Network Effects and Learning in Crowdfunding

by

Paul Belleflamme, Thomas Lambert, and
Armin Schwienbacher

Discussant:

Vesa Pursiainen

University of Hong Kong & Imperial College London

Research question

- How do network effects, combined with backer learning, affect reward-based crowdfunding platforms?

Why is this interesting?

- Network effects are very pervasive, particularly in the digital economy
  - Huge amount of policy debate about how to deal with platforms that become dominant due to network effects
- Crowdfunding provides a nice setup to empirically quantify network effects on a two-sided platform with a large number of users
- On the flipside, it's not clear that crowdfunding is the environment where network effects are most prominent

Methodology & findings

- Crowdfunding data from Ulule (and KKBB)
- Regression analysis of daily campaign contributions
  - Within project:
    - Current contributions increase with prior contributions
  - Cross-project:
    - Current contributions to a project increase with past contributions to other contemporaneous projects
- Regression analysis of campaign success likelihood and contribution timing
  - Recurrent backers more likely to back successful projects
  - Recurrent backers (may) contribute earlier in campaign?

What does a network effect mean?

- “The utility that a user derives from consumption of the good increases with the number of other agents consuming the good” (Katz and Shapiro, 1985 AER)
- In the case of a crowdfunding platform:
  - Backer-entrepreneur
  - Backer-backer
  - Entrepreneur-entrepreneur

Network effects estimation
Regression model

- Regression equation:

$$y_{ijt} = \alpha_i + \alpha_t + \beta_1 Y_{i,t-1} + \beta_2 Y_{-i,t-1} + \beta_3 Y_{-j,t-1} + \boldsymbol{\gamma X}_{i,t-1} + \varepsilon_{it}$$

"Auto-correlation" of daily contributions

"Correlation" of daily contributions between projects within category

"Correlation" of daily contributions between project categories

Network effects estimation
Baseline results

	(1)	(2)	(3)	(4)
# $\text{contributions}_{i,t-1}$	0.185***			0.183***
	(0.002)			(0.002)
# $\text{contributions}_{-i,t-1}$		0.027***		0.013***
		(0.002)		(0.002)
# $\text{contributions}_{-j,t-1}$			0.075***	0.047***
			(0.003)	(0.003)
# $\text{projects}_{i,t}$	0.001	-0.030***	-0.024***	-0.029***
	(0.007)	(0.009)	(0.008)	(0.007)
% goal_t	0.286***	0.369***	0.368***	0.284***
	(0.006)	(0.007)	(0.007)	(0.006)
Popular_t	1.161***	1.252***	1.253***	1.163***
	(0.010)	(0.012)	(0.012)	(0.010)
% recurrent backers_t	0.662***	0.675***	0.674***	0.661***
	(0.002)	(0.003)	(0.003)	(0.002)
Project Fixed Effects	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Day of week Fixed Effects	Yes	Yes	Yes	Yes
Funding cycle day Fixed Effects	Yes	Yes	Yes	Yes
# observations	814,960	814,960	814,960	814,960
# projects	23,022	23,022	23,022	23,022
R^2	0.548	0.529	0.529	0.548

More backers in t-1 is ***positively*** associated with contributions in t

More projects in t is ***negatively*** associated with contributions in t – this result is important but not really discussed in paper

