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

Ran Chang[§] Lin William Cong[†]

First draft: December 2021; current draft: March 2022.

Abstract

Despite the explosive growth of cryptocurrencies, whether the underlying technology adds significant value and will thus sustain broad adoption remains unclear. Using firm-level blockchain data from 2015 to 2021, we conduct the first large-sample study linking on-chain data to firm fundamentals and asset valuation in a country where cryptocurrencies are completely banned. We find that year-over-year quarterly blockchain data growth (BDG) contains value-relevant information for nowcasting and forecasting assets growth, sales growth, ROA, standardized unexpected earnings (SUE), and innovation outcomes measured through patents. BDG also predicts stock returns, especially around future earnings announcements, with a long-short BDG-sorted portfolio generating a 10.56% risk-adjusted return annually. The findings are robust across industries and regions, superior compared to other nowcasters, and hold in international samples. We further discuss the underlying economic channels (e.g., continuous disclosure and reduction in information asymmetry) and propose strategies for identifying the blockchain impact. We find results consistent with the aforementioned channels, actual use cases, and heterogeneity analyses that reveal firms with greater information asymmetry, lower disclosure quality, more industry competition, and less public trust benefit more from blockchain adoption and on-chain data growth, especially in the category of firm operations and financials.

JEL Classification: G10, G11, G12, G14.
Keywords: Alternative Data, Blockchain, Digitization, Disclosure, Nowcasting.

[§]Antai College of Economics & Management, Shanghai Jiao Tong University. E-mail: r.zhang@sjtu.edu.cn.

[†]Johnson Graduate School of Management and the FinTech Initiative, Cornell University SC Johnson College of Business. E-mail: will.cong@cornell.edu.

*We thank Greg Buchak (discussant), Bing Han, Mei Lin Hu, Wei Jiang, Tao Li, Penghai Lin, Jacopo Ponticelli, Zheng (Michael) Song, Bernard Yeung, and seminar and conference participants at the 2022 Asian Bureau of Finance and Economic Research (ABFER), the 2022 Halle Institute for Economic Research (IWH), the 2022 China FinTech Research Conference, HKUST (Guangzhou Campus), and Zhejiang University QingYuan Academy Digital Social Science Public Lecture Series for helpful comments. Chang acknowledges financial support from the Shanghai Institute of International Finance and Economics and the Shanghai Pujiang Program. Cong was formerly an advisor for DataYes and is currently the senior economic advisor for Chainlink.

1 Introduction

"You can't have blockchain without crypto."

— CZ_binance, *Twitter, 2022/2/1*

Blockchains are distributed ledgers for relatively decentralized consensus and recordkeeping. Because they prevent single points of failure (and thus systematic risk), encourage competition, and allow multi-party computation and communications that are potentially privacy-preserving (Cong and He, 2019), they have been hailed as the next biggest innovation in general purpose technologies and the engine for Web3. However, much of the debate has centered around its payment applications in the form of digital currencies or (utility) tokens in Decentralized Finance (DeFi). This is epitomized in the tweet by Changpeng Zhao (under Twitter ID CZ_binance), the founder of Binance and the richest Canadian (the 14th richest person in the world, according to Bloomberg Billionaires Index) at the time. He was probably underscoring the importance of on-chain value exchange and the attention is understandable given the explosive growth in cryptocurrency (with a market capitalization of 2.5 trillion USD by the end of 2021) and general excitement on DeFi (Harvey et al., 2021) (with a market capitalization of over 130 billion USD as of Feb 2022). Yet one naturally wonders whether blockchains benefit the real economy and business operations without cryptocurrencies and beyond payment applications, a question crucial for the technology's wide adoption and long-term viability. Despite firm executives' favorable sentiments towards the technology (e.g., Pawczuk et al., 2018) and an abundance of anecdotes about the potential beyond payments (with the latest ones focusing on carbon offset using blockchains, e.g., Bruce, 2021; Gkritsi, 2021), the existing literature has shown little evidence of value creation associated with corporate blockchain adoption and usage.

