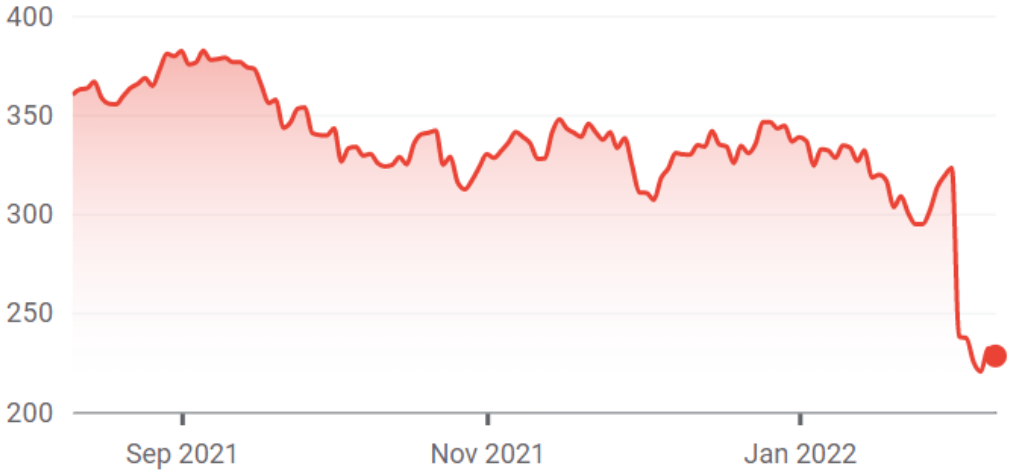
Blockchain without Crypto? Linking On-Chain Data Growth to Firm Fundamentals and Stock Returns

Ran Chang Will Cong

Discussion by Greg Buchak
Stanford GSB

April 5, 2022

Motivation: Facebook (Meta) Stock Price



Motivation: Facebook (Meta) Stock Price

TECH

Facebook says Apple iOS privacy change will result in $10 billion revenue hit this year

PUBLISHED WED, FEB 2 2022•7:54 PM EST | UPDATED THU, FEB 3 2022•11:06 AM EST

Motivation

Data increasingly seen as a valuable asset

Motivation

Data increasingly seen as a valuable asset

Growing literature on how firms use data to their advantage

- Production/management decisions (Jones & Tonetti (2020); Farboodi et al. (2021))
- Screening/marketing (He et al. (2021); Babina et al. (2022))

Motivation

Data increasingly seen as a valuable asset

Growing literature on how firms use data to their advantage

- Production/management decisions (Jones & Tonetti (2020); Farboodi et al. (2021))
- Screening/marketing (He et al. (2021); Babina et al. (2022))

Preceding use cases are "profit" focused

- Informative about customers/market
- Relevant for increasing firm revenues/reducing firm costs
- Key: data the firm uses to improve performance

Motivation

Data increasingly seen as a valuable asset

Growing literature on how firms use data to their advantage

- Production/management decisions (Jones & Tonetti (2020); Farboodi et al. (2021))
- Screening/marketing (He et al. (2021); Babina et al. (2022))

Preceding use cases are "profit" focused

- Informative about customers/market
- Relevant for increasing firm revenues/reducing firm costs
- Key: data the firm uses to improve performance

Can data credibly inform investors about firm performance/management?

Motivation

Basic idea: data *about* the firm can alleviate classic corporate finance frictions:

- Information asymmetries, i.e., inform investors about state of the world
- Incentive problems (e.g., harder to expropriate/take bad actions)

Motivation

Basic idea: data *about* the firm can alleviate classic corporate finance frictions:

- Information asymmetries, i.e., inform investors about state of the world
- Incentive problems (e.g., harder to expropriate/take bad actions)

Technological problem: how to make firm-generated data credible?

- Firm can easily lie/change data ex-post
- Old-school technology: external audit
- New technology: record the data on the blockchain

Motivation

Basic idea: data *about* the firm can alleviate classic corporate finance frictions:

- Information asymmetries, i.e., inform investors about state of the world
- Incentive problems (e.g., harder to expropriate/take bad actions)

Technological problem: how to make firm-generated data credible?

- Firm can easily lie/change data ex-post
- Old-school technology: external audit
- New technology: record the data on the blockchain

How does this solve the problem?

- Firm can't edit results ex-post
- Firm can't falsify transaction records

This paper: Looking for this in the data

Does firm blockchain adoption...

1. Predict (cause?) firm fundamentals?
2. Show up in asset prices?

This paper: Looking for this in the data

Does firm blockchain adoption...