Network effects estimation
Model specification

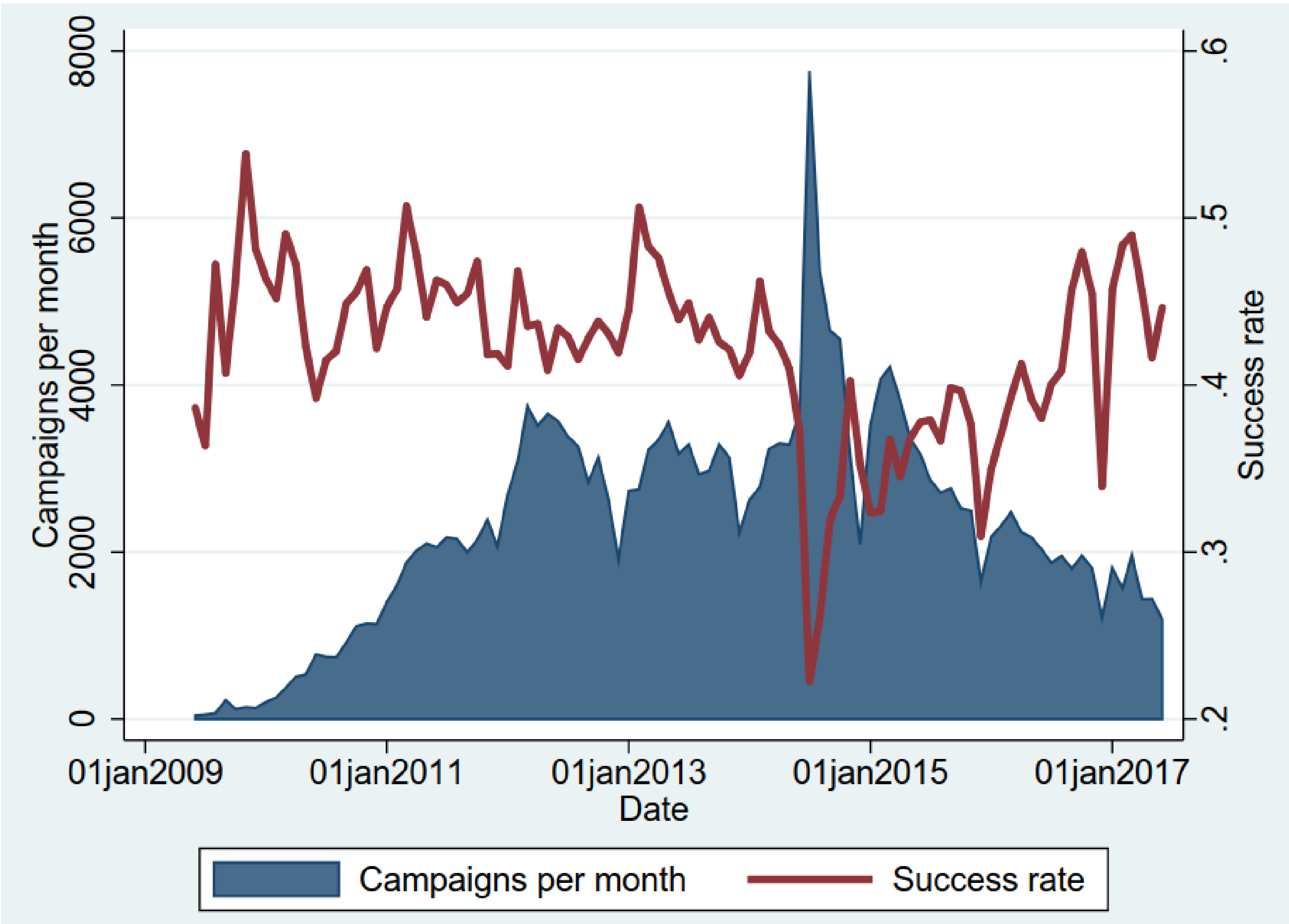
- Regression methodology assumes a linear functional form for all network effects:

$$y_{ijt} = \alpha_i + \alpha_t + \beta_1 Y_{i,t-1} + \beta_2 Y_{-i,t-1} + \beta_3 Y_{-j,t-1} + \boldsymbol{\gamma} \boldsymbol{X}_{i,t-1} + \varepsilon_{it}$$

- This seems to miss some important nuance, e.g.:
  - ***Inter-project***: Positive and negative inter-project effects may dominate at different times (e.g. platform liquidity constraints?)
  - ***Intra-project***: Pledging is (Mollick, 2014 JBV) and should be (Strausz, 2017 AER) highly conditional on current funding status