To this end, we provide the first large-sample analysis on blockchains as a data infrastructure, which indicates that blockchain applications are progressing faster than we might have anticipated and are not only about the hype over cryptocurrencies.[1] Specifically, we find that year-over-year quarterly blockchain data growth (BDG), especially concerning firm operations and financials, contains value-relevant information for nowcasting (in the current quarter) and forecasting (in the next quarter) assets growth, sales growth, ROA, standardized unexpected earnings (SUE), and innovation outcomes measured through patents. BDG also predicts stock returns, especially around

[1] Lakhani and Iansiti (2017) speculated, "It has the potential to create new foundations for our economic and social systems. But while the impact will be enormous, it will take decades for blockchain to seep into our economic and social infrastructure."

future earnings announcements, with a long-short BDG-sorted portfolio generating a 10.56% risk-adjusted return annually. BDG's incremental predictive power on top of other accounting signals and recently discovered nowcasters remains in an array of robustness tests.

We then discuss how in practice blockchains as decentralized ledgers guarantee the fidelity and security of data and create trust without the need for a trusted third party. One key difference between a typical database and a blockchain is that the latter stores data using a linked-list structure with time-stamps. This means that blockchain data are immutable and easily shared. Firm can therefore hardly manipulate data ex post and the speed and fidelity concerns in traditional external audits are potentially mitigated. We describe several real-life use cases in Section **??**, which are largely missing from the prior academic literature.[2] A heterogeneity analysis also reveals that small and private firms with low institutional ownership, market power, and analyst coverage tend to benefit the most from BDG, consistent with that blockchains reduce information asymmetry, mitigate agency issues, and save monitoring and intermediary costs. We propose an IV strategy and a difference-in-difference strategy aimed at providing suggestive evidence that blockchains' effects on firm fundamentals and future stock returns are causal, which may be interpreted as operating through the aforementioned channels.[3]

We obtain blockchain data records from 2015 to 2021 from a leading blockchain service platform in China that parallels IBM blockchain platform. As of 2021, our data provider covers more than 25% of listed firms, 70% of all technology firms, 85% of all cities, and all major industries in China, through its provision of permissioned blockchain solutions or "open enterprise blockchain" services.[4] The number of firms in our sample grew from around 100 in 2016 to more than 700 in 2021. Our data also cover firms from other countries, which we use when analyzing the international sample for robustness. The key firm-level metric we construct is the quarterly blockchain data growth (BDG)—the year-over-year growth rate in blockchain data for a firm in a quarter, which alleviates

[2] Across global supply chains, financial services, healthcare, government, and many other industries, many innovators have already utilized blockchains to generate significant business benefits, including greater transparency, enhanced security, improved traceability, increased efficiency and speed of transactions, and reduced costs.

[3] Note that reverse causality does not diminish the fact that blockchain data are observed more real time and thus serve as good predictors of firm performance that are observed with delays. Ex post, our findings are intuitive because after all, blockchains are fundamentally ledgers/databases. It is not only about transaction recording for value exchanges, but also about other types of recordkeeping, disclosure, and information exchange. Virtually any asset of value, tangible (house, car, cash, land) or intangible (intellectual property, patents, copyrights, branding), can be tracked and traded on a blockchain network, reducing risk and costs for all involved. Blockchains also can deliver more frequent updates, providing immediate, transparent, and immutable information accessible to public or permissioned network participants, potentially with privacy-preservation and secure multi-party recordkeeping.

[4] Henceforth referred to as refer to it as "open blockchain." For more information about the concept of open enterprise blockchain, please refer to https://antchain.antgroup.com/products/openchain.

potential calendar-year seasonality. While BDG in quarter $q+1$ can be observed at the end of that quarter, firm fundamentals and accounting variables in that quarter are released with a delay and the earliest in the next quarter, making BDG a more timely variable for forecasting and nowcasting other accounting variables and stock returns.

We confirm that BDG's nowcasting power holds up at the firm-level. We first use BDG in $q+1$ to predict (or nowcast) fundamental variables in $q+1$ after controlling for their lags and other stock characteristics and quarter-q accounting variables (which are only observable in $q+1$). We find that BDG has significant incremental nowcasting power on a firm's outputs. For example, a ten percentage-point increase in BDG predicts an increase of 7.66 percentage points in the return on assets (ROA), 3.61 percentage points in asset growth (AG), 0.57 percentage points in sales growth (SG), 2.63 percentage points in the growth rate of patents applied (PA), and 1.79 percentage points in the growth rate of patents granted (PG) during the same quarter. The fundamentals predictive power of BDG goes beyond nowcasting. A 10 percentage-point increase in $q+1$ BDG also predicts an increase of 4.43 percentage points in ROA, 2.19 percentage points in AG, 0.49 percentage points in SG, 2.09 percentage points in PA, 1.29 percentage points in PG during the next quarter ($q+2$), consistent with that BDG contains information regarding the firm's earnings power in the long run.