1. Predict (cause?) firm fundamentals?
2. Show up in asset prices?

Setting: Data on blockchain use for Chinese firms for data
Blockchain allows real-time, credible information reporting

This paper: Looking for this in the data

Does firm blockchain adoption...

1. Predict (cause?) firm fundamentals?
2. Show up in asset prices?

Setting: Data on blockchain use for Chinese firms for data
Blockchain allows real-time, credible information reporting

Main findings: More blockchain use for data needs $\rightarrow$ better firm performance

- Better fundamentals (asset growth, sales growth, ROA, etc.)
- Better returns (standard long-short XS approach)

This paper: Looking for this in the data

Does firm blockchain adoption...

1. Predict (cause?) firm fundamentals?
2. Show up in asset prices?

Setting: Data on blockchain use for Chinese firms for data
Blockchain allows real-time, credible information reporting

Main findings: More blockchain use for data needs $\rightarrow$ better firm performance

- Better fundamentals (asset growth, sales growth, ROA, etc.)
- Better returns (standard long-short XS approach)

In support of credible disclosure mechanism:
Stronger results for small firms, opaque firms

This discussion

1. **Summary of setting and data**

2. **Predictive results**
   - I'm entirely convinced by predictive relationships
   - Extremely thorough analysis
   - TONS of results and robustness checks

3. **Interpreting the results as causal**
   - More work to do on causal claims
   - Some questions and suggestions

Data

Key metric: Blockchain data growth (BDG_{iq})

- Firm i quarter q level
- (Log) 4-quarter change in amount of blockchain data

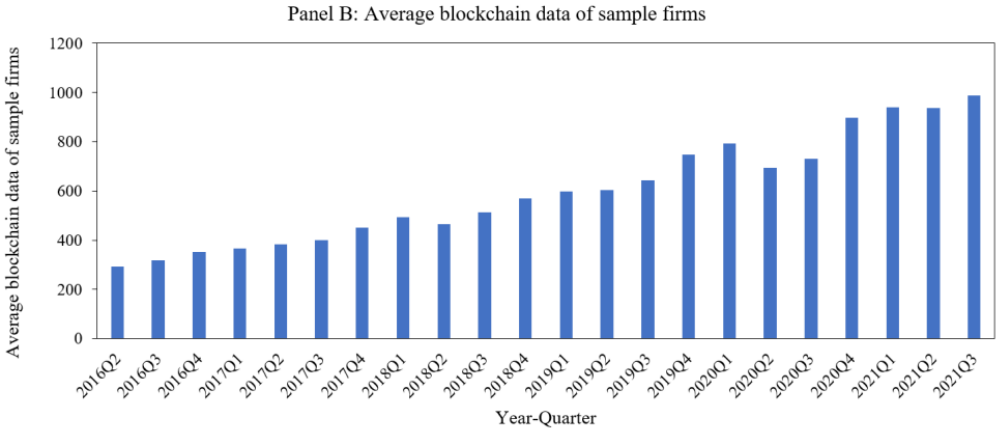
Real-time data: Blockchain data is available to essentially in real time

What is on the blockchain? Exploit this more?

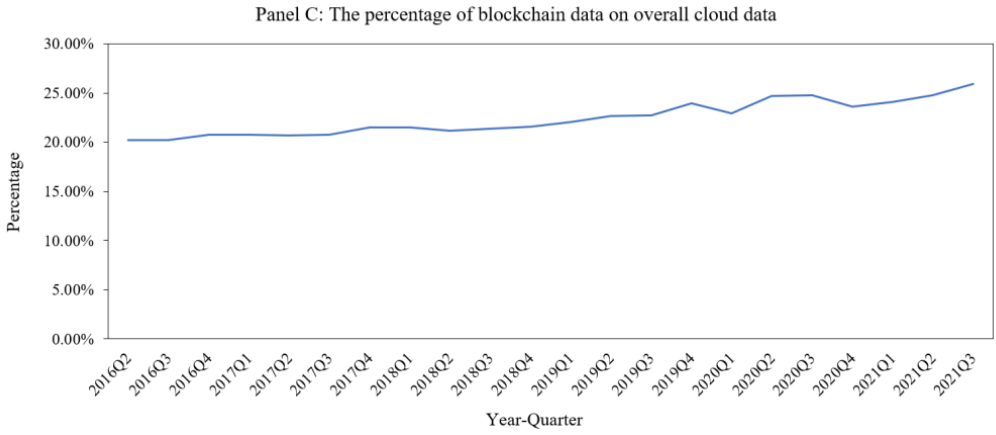
- Accounting data? (i.e., '10-K in blockchain form')
- Sales logs? (i.e., ledger of all firm transactions)
- Data relevant for firm operation (e.g., sales leads/R&D datasets)