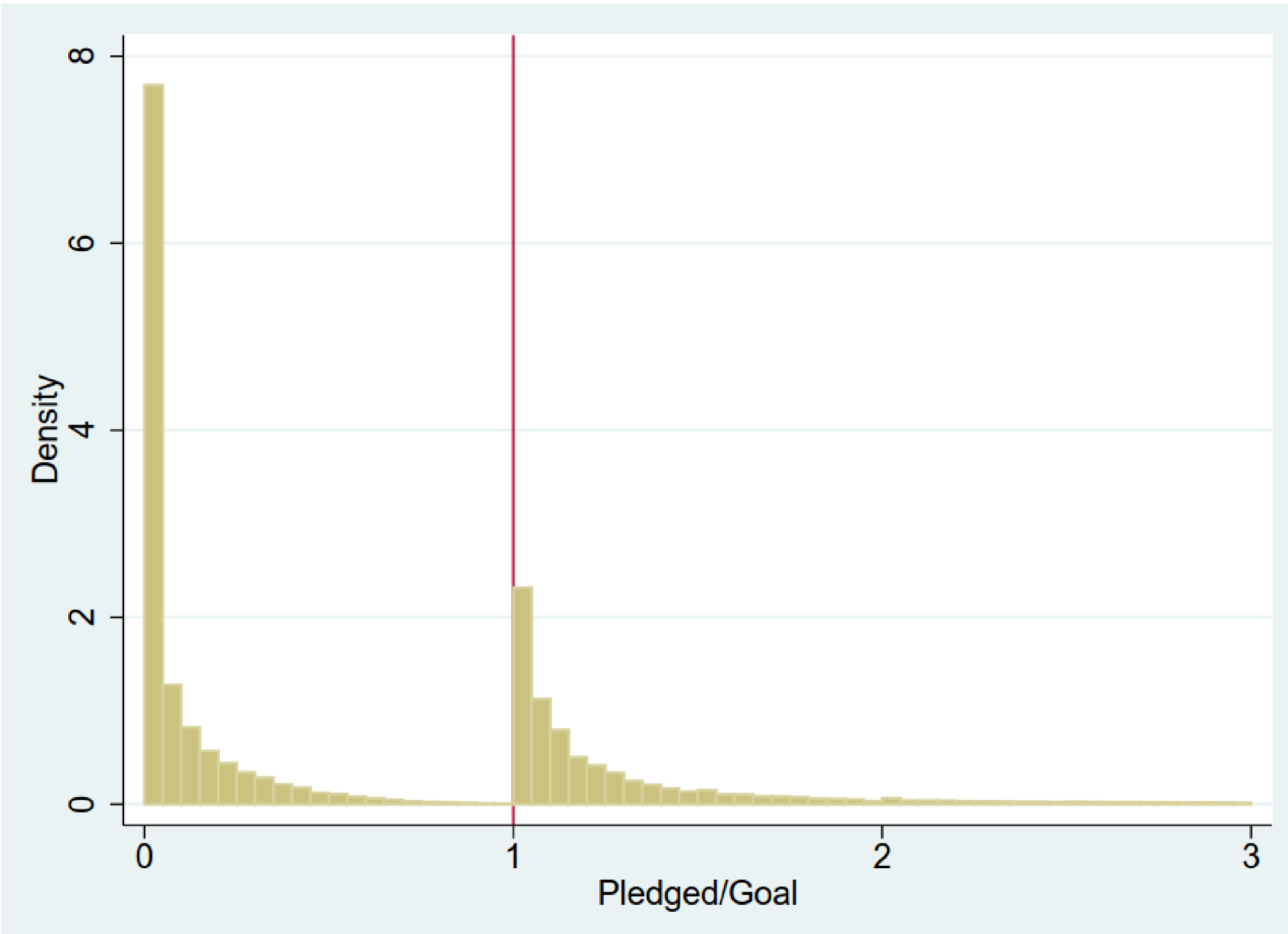
Network effects estimation

Inter-project dynamics (Kickstarter)



- There seem to be limits to platform liquidity
- Inter-project network effects are likely to depend on the ratio of projects to backers