We run a horse race of BDG against a battery of alternative nowcasters. They include the year-over-year quarterly growth rates of search volume for firms' products (SEAG), firms' App traffic (APPG), firms' customer product ratings (CUSG), firms' employer ratings (EMPG), number of cars in firms' parking lots (CARG), and credit card spending on firms' products and services (SPEG). Even after controlling for them simultaneously, BDG remains powerful in forecasting ROA, AG, SG, PA and PG in both the current and the next quarter. Compared to BDG, the forecasting power of the other nowcasters is more sporadic. For example, SEAG and CARG only predict ROA, AG and SG, while SPEG only predicts PA and PG.

We then examine BDG's predictive power towards earnings surprises and market reactions during earnings announcements in the next two quarters. We find that BDG has incremental predictive power on contemporaneous quarter's earnings which is released in the next quarter. A 10% increase in BDG predicts a standardized unexpected earnings (SUE) that is 2.51% higher. BDG also predicts the stock earnings announcement return in the next quarter. A 1% increase in BDG predicts an earnings announcement window abnormal return (CAR) that is 1.84% higher, suggesting that BDG contains new information not fully processed by the market before the announcement. Again, the positive predictive power of BDG on SUE and CAR goes beyond quarter $q+1$ and remains significant in quarter $q + 2$, even after controlling for other nowcasters.

Finally, we show that BDG has strong return predictive power. In a quarterly-rebalanced quintile portfolio sorting exercise, a long-short strategy that buys (sells) stocks in the top (bottom) BDG-quintile generates a monthly profit of 0.88% (value-weighted) or 1.46% (equal-weighted). The risk-adjusted value-weighted returns (alphas) are 0.69%, 0.76%, 0.66%, and 0.66% per month, respectively, for the Chinese q-factor model based on Hou et al. (2015), Chinese five-factor model based on Fama and French (2015), and the LSY3 and LSY4 factor models of Liu et al. (2019a), and remain highly significant. Two-thirds of the profit accrues during the earnings announcement month even though such a month accounts for only one third of a quarter. This result supports the view that BDG contains novel information about a firm's fundamentals and such information is incorporated into the price when it is released to the public during the earnings announcement. The return predictive power of BDG is long lasting. The long-short strategy continues to deliver positive returns for up to a year and we do not observe long-run reversals beyond a year. The cumulative return patterns suggest that BDG's return predictability is unlikely driven by a persistent price pressure, which could eventually dissipate to cause a price reversal.

Fama-MacBeth cross-sectional regression approach allows us to tease out the incremental return predictive power of BDG. BDG shows a robust, positive, and statistically significant relation with future excess, industry-adjusted, and geographic-adjusted returns in multi-variate Fama–MacBeth regressions when we control for a number of firm characteristics and risk factors, including firm's size (SIZE), book-to-market ratio (BM), return on assets (ROA), leverage (LEV), PPE growth (PG), intangible growth (IG), one-month lagged return (STR), price momen- tum (MOM), earnings surprise (SUE), Amihud illiquidity measure (ILLIQ), idiosyncratic volatility (IVOL), turnover ratio (TO), analyst coverage (ANA), and institutional ownership (IO). The slope coefficient on BDG drops only slightly from 0.407 in the univariate case to 0.388 in the multivariate case when all other controls are included. Similarly, BDG remains significant when controlling for other nowcasters, either in pair-wise comparisons or in a multivariate-regression-based horse race where all these nowcasters are included.