What's stored on the blockchain will be highly relevant for interpretation

Data


Panel B: Average blockchain data of sample firms
Average blockchain data of sample firms
1200
1000
800
600
400
200
0
2016Q2
2016Q3
2016Q4
2017Q1
2017Q2
2017Q3
2017Q4
2018Q1
2018Q2
2018Q3
2018Q4
2019Q1
2019Q2
2019Q3
2019Q4
2020Q1
2020Q2
2020Q3
2020Q4
2021Q1
2021Q2
2021Q3
Year-Quarter

Data


Panel C: The percentage of blockchain data on overall cloud data
30.00%
25.00%
20.00%
15.00%
10.00%
5.00%
0.00%
Percentage
2016Q2
2016Q3
2016Q4
2017Q1
2017Q2
2017Q3
2017Q4
2018Q1
2018Q2
2018Q3
2018Q4
2019Q1
2019Q2
2019Q3
2019Q4
2020Q1
2020Q2
2020Q3
2020Q4
2021Q1
2021Q2
2021Q3
Year-Quarter

Predictive results

	ROA_{q+1}	ROA_{q+2}	AG_{q+1}	AG_{q+2}	SG_{q+1}	SG_{q+2}	PA_{q+1}	PA_{q+2}	PG_{q+1}	PG_{q+2}
BDG_{q+1}	0.766***	0.443***	0.361***	0.219***	0.057***	0.049***	0.263***	0.209***	0.179***	0.129***
	(5.26)	(3.52)	(4.28)	(2.91)	(4.27)	(2.86)	(3.84)	(2.99)	(4.69)	(3.00)
					...					
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year-Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	11266	11147	10807	10692	11037	10920	8047	7962	8047	7962
Adj. R2	0.58	0.54	0.49	0.41	0.35	0.34	0.25	0.21	0.17	0.13

Predictive results

Panel A: nowcasting and forecasting firm fundamentals with interaction terms

	ROA_{q+1}	ROA_{q+2}	AG_{q+1}	AG_{q+2}	SG_{q+1}	SG_{q+2}	PA_{q+1}	PA_{q+2}	PG_{q+1}	PG_{q+2}
BDG_{q+1}	0.152***	0.101**	0.085**	0.052*	0.015**	0.013*	0.055**	0.058*	0.050***	0.028*
	(3.15)	(2.10)	(2.43)	(1.77)	(2.51)	(1.85)	(2.26)	(1.74)	(2.73)	(1.88)
BDG_{q+1} * DSmall	0.185***	0.126**	0.067***	0.087***	0.021***	0.013*	0.096***	0.046**	0.042***	0.041*
	(3.14)	(2.47)	(4.02)	(2.66)	(3.31)	(1.75)	(3.58)	(2.33)	(3.27)	(1.91)
BDG_{q+1} * DLowIO	0.209***	0.149*	0.070***	0.051**	0.013**	0.012*	0.078***	0.048**	0.028**	0.035*
	(3.56)	(1.84)	(3.11)	(2.21)	(2.43)	(1.88)	(3.30)	(2.39)	(2.32)	(1.96)
BDG_{q+1} * DLowCov	0.188***	0.153**	0.108***	0.047**	0.017***	0.013	0.082***	0.046**	0.048***	0.029**
	(3.99)	(2.01)	(3.48)	(2.41)	(3.07)	(1.60)	(3.09)	(2.05)	(2.65)	(2.01)
BDG_{q+1} * DPrivate	0.179***	0.124**	0.090***	0.051**	0.018***	0.017**	0.074***	0.074*	0.048***	0.034*
	(4.13)	(1.98)	(3.83)	(2.43)	(3.25)	(2.04)	(3.49)	(1.78)	(3.09)	(1.86)
BDG_{q+1} * DLowHHI	0.250***	0.169**	0.079***	0.060**	0.015***	0.014**	0.090***	0.057***	0.033***	0.042**
	(4.15)	(2.10)	(3.58)	(2.51)	(2.81)	(2.12)	(3.86)	(2.87)	(2.80)	(2.33)
Dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year-Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	11266	11147	10807	10692	11037	10920	8047	7962	8047	7962
Adj. R2	0.69	0.65	0.60	0.49	0.43	0.42	0.30	0.25	0.20	0.16

Predictive results

Panel A: Nowcasting and forecasting firm fundamentals after controlling digital economy variables