Network effects estimation
Intra-project dynamics (Kickstarter)



- Effect of past contributions depends on the current pledged/goal ratio
- Linear control does not capture this
- More flexible functional form might provide a better fit

Network effects estimation
Control variables

	(1)	(2)	(3)	(4)
# contributions$_{i,t-1}$	0.185***			0.183***
	(0.002)			(0.002)
# contributions$_{-i,t-1}$		0.027***		0.013***
		(0.002)		(0.002)
# contributions$_{-j,t-1}$			0.075***	0.047***
			(0.003)	(0.003)
# projects$_{i,t}$	0.001	-0.030***	- 0.024***	- 0.029***
	(0.007)	(0.009)	(0.008)	(0.007)
% goal$_t$	0.286***	0.369***	0.368***	0.284***
	(0.006)	(0.007)	(0.007)	(0.006)
Popular$_t$	1.161***	1.252***	1.253***	1.163***
	(0.010)	(0.012)	(0.012)	(0.010)
% recurrent backers$_t$	0.662***	0.675***	0.674***	0.661***
	(0.002)	(0.003)	(0.003)	(0.002)
Project Fixed Effects	Yes	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Day of week Fixed Effects	Yes	Yes	Yes	Yes
Funding cycle day Fixed Effects	Yes	Yes	Yes	Yes
# observations	814,960	814,960	814,960	814,960
# projects	23,022	23,022	23,022	23,022
R^2	0.548	0.529	0.529	0.548

- A more flexible functional form would make sense (range FE?)
- Should probably be calculated at t-1, as it now seems to include the current contribution (LHS variable)

On the fixed effects:

- Projects last about one month, so project FE are pretty close to including year-month FE – not clear if the year and month FE do much here
- The specification would allow day FE as well

Network effects estimation
Causality?

- Do contributions to a project *cause* more contributions to
  - The same project?
  - Other projects?
- How would this happen?
  - Within project:
    - Likelihood of completion?
    - Information or herding effect (e.g., Astebro et al, 2018)?
  - Cross-project:
    - Larger pool of backers reviewing projects makes matching more likely?
- But, the results could also be caused by variation in participation due to omitted variables
  - In some sense, the high frequency of the data (daily) makes this concern worse, as last day’s volume may measure short-term fluctuations instead of “scale” of the platform

Network effects estimation
Diff-in-Diff analysis

- The authors exploit unexpected fast campaign starts as “exogenous” shocks to campaign contributions
- This does not completely remove the concern that increased participation is driven by some omitted variable, which also makes fast starts more likely
- However, it’s of course better than pure correlation

Network effects estimation
Diff-in-Diff analysis

	>200	
	(1)	(2)
Fast $start_t$	0.013***	
	(0.004)	
Fast $start_{j,t}$ [1]		0.038**
		(0.015)
Fast $start_{-j,t}$ [2]		0.011**
		(0.004)
p-value [1] = [2]		[0.0937]
Controls	Yes	Yes
Project Fixed Effects	Yes	Yes
Month Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
Day of week Fixed Effects	Yes	Yes
Funding cycle day Fixed Effects	Yes	Yes
# observations	813,983	813,983
# projects	22,995	22,995
R^2	0.518	0.518

- DiD requires a treatment group and control group, and that these groups would look similar in the absence of treatment
- In this case, the shock is over time – it's not clear if there is a control group that is not affected
- Column 2 looks like DiD
  - But treatment is not randomly assigned – are the treatment and control groups similar?
  - Presentationally cleaner to present column 2 as interaction (Fast start x Same category)?