During our sample period, firm-level on-chain data are accessible to permissioned network members or the public (with a fee), depending on the type of blockchains used. Real life cases suggest that BDG's predictive power should reflect the value of more transparent and timely information disclosure and the cost-saving from dis-intermediation. In the cross section, we expect the value of BDG to be the biggest for small firms, private firms, and firms with low institutional ownership, market power, and analyst coverage, because such firms tend to have opaque and infrequent disclosures that subject to retroactive manipulation, and are less trusted by investors and business partners. Consistent with this notion, we find that the fundamental and return predictive powers of BDG are indeed lower among large firms, state-owned firms, and firms with higher institutional ownership, market power, and analyst coverage. In addition, we appeal to knowledge spillover and use the number of firms in a focal firm's industry or headquarter city that are considered as blockchain leaders to construct instruments to plausibly identify the impact of BDG. We also use propensity score matching in a difference-in-difference framework to analyze the impact of blockchain adoption. All findings are consistent with that BDG causes firm fundamentals to improve and leads to positive excess returns.

A battery of international tests and subsample analyses confirm the robustness of BDG's fundamental and return predictive power. For example, we find BDG to have strong fundamental and return predictive power among other Asian countries such as Indonesia, Malaysia, South Korea, and Thailand. The predictive power is also present in various subsamples: (1) firms in manufacturing industries vs. firms in other industries; (2) firms headquartered in Top 5 provinces in China vs. other firms; (3) pre- vs. post-Covid sample periods. Finally, we explore blockchains' impact beyond individual firms. We first document that customers' on-chain data growth nowcasts and forecasts firm fundamentals and earnings surprises of suppliers as well as the supplies' own BDG does. Similar results hold using suppliers' BDG to predict customer firms' fundamentals and stock returns. Consistently, a complementary analysis shows on-chain data growth from firms' peers in the same blockchain networks also predicts firm fundamentals, earnings surprises, and stock returns.

Literature — Our study adds to the fast-emerging literature on blockchain economics, especially recent attempts documenting the fundamental value-creation and real effects of blockchains beyond cryptocurrencies.[5] While previous studies have discussed theoretical implications and designs of blockchain systems for auditing, information disclosure, supply chains, and secure multi-party computation (e.g., Cao et al., 2018, 2020; Hastings et al., 2021; Iyengar et al., 2021), empirical analyses of blockchains without crypto are scarce. Among them, Chen et al. (2021a) utilize the introduction

[5] Extant studies mostly examine issues related to consensus algorithms (Biais et al., 2019; Saleh, 2021), cryptocurrency mining (e.g., Cong et al., 2021g; Lehar and Parlour, 2020), scalability (e.g., Abadi and Brunnermeier, 2018; John et al., 2020), fee designs Easley et al. (2019); Basu et al. (2019); Huberman et al. (2021), DeFi (e.g., Harvey et al., 2021; Capponi and Jia, 2021), tokenomics (e.g., Cong et al., 2021f,e; Malinova and Park, 2018; Cong et al., 2022b), ICOs (e.g., Lyandres et al., 2019; Howell et al., 2020), pricing of crypto assets (e.g., Liu et al., 2019b; Cong et al., 2021c), manipulation and regulation (e.g., Griffin and Shams, 2020; Li et al., 2021; Cong et al., 2021d, n.d.), or digital currencies (e.g., Gans et al., 2015; Bech and Garratt, 2017; Chiu et al., 2019; Cong and Mayer, 2021).

of state blockchain laws to capture exogenous increase in firms' ability to develop, adopt, and use smart-contracting technology. The authors find that the technology has very asymmetric impact on customers versus suppliers and reshapes the balance of power in supply chain relationships. Chiu (2021) uses descriptions in 8-K filings of U.S. listed firms to identify blockchain adopters and shows that blockchain adoption potentially improves investment efficiency through increasing firm investment sensitivity to stock prices. Several other studies also document positive stock market reactions to plans or announcements of blockchain adoption (Cheng et al., 2019; Cahill et al., 2020).

These studies use U.S. data or focus on announcements of blockchain adoption, and therefore cannot distinguish the fundamental effect of using blockchains from market speculation or cryptocurrency trading. For example, adoption plan is not equivalent to actual usage: real adoption could be lengthy and costly (Guo et al., 2021) while speculative adoption typically leads to reversal of stock prices (Cheng et al., 2019). One notable exception is Chen et al. (2021b) which documents how blockchains reduce information asymmetry in the asset-backed securities market in China. We similarly take advantage of the Chinese setting, which has the world's largest market for FinTech innovation and the absence of cryptocurrencies. BDG allows us to go beyond the announcement effect of adoption and examine asset pricing implications of actual blockchain usage. Instead of focusing on a small sample of firms in a particular industry as in Chen et al. (2021b), we analyze a large cross-section of various industries and offer the first large-sample analysis of the fundamental value of blockchain independent of cryptocurrencies.

