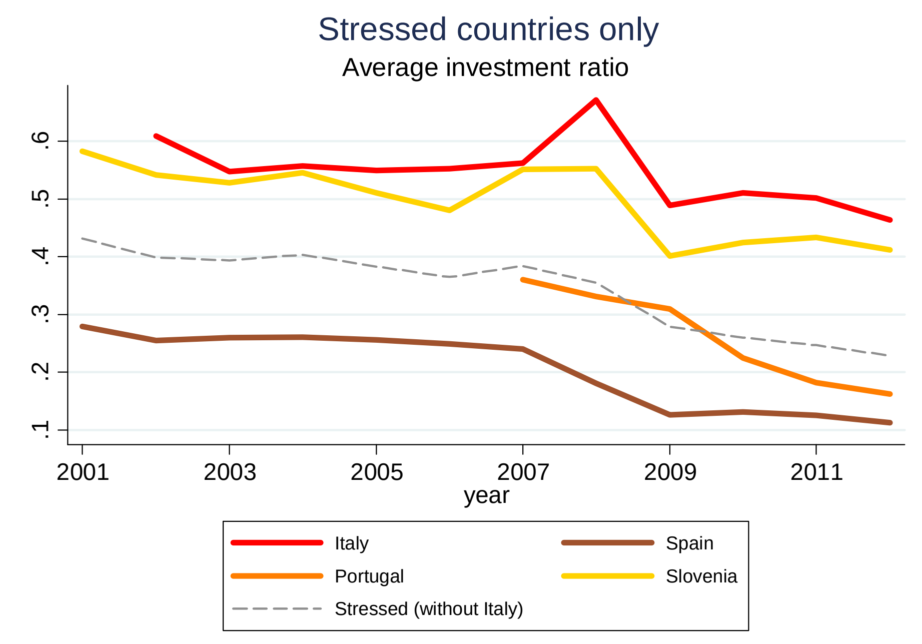
2. Next steps: Inference for the bounds

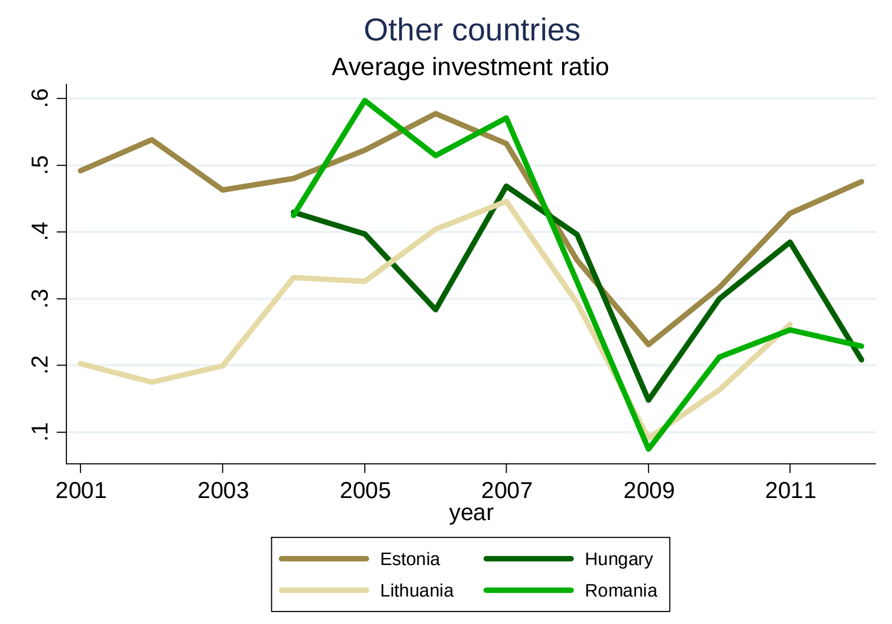
✓ Confidence interval: to be done

Determinants of corporate investment and role of leverage

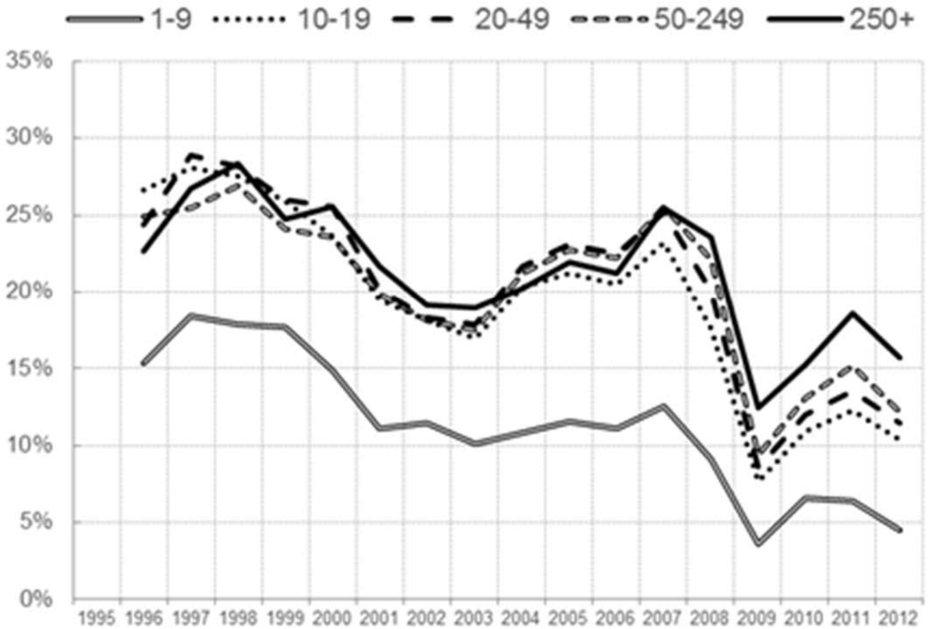
The 2009 investment collapse

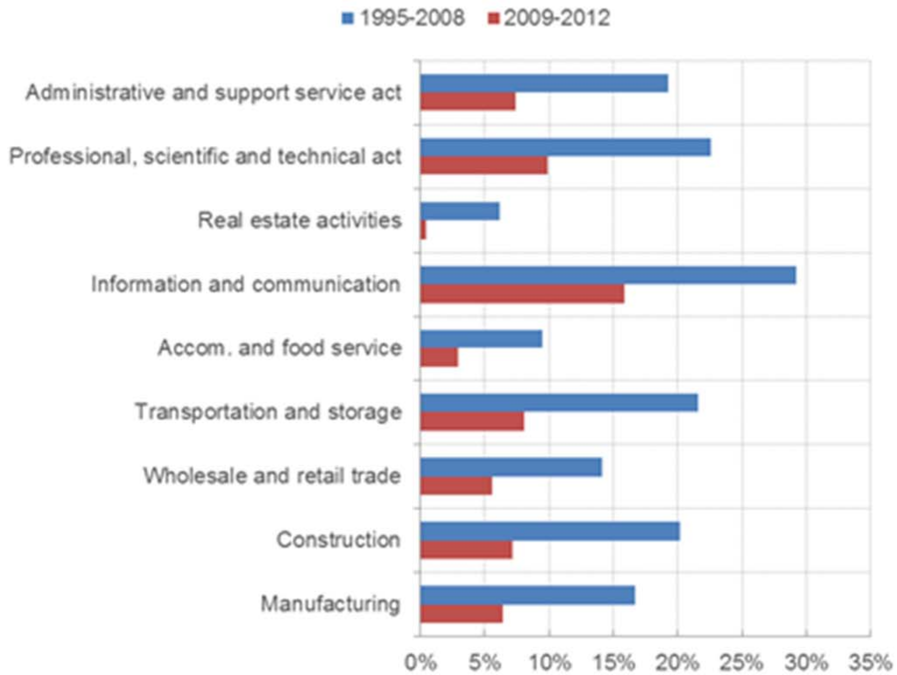






The 2009 investment collapse





Theory: fundamental vs. financial determinants

With perfect capital markets,

- Modigliani-Miller capital-structure irrelevance proposition.
- Tobin's q theory: present value of future marginal productivity of capital.

With capital markets imperfections (e.g. asymmetric information),

- Internal and external capital are not perfect substitutes.
- Liquidity and strength of balance sheet matter: dependence on external funds, external finance premium, collaterals
- Empirical literature on investment-cash flow sensitivity since Fazzari *et al*. (1988).