Network effects estimation
The role of recurrent backers

- *"Our analyses establish that recurrent backers act as the main transmission channel of cross-project funding dynamics"*
- The analysis may not quite justify this statement
  - Table 6 Panel A suggests the opposite
  - Even in Panel B (fast starts) it's not clear there is significant difference between new and recurrent backers

Learning analysis
Likelihood of campaign success

	Success$_i$ (Ulule)			
	(1)	(2)	(3)	(4)
Recurrent backer$_i$	0.028***			0.029***
	(0.002)			(0.002)
Recurrent backer$_{-i}$		0.004***		0.003***
		(0.001)		(0.001)
Recurrent backer$_{-j}$			0.007***	0.009***
			(0.001)	(0.001)
Age	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
€-value first contribution	0.017***	0.017***	0.017***	0.017***
	(0.000)	(0.000)	(0.000)	(0.000)
Campaign duration	-0.053***	-0.053***	-0.053***	-0.053***
	(0.001)	(0.001)	(0.001)	(0.001)
Cash contribution	0.063***	0.062***	0.062***	0.063***
	(0.001)	(0.001)	(0.001)	(0.001)
Country of residence Fixed Effects	Yes	Yes	Yes	Yes
Category Fixed Effects	Yes	Yes	Yes	Yes
Day Fixed Effects	Yes	Yes	Yes	Yes

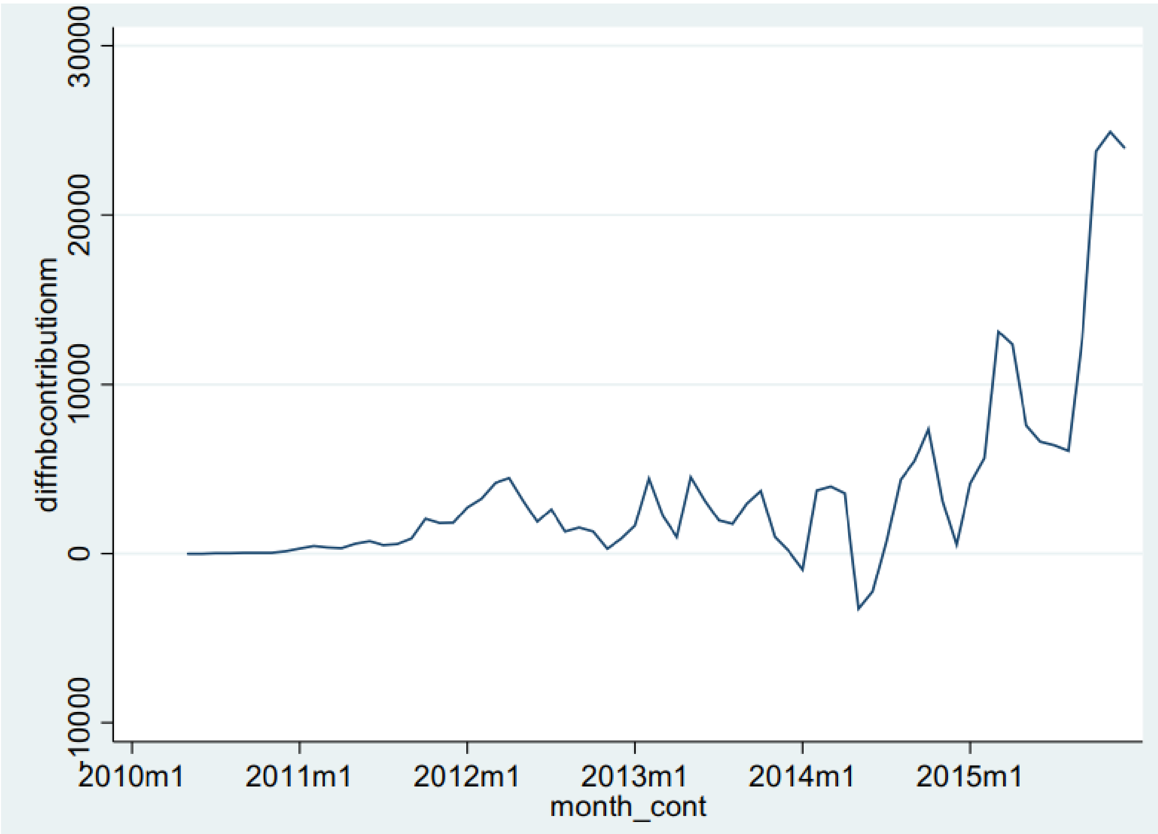
- The unit of observation is contribution
- Some contributions are for projects which are already successful (amount pledged > goal), so there is no uncertainty at all?
- What if you only look at the first contributions by the people who then become recurrent backers? Maybe they are just inherently different?