Our paper is also related to the literature on alternative dataand their use in nowcasting. Cong et al. (2021a) provides an overview and recent contributions include Barwick et al. (2020) and Berg et al. (2021). Related to corporate finance and asset pricing, Google search volume, web traffic, customer product ratings, crowd-sourced employer ratings, satellite images, and credit card spending are shown to nowcast and forecast firm fundamentals, earnings surprises, earnings announcement returns, etc. in some developing countries (Da et al., 2011; Rajgopal et al., 2003; Huang, 2018; Green et al., 2019; Katona et al., 2018; Zhu, 2019; Agarwal et al., 2021). We verify that search volume, employer rating, parking lot occupancy, Most recently, Chang and Da (2022) show that firm-level cloud data serve as a powerful nowcaster and may facilitate insider trading. We identify a nowcaster that is potentially more publicly available and possesses significant incremental predictive power of accounting fundamentals and market outcomes.

More broadly, the study adds to the literature on the data economy. While earlier literature emphasize how data directly enter production function or enhance products (e.g., Jones and Tonetti, 2020; Cong et al., 2021b, 2022a), we show that data can inform investors, partners, and stakeholders and firms can potentially use technology to improve data sharing and disclosure. In particular, the features of blockchains make them infrastructure candidates for data sharing and are broadly related to debates surrounding data privacy, intermediation, and open banking (e.g., Tang, 2019; Liu et al., 2020; He et al., 2020; Babina et al., 2022).

The rest of the paper is organized as follows. Section 2 describes the data and variables. Section 3 presents findings linking blockchains to firm fundamentals, earnings surprises, and innovation. Section 4 relates on-chain data growth to firm valuation and cross-sectional stock returns. Section 5 examines economic channels and identify the impacts of on-chain data through case studies, heterogeneity analysis, and instrumental variables and DID analysis. Section 6 conducts robustness tests and provides further discussion. Section 7 concludes.

2 Data and Variables of Interest

We introduce our data and describe key dependent and independent variables in our analyses, which include nowcasters using alternative data as well as conventional accounting variables.

2.1 Data Sample

We obtain proprietary blockchain data from a leading blockchain service platform, which mainly operates in China and Asian countries nearby. Similar to the business model of IBM Blockchain Platform, this platform offers blockchain technology services to enterprise clients. Firms select variable categories to automate the uploading of information onto the blockchain and can rarely misreport because the data have to be consistent and are accessible by auditors. As we describe in Section 5.1, most firms in our sample upload information daily, although our analyses primarily use data quantity aggregated at quarterly level.

Our study focuses on publicly listed firms in China. Because each registered business entity has a Unified Social Credit Code (USCC) issued by the Chinese government, we extract USCC information about the platform's clients and match the platform data with the China Stock Market and Accounting Research Database (CSMAR).[6] Specifically, we observe on-chain data size, firm

[6]CSMAR provides comprehensive information about stock prices, financial statements, corporate governance, and ownership structure for all publicly listed firms in Shanghai and Shenzhen stock exchanges. It is widely used for both academic and industry research.

basic info, accounting variables, financial and operation variables, number of employees, etc. On-chain data are further divided into seven categories: Operation, Financials, Human Resources, Marketing, Supply Chain, and others. Industrial firms have more operation information such as number of raw materials, products, storage, and waste, but financial firms have more financial information, payment, borrowing, lending, investing and so on.

Our data cover all publicly listed firms on the Shanghai and Shenzhen stock exchanges. We apply several filters in constructing our main sample. First, we exclude stocks having less than 15 days of trading records during the most recent month. Second, we remove firm-quarter observations with missing financial information. While we have the firm-level blockchain data since 2015, the need to compute year-over-year quarterly growth rates require us to begin our analyses in 2016. Our final sample includes 11,497 firm-quarter observations, which cover 1,149 unique firms. The sample period includes 22 quarters in total, beginning in 2016/Q2 and ending in 2021/Q3. On average, our sample covers around 523 firms per quarter, which is much larger than those in previous studies exploiting alternative data in the U.S. market.[7]