	ROA_{q+1}	ROA_{q+2}	AG_{q+1}	AG_{q+2}	SG_{q+1}	SG_{q+2}	PA_{q+1}	PA_{q+2}	PG_{q+1}	PG_{q+2}
BDG_{q+1}	0.424***	0.340**	0.144***	0.125**	0.041**	0.030**	0.198**	0.146**	0.089***	0.066*
	(2.61)	(2.48)	(2.67)	(2.26)	(2.29)	(2.23)	(2.22)	(2.39)	(2.74)	(1.93)
$IoTG_{q+1}$	0.281	0.212	0.377**	0.315*	0.402**	0.302	0.362*	0.271	0.420*	0.326
	(1.55)	(1.24)	(2.16)	(1.67)	(2.21)	(1.65)	(1.95)	(1.46)	(1.83)	(1.60)
$ROBG_{q+1}$	0.059	0.045	0.072**	0.06**	0.271**	0.219	0.204**	0.161*	0.271	0.200
	(1.17)	(0.88)	(2.38)	(2.00)	(1.98)	(1.50)	(2.26)	(1.73)	(1.23)	(0.94)
$STEMG_{q+1}$	0.204**	0.165	0.148	0.121	0.307	0.237	0.266	0.211	0.355***	0.268**
	(2.09)	(1.62)	(1.28)	(1.02)	(1.48)	(1.15)	(1.14)	(0.94)	(2.80)	(2.19)
CDG_{q+1}	0.505***	0.399***	0.202***	0.160**	0.044**	0.036**	0.184***	0.142**	0.085**	0.070**
	(3.11)	(2.63)	(3.04)	(2.50)	(2.56)	(2.13)	(2.72)	(2.29)	(2.43)	(2.02)
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year-Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	4169	4124	3998	3956	4084	4040	2978	2946	2978	2946
Adj. R2	0.68	0.63	0.58	0.48	0.40	0.40	0.29	0.24	0.19	0.16

What's the interpretation?

1. Purely predictive interpretation:
   - E.g., positive firm shock $\rightarrow$ more data $\rightarrow$ utilizes blockchain
   - Results (fundamentals & returns) are caused by positive firm shock

What's the interpretation?

1. Purely predictive interpretation:
   - E.g., positive firm shock $\rightarrow$ more data $\rightarrow$ utilizes blockchain
   - Results (fundamentals & returns) are caused by positive firm shock

2. Weaker causal interpretation: Data (collecting and using) improves firm performance
   - More data $\rightarrow$ manage the firm better (fewer agency frictions, etc.)
   - More data $\rightarrow$ better understanding of demand/supply/marketing
   - See, e.g., Babina et al (2022)

What's the interpretation?

1. Purely predictive interpretation:
   - E.g., positive firm shock $\rightarrow$ more data $\rightarrow$ utilizes blockchain
   - Results (fundamentals & returns) are caused by positive firm shock

2. Weaker causal interpretation: Data (collecting and using) improves firm performance
   - More data $\rightarrow$ manage the firm better (fewer agency frictions, etc.)
   - More data $\rightarrow$ better understanding of demand/supply/marketing
   - See, e.g., Babina et al (2022)

3. Stronger causal interpretation: Blockchain *per se* improves firm performance
   - Blockchain technology *itself* is an improvement over other methods of managing data
   - E.g., blockchain allows firm to credibly report fundamentals

What's the interpretation?

1. Purely predictive interpretation:
   - E.g., positive firm shock $\rightarrow$ more data $\rightarrow$ utilizes blockchain
   - Results (fundamentals & returns) are caused by positive firm shock

2. Weaker causal interpretation: Data (collecting and using) improves firm performance
   - More data $\rightarrow$ manage the firm better (fewer agency frictions, etc.)
   - More data $\rightarrow$ better understanding of demand/supply/marketing
   - See, e.g., Babina et al (2022)

3. Stronger causal interpretation: Blockchain *per se* improves firm performance
   - Blockchain technology *itself* is an improvement over other methods of managing data
   - E.g., blockchain allows firm to credibly report fundamentals

All three are interesting to different audiences

But I would like the paper to be sharper about the preferred explanation

Causal analysis—IV

Panel A: 1st stage regression

	BDG_{q+1}	BDG_{q+2}
BSI	0.421***	0.355***
	(3.35)	(2.82)
Controls	Y	Y
N	10140	10032
Adj. R2	0.38	0.35