Fundamental	Financial
Sales growth	Liquidity Leverage
Cash flow	

Specific role of leverage

Financial accelerator (Bernanke-Gertler,1989) :

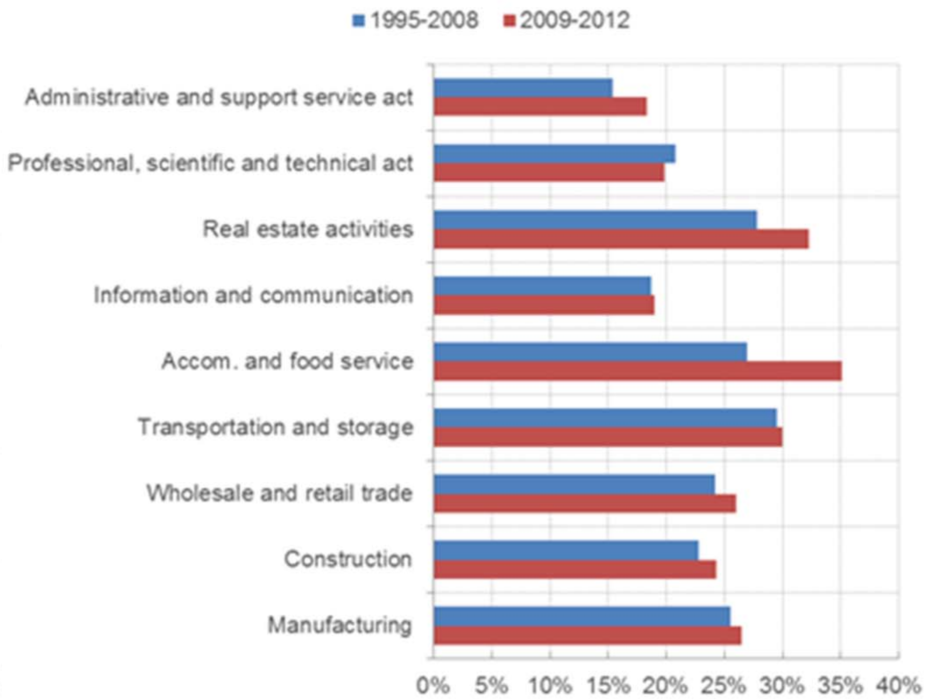
- Greater access to bank credit/ diversification of funding options
    - *boost productivity levels/reduction macro volatility*

- Excess indebtdeness can more than offset benefits
    - *raise corporate vulnerabilities/ amplify firms‘ sensitivity to income and interest shocks.*

- **Important asymmetric effects between investment decisions and balance sheet positions**
    - *(Cecchetti, Mohanty and Zampolli, 2010, Coricelli et al., 2010, Buca and Vermeulen, 2013, Goretti and Souto, 2013, Ferrando, Marchica and Mura, 2014, SIR 2015)*

- **Firms‘ high leverage is legacy of pre-crisis period** *(SIR 2013, Kalemli- Ozcan, Laeven and Moreno, 2015)*

Leverage across firm size and sectors



Analysis of Investment determinants in the financial module

$$\bar{\dot{IK}}_{ct} = \beta_1 \overline{IK}_{ct-1} + \beta_{21}\overline{SG}_{ct} + \beta_{22}\overline{SG}_{ct}\text{x } 1\{t \geq 2009\} + \beta_{31}\overline{CFK}_{ct-1} + \beta_{32}\overline{CFK}_{ct-1}\text{x } 1\{t \geq 2009\} + \beta_{41}\overline{Lev}_{ct-1} + \beta_{42}\overline{Lev}_{ct-1}\text{x } 1\{t \geq 2009\} + \bar{\mu}_c + \bar{v}_{ct}$$

	(A)	(B)
IK_{it-1}	0.0148	0.148*
SG_{it}	0.211*	0.292**
SG_{it} x $1\{t \geq 2009\}$	-0.109	-0.134
CFK_{it-1}	1.354*	0.833
CFK_{it-1} x $1\{t \geq 2009\}$	0.0546	-1.691***
Lev_{it-1}	-0.236	2.394**
Lev_{it-1} x $1\{t \geq 2009\}$	-0.132	1.873***
Lev^2_{it-1}		-0.0482**
Lev^2_{it-1} x $1\{t \geq 2009\}$		-0.0489***
Observations	1,049	1,049
Number of cells	157	157
Number of instruments	35	39
AR2(p-value)	0.878	0.579
H-test(p-value)	0.0182	0.953

Cell-based model for 4 countries (BE, DE, ES and IT) over the 2000-2012 period, 9 macro-sectors and 5 size classes.