Figure 1 displays the number of firms in our sample by quarter (Panel A) as well as the average size of on-chain data for a firm (Panel B). On-chain data as a percentage of a firm's overall "cloud-based" (non-nocal) data has been growing too (Panel C). Table 1 reports the summary statistics of the dependent and independent variables in our main analysis. Panel A reports the firm characteristics. The blockchain data (BD) measure 599.122 terabytes on average. Other statistics in Panel A suggests that firms in our sample, on average, have quarterly return on assets of 1.351%, market capitalization of RMB 5.66 billion RMB, book-to-market ratio of 0.445, book leverage of 0.179, percentage ownership by institutional investors of 6.589%, and 7.455 analysts. Panel B reports the characteristics of firm fundamentals, earnings surprise, and innovation performance. The average growth of total assets (AG), growth of sales (SG), return on assets (ROA), standardized unexpected earnings (SUE), earnings announcement abnormal returns (CAR), the log of one plus quarterly number of patents applied (PA), and the log of one plus quarterly number of patents granted (PG). are 0.179, 0.175, 1.455, 0.124, 0.300, 1.176, 1.040, respectively. Panel C reports our main variable of interest in this paper, the BDG. We note that the mean (median) value of this measure is 0.091 (0.072). The variation of this measure is also large, with the 5th and 95th percentiles being -14.3% and 36.2%, respectively. Panel C also report other nowcasters. The mean

[7] For example, using the customers' review data, Huang (2018) covers 150 firms each month on average. Using employers' review data, Goldstein and Yang (2019) covers 508 firms each quarter on average.

of these nowcasters ranges from 0.061 to 0.127. Finally, Panel D reports various types of on-chain data as a percentage of total on-chain data.

2.2 Key Variables

Our key variable is blockchain data growth (BDG). In the literature, many nowcasters have been shown to forecast firm fundamentals and earnings surprise. In our analyses, we construct the Chinese version of these variables to (i) examine whether they have predictive power in the Chinese market, and (ii) make sure that the impact of BDG is not subsumed by these other nowcasters using alternative data.

Blockchain data growth (BDG). We first aggregate blockchain data at the quarterly frequency to reduce noise and better match quarterly financial reports.[8] We then construct $BDG_{i,q}$, the natural logarithm of the amount of blockchain data of Firm i in quarter q (# of $BD_{i,q}$) minus the natural logarithm of the amount of blockchain data of the firm in the same quarter last year $q-4$ (# of $BD_{i,q-4}$):

$$BDG_{i,q} = Ln\left(\frac{\#\ of\ BD_{i,q}}{\#\ of\ BD_{i,q-4}}\right).$$

Intuitively, a bigger BDG reflects faster data growth and more service usage of the technology.

Search volume growth (SEAG). Da et al. (2011) find that Google search volume in the United States for firms' products can predict revenue surprises, earnings surprises, and earnings announcement returns. We construct a similar variable in the Chinese context using product search data from Baidu (https://index.baidu.com). Specifically, we look at Baidu Index, the equivalent of "Google Trends" for the year-over-year quarterly growth of search volume for firms' products ($SEAG_{i,q}$), defined as the natural logarithm of the search volume of products of Firm i in quarter q (# of $SEA_{i,q}$) minus the natural logarithm of the search volume of products of the firm in the same quarter last year $q-4$ (# of $SEA_{i,q-4}$):

$$SEAG_{i,q} = Ln\left(\frac{\#\ of\ SEA_{i,q}}{\#\ of\ SEA_{i,q-4}}\right).$$

[8]Our main results remain robust with higher frequency aggregations.

A bigger SEAG means more growth of search volume for firms' products and indicates high attention to firms' products.

App traffic growth (APPG). Rajgopal et al. (2003) find that website traffic has substantial explanatory power for stock prices and can forecast earnings and book value of equity. In the same spirit, we construct the year-over-year quarterly growth of firms' App traffic ($APPG_{i,q}$), defined as the natural logarithm of the visiting volume of App of Firm i in quarter q (# of $APP_{i,q}$) minus the natural logarithm of the visiting volume of App of the firm in the same quarter last year $q-4$ (# of $APP_{i,q-4}$):

$$APPG_{i,q} = Ln\left(\frac{\#\ of\ APP_{i,q}}{\#\ of\ APP_{i,q-4}}\right).$$

A larger APPG means higher growth in traffic for firms' App and indicates high attention to firms' information. The App traffic data are from Qianfan (https://qianfan.analysys.cn), an authoritative digital economy insight platform covering 45 domestic fields, more than 300 industries and over 50000 apps, and serving more than 1000 enterprise customers.