Panel B: 2nd stage regression of fundamentals

	ROA_{q+1}	ROA_{q+2}	AG_{q+1}	AG_{q+2}	SG_{q+1}	SG_{q+2}	PA_{q+1}	PA_{q+2}	PG_{q+1}	PG_{q+2}
BDG_{q+1}	3.538***	1.964***	1.566***	0.992**	0.276***	0.243**	1.056***	0.952***	0.870***	0.533**
	(4.34)	(3.06)	(3.85)	(2.35)	(3.71)	(2.46)	(3.33)	(2.67)	(4.19)	(2.45)
1st stage residual	1.542***	1.120***	0.930***	0.515**	0.160***	0.139**	0.698***	0.556**	0.484***	0.376**
	(4.03)	(2.71)	(3.34)	(2.17)	(3.02)	(2.25)	(2.97)	(2.10)	(3.64)	(2.22)
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year-Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	9577	9475	9186	9088	9381	9282	6840	6768	6840	6768
Adj. R2	0.67	0.62	0.57	0.47	0.40	0.39	0.28	0.24	0.19	0.15

Why don't I think the paper makes the causal claim yet?

IV analysis: preferred interpretation

- Relevance condition: knowledge spillovers
- Exogeneity condition: industry blockchain investment uncorrelated with industry performance

Why don't I think the paper makes the causal claim yet?

IV analysis: preferred interpretation

- Relevance condition: knowledge spillovers
- Exogeneity condition: industry blockchain investment uncorrelated with industry performance

My worry: industry-level shocks you can't control for

- Industry invests in blockchain tech in anticipation of good growth
- OR, industries that are doing well have money to burn on blockchain technology
- (And you can't have time-industry FE)

Causal analysis—difference in difference

Causal analysis—difference in difference

Panel A: difference-in-difference tests of fundamentals

	ROA	AG	SG	PA	PG
Treat*Post	0.152***	0.112***	0.058**	1.224***	1.286**
	(3.91)	(3.13)	(2.32)	(3.47)	(2.57)
Treat	0.099	0.069	0.040	0.819	0.817
	(1.49)	(1.02)	(0.86)	(0.99)	(0.74)
Post	0.060	0.047	0.025	0.546	0.462
	(0.89)	(0.59)	(0.36)	(0.45)	(0.64)
Controls	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y
Year-Quarter FE	Y	Y	Y	Y	Y
N	4056	3890	3973	2897	2897
Adj. R2	0.33	0.28	0.19	0.14	0.09

Why don't I think the paper makes the causal claim yet?

Diff in diff analysis: preferred interpretation

- Timing of blockchain adoption is exogenous
- Look at outcomes post blockchain adoption

Why don't I think the paper makes the causal claim yet?

Diff in diff analysis: preferred interpretation

- Timing of blockchain adoption is exogenous
- Look at outcomes post blockchain adoption

My worry: firm-level shocks you can't control for

- Firm invests in blockchain tech in anticipation of good growth
- OR, firms that are doing well have money to burn on blockchain technology
- (And you can't have firm-time FE)

Thoughts on strengthening the causal claim

What data is actually on the blockchain? Extremely helpful for interpretation...

Data about the firm (e.g., firm accounting fundamentals)

- Suggests blockchain technology helps solve information asymmetry
- Interpretation is more likely to be causal

Data generated through firm's operation (e.g., sales records)

- Early predictor of firm growth/performance
- Interpretation more likely to be about correlations

Sounds like the data classifies this—push harder here?

Thoughts on strengthening the causal claim

Exploit geographical knowledge spillovers

1. Does blockchain adoption spread geographically from tech hubs?

   $\rightarrow$ use geographical HQ distance from Shenzhen as city-level instrument

   Likely need some time variation...

2. Do local firms adopt blockchain after large local player adopts blockchain?

   $Post_{it} = 1$ if a local firm as adopted blockchain prior to time t

   Can include industry-time FE at least...

   Even better: $Post_{it} = 1$ if local firm *in same industry* has adopted

Produces more exogenous blockchain adoption variation

Thoughts on strengthening the causal claim

What drives firms to adopt blockchain?

- E.g., what types, industries, etc.
- Quality of local institutions?

Are firms that adopted blockchain differentially exposed to macro shocks?

- Idea—want to tie blockchain adoption to the types of problems it's supposed to solve/create
- E.g., following big accounting scandal, good shock for blockchain firms?
- Empirical design: $Outcome_{it} = \beta Shock_t \times Adopted_i + ...$

Conclusion

Very interesting preliminary evidence that

- Data on blockchain nowcasts firm fundamentals
- Data on blockchain has asset pricing predictions

Next steps

- Sharpen interpretation (causal vs. correlations—both interesting)
- Delve into what's actually on the firms' blockchains

Check out the paper!