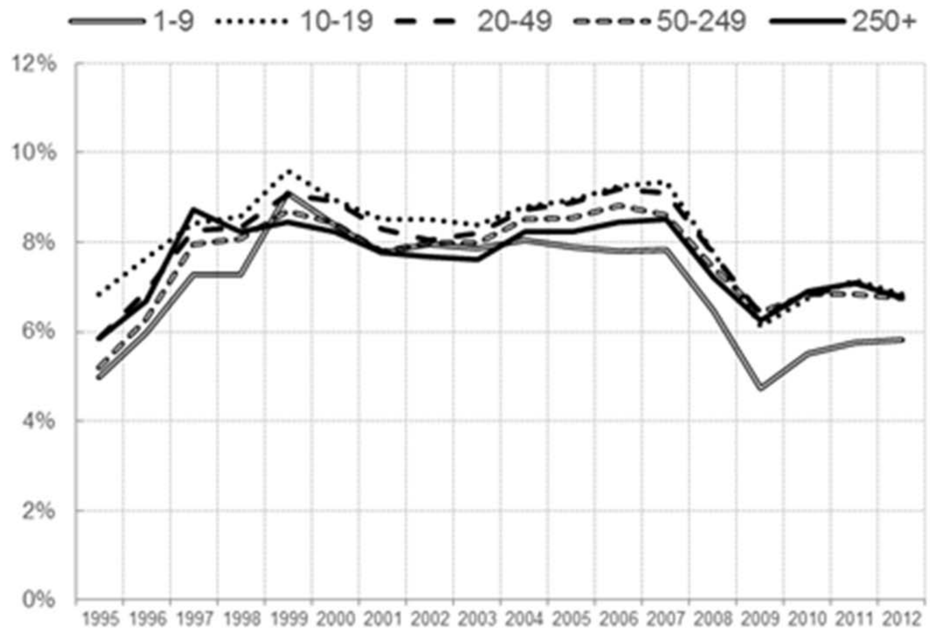
Conclusions

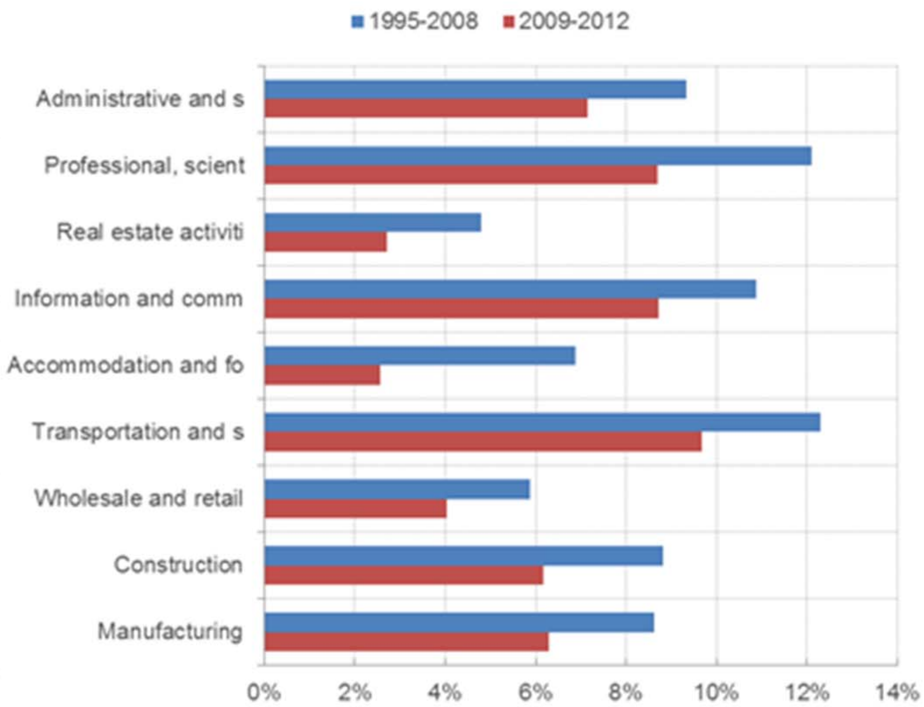
- Non-linearities at the firm level imply that the aggregate relationship differ from the individual relationship evaluated at the average.
- Points identified if all moments are known but this information is partly missing in micro-aggregated databases like CompNet
- **Partial identification** is the best we can do.
- **New methodology** to find bounds for $E\left(Y_{it_2}X_{it_2}\right)$ using information on $F_{Y_{t_2}}$ and $F_{X_{t_2}}$:
  1. We apply interpolation techniques using percentiles of the distribution available in CompNet
  2. We apply precision-correction techniques to control for sampling errors