Customer product rating growth (CUSG). Huang (2018) find abnormal customer ratings positively predict revenues and earnings surprises. The consumer opinions contain information about firms' fundamentals and stock pricing. We construct the year-over-year quarterly growth of customer product ratings of firms ($CUSG_{i,q}$), defined as the natural logarithm of the customer product ratings of Firm i in quarter q (# of $CUS_{i,q}$) minus the natural logarithm of the customer product ratings of the firm in the same quarter last year $q-4$ (# of $CUS_{i,q-4}$):

$$CUSG_{i,q} = Ln\left(\frac{\#\ of\ CUS_{i,q}}{\#\ of\ CUS_{i,q-4}}\right).$$

A larger CUSG means higher growth of customer product ratings of firms and indicates high customer satisfaction. The customer product ratings are from ECdataway (https://www.ecdataway.com), which provides customer product analysis and ratings of China's leading E-commerce platforms, including T-Mall, Taobao, JD, Pinduoduo, Suning, Jumei, etc.

Employer rating growth (EMPG). Green et al. (2019) find firms experiencing improvements in crowd-sourced employer ratings significantly outperform firms experiencing declines. Employer rating changes are associated with growth in sales and profitability and help forecast one-quarterahead earnings announcement surprises. We construct the year-over-year quarterly growth of employer ratings of firms ($EMPG_{i,q}$), defined as the natural logarithm of the employer ratings of Firm i in quarter q (# of $EMP_{i,q}$) minus the natural logarithm of the employer ratings of the firm in the same quarter last year $q-4$ (# of $EMP_{i,q-4}$):

$$EMPG_{i,q} = Ln\left(\frac{\#\ of\ EMP_{i,q}}{\#\ of\ EMP_{i,q-4}}\right).$$

A bigger EMPG means more growth of employer ratings of firms and indicates high employee satisfaction. The employer ratings are from Kanzhun (https://www.kanzhun.com), the Chinese counterpart of "Glassdoor," which provides statistics of job, salary, and employer ratings information.

Parking lot occupancy growth (CARG). Katona et al. (2018) and Zhu (2019) use satellite images to count the number of cars in parking lots to construct abnormal changes in parking lot fill rates that can positively forecast revenue, earnings, and earnings announcement returns. We construct the year-over-year quarterly growth of number of cars in parking lots of firms ($CARG_{i,q}$), defined as the natural logarithm of number of cars in parking lots of Firm i in quarter q (# of $CAR_{i,q}$) minus the natural logarithm of number of cars in parking lots of the firm in the same quarter last year $q-4$ (# of $CAR_{i,q-4}$):

$$CARG_{i,q} = Ln\left(\frac{\#\ of\ CAR_{i,q}}{\#\ of\ CAR_{i,q-4}}\right).$$

A bigger CARG means more growth of number of cars in parking lots of firms and indicate high working time. The number of cars in parking lots of firms is from Wywxdata (https://www.wywxdata.cn), the leading satellite remote sensing and data analysis company in China.

Credit card spending growth (SPEG). Agarwal et al. (2021) show that transaction-level credit-card spending provides accurate and persistent signals of customer demand relevant to a firm's stock pricing. We similarly construct the year-over-year quarterly growth of credit card spending on firms' products and services ($SPEG_q$). Specifically, $SPEG_{i,q}$ is defined as the natural logarithm of credit card spending on the products and services of Firm i in quarter q (# of $SPE_{i,q}$) minus the natural logarithm of credit card spending to the products and services of the firm in the same quarter last year $q-4$ (# of $SPE_{i,q-4}$),

$$SPEG_{i,q} = Ln\left(\frac{\#\ of\ SPE_{i,q}}{\#\ of\ SPE_{i,q-4}}\right)$$

A larger value of SPEG means more growth of credit card spending to the products and services of firms and indicates high popularity of products and services of firms. The credit card spending data are from one of the largest commercial banks in China.