Learning analysis
Timing of contribution

	Timing$_i$ (Ulule)			
	(1)	(2)	(3)	(4)
Recurrent backer$_i$	0.180***			0.179***
	(0.002)			(0.002)
Recurrent backer$_{-i}$		-0.025***		-0.008***
		(0.001)		(0.001)
Recurrent backer$_{-j}$			0.002**	0.008***
			(0.001)	(0.001)
Age	-0.001***	-0.000***	-0.000***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
€-value first contribution	0.005***	0.008***	0.008***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)
Campaign duration	0.014***	0.016***	0.017***	0.014***
	(0.001)	(0.001)	(0.001)	(0.001)
Cash contribution	0.032***	0.030***	0.030***	0.032***
	(0.001)	(0.001)	(0.001)	(0.001)
Country of residence Fixed Effects	Yes	Yes	Yes	Yes
Category Fixed Effects	Yes	Yes	Yes	Yes
Day Fixed Effects	Yes	Yes	Yes	Yes
# observations	1,302,899	1,302,899	1,302,899	1,302,899
R^2	0.098	0.081	0.081	0.098

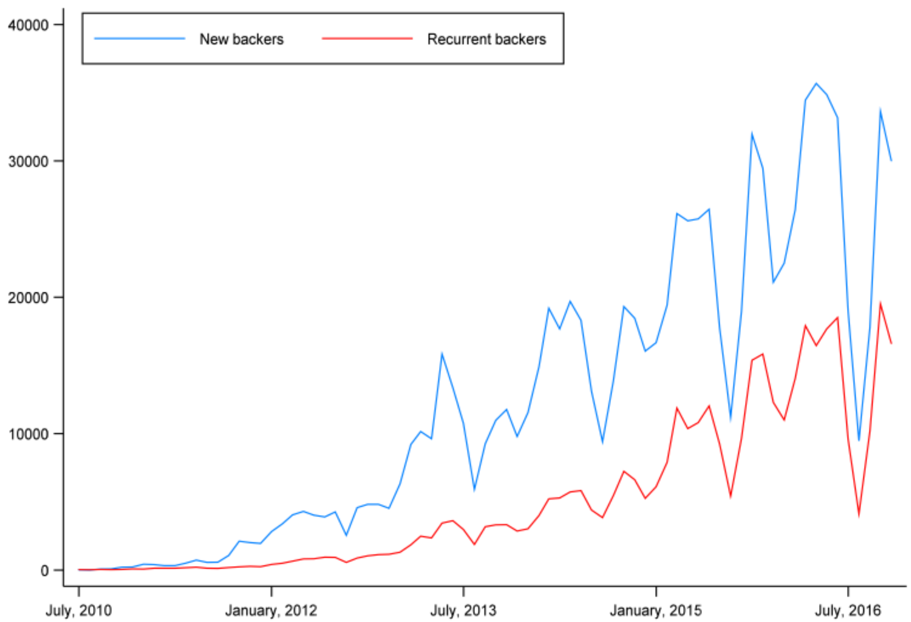
- The signs are different across the different categories – not clear what the conclusion is
- “[recurrent backers] are more likely to contribute at earlier stages of the campaign than other backers.”
- “..explain why recurrent backers exert a significant influence on later backers.”
- Not sure these statements accurately reflect the results