- **We believe there is need to obtain more precision in the relationship between investment and leverage**

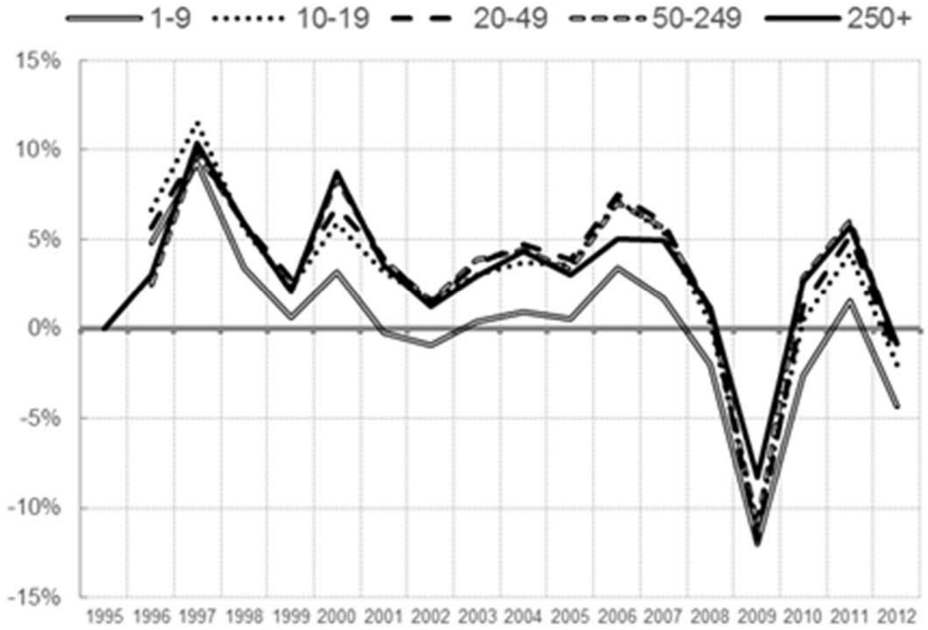
Thank you

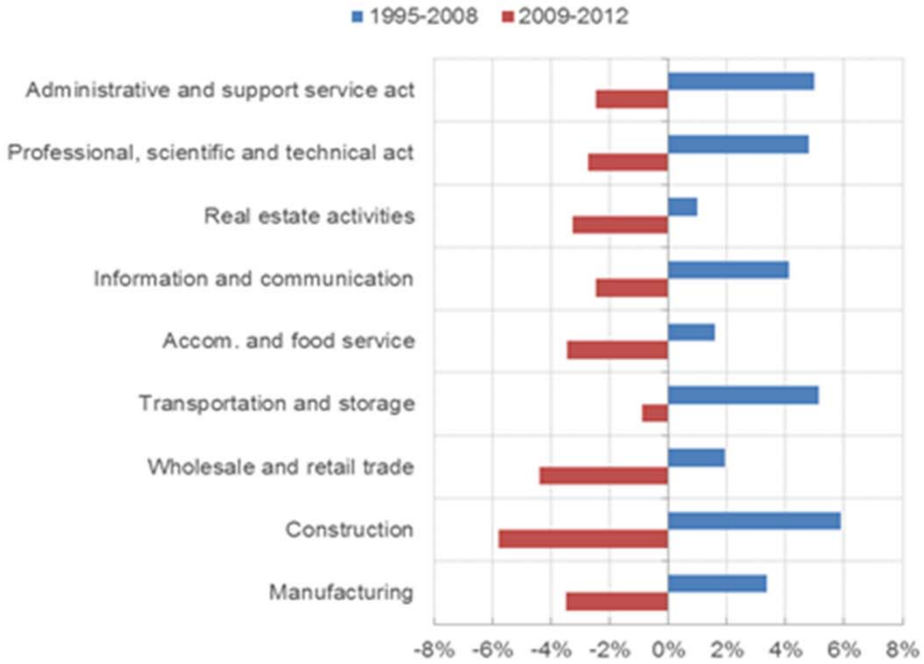
Explanatory variables: cash flow





Explanatory variables: sales growth





- Sales growth falls in 2009 and 2012.

Explanatory variables: Cash holding