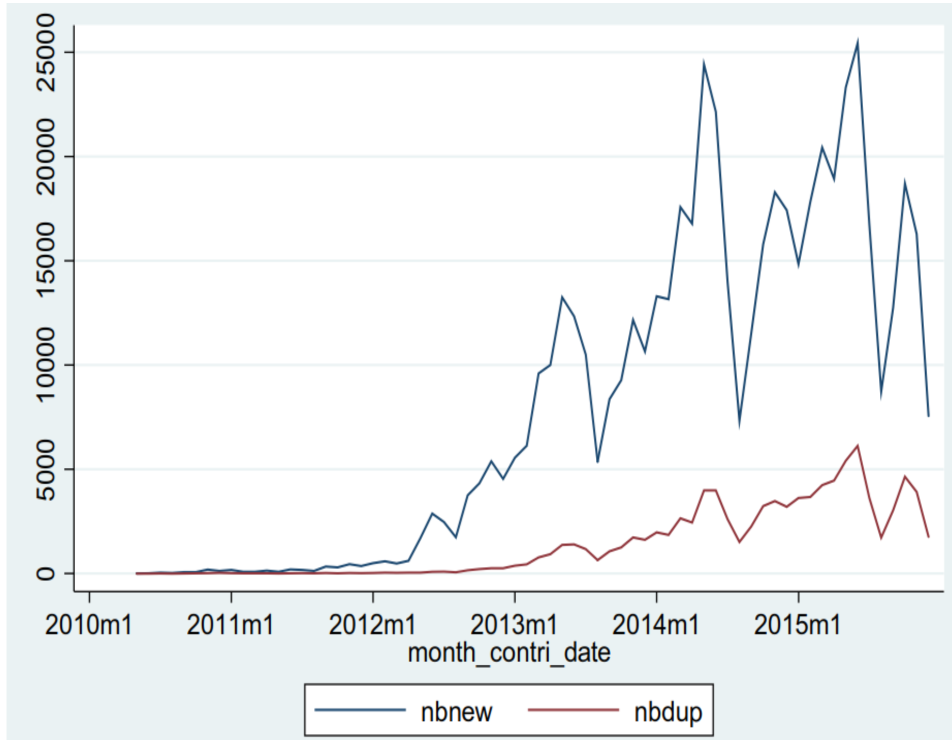
Winner takes all?
Interpretation of results



- "Our results suggest that reward-based crowdfunding is a 'winner-takes-all' type of market"
- "We take as evidence the widening gap between Ulule and KKBB"

Winner takes all?
Interpretation of results





Random thoughts

- How about entrepreneur learning?
  - Do entrepreneurs get better at getting funded (this happens on Kickstarter)?
  - Do entrepreneurs time their projects considering backer activity and competition from other projects? (this hasn't been studied as far as I'm aware)
  - The latter question might generate important prescriptions for entrepreneurs looking to fund projects
- Daily frequency for identifying network effects seems very high – it would be interesting to show some analysis at lower frequencies
  - Or alternatively, have the network effect variables of interest as rolling averages over, e.g., last month

Conclusion

- Interesting paper on an important topic
- Great data and interesting empirical analysis
- Still room to extend in several directions
- The story and results need some tidying up
  (which is not surprising given the version I read did not even include discussion of the new results yet)
- Good luck!

Appendix

Small comments

- Why not run some version of network analysis using a project level success dummy as the LHS variable, instead of daily contributions?
  - More direct measure of entrepreneur utility
- There is not much analysis on the "interplay" between network effects and learning in the current version, even though it's stated as the main research question of the paper

Small comments

- The tone in the text is pretty strong
  - E.g. “we *show* that current backers’ contributions to a particular project are positively *influenced* by previous backers’ contributions to that project”
  - This empirical result is basically a simple correlation, which makes the statement a bit aggressive in terms of claiming causality
- I didn’t understand this: “we estimate a dynamic panel model using the standard within estimator as well as a moment-based estimator with better asymptotic properties” – didn’t see any explanation elsewhere
- The paper talks about the importance for platforms to manage project mix but does not provide the prescriptions for what/how they should do that

Small comments

- In Tables 7 and 8, could probably control for more campaign and entrepreneur features (size, entrepreneur experience and characteristics, more accurate location etc.) and add controls/analysis on backer-entrepreneur pairs
- Typo on “KissKissBanBank” on front page
- Typos and outdated versions in the list of references