Accounting fundamentals and market signals. We include the following control variables:

- SIZE is the firm's market capitalization computed as the logarithm of the market value of the firm's outstanding equity at the end of quarter q-1.
- BM is the logarithm of the firm's book value of equity divided by its market capitalization, where the BM ratio is computed following Fama and French (2008). Firms with negative book values are excluded from the analysis.
- ROA is the quarterly operating income scaled by lagged assets.
- LEV is the quarterly sum of long-term debt and short-term borrowing scaled by total assets.
- Short-term reversal (STR) is the stock's lagged-one monthly return.
- MOM is the stock's cumulative return from the start of lagged-twelve month to the end of lagged-two month (skipping the STR month), following Jegadeesh and Titman (1993).
- PPE Growth (PG) is the year-over-year quarterly growth in property, plant, and equipment scaled by total assets.
- Intangible Growth (IG) is the year-over-year quarterly growth in intangible assets scaled by total assets.
- TO is the quarterly turnover computed as the number of shares traded divided by the total number of shares outstanding in quarter $q - 1$.
- ILLIQ is the quarterly illiquidity measure computed as the absolute daily return divided by daily dollar trading volume, averaged in quarter $q - 1$.
- IVOL is the idiosyncratic volatility defined as the standard deviation of daily residuals estimated from the regression of daily excess stock returns on the daily market, size, and value factors of Fama and French (1993) in quarter $q - 1$.

- SUE is the standardized unexpected earnings defined as actual earnings in the current quarter minus earnings 4 quarters ago, scaled by stock price in the current quarter following Livnat and Mendenhall (2006).

- ANA is defined as the number of analysts following the firm in quarter $q-1$, and IO is the percentage of tradable shares held by institutional investors in quarter $q-1$.

Finally, to measure firms' innovations, we use (i) the log of one plus quarterly number of patents applied of the firm (PA) and (ii) the log of one plus quarterly number of patents granted of the firm (PG). Patent data concerning the firms are from DataYes, the leading FinTech data provider in China.[9] We winsorize all control variables at the 1st and 99th cross-sectional percentiles.

3 On-Chain Data and Changes in Firm Fundamentals

3.1 BDG and firm fundamentals

In this section, we examine how on-chain data growth relates to and actually contains valuable information about firm fundamentals. If blockchain creates value for firms, we should expect companies with faster blockchain data growth to perform better going forward than their peers. Thus, we conduct quarterly panel data regressions of firm fundamentals on BDG and other variables:

$$FF_{i,q+n} = \alpha_d + \beta_1 * BDG_{i,q+1} + \beta_2 * FF_{i,q} + \gamma_{q+n} + controls_{i,q} + e_{i,q+n}, \tag{1}$$

where $FF_{i,q+n}$ is Firm i's fundamentals in quarter $q+n$ (n=1 or 2), α_d is industry fixed effect, $BDG_{i,q+1}$ is Firm i's quarterly blockchain data growth in quarter $q+1$, γ_{q+n} is year-quarter fixed effect. We include the past firm fundamentals in the model to account for persistence in firm fundamentals. We also include control variables listed in Panel A of Table 1 in the regressions.

We focus on three proxies for firm fundamentals that reflect operation performance and are prevalent in the literature (Hirshleifer et al., 2013, 2018): return-on-asset (ROA, quarterly operating income scaled by lagged assets), assets growth (AG, quarterly growth in total assets), and sales growth (SG, quarterly growth in sales). To reduce the influence of outliers, we winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a mean of zero and standard deviation of one. Standard errors are double clustered by industry and by year-quarter.

[9] https://www.datayes.com/

Table 2 presents the average slope coefficients and the corresponding t-statistics from the quarterly panel data regressions. The results show a significantly positive relationship between the BDG and the proxies of firm fundamentals in quarter q+1 or quarter $q+2$. Specifically, we regress ROA, assets growth, or sales growth, in quarter $q+1$ or quarter $q+2$ on the BDG in quarter $q+1$ as well as the assets growth, sales growth, or ROA in quarter q.

For quarter $q+1$, the coefficients between the BDG and firm fundamentals are significant at the 1% level after accounting for the control variables and the industry and year-quarter fixed effects. The coefficients for quarter $q+1$ ROA, assets growth, sales growth, PA, and PG on $q+1$ BDG are 0.766, 0.361, 0.057, 0.263, and 0.179, respectively. For quarter $q+2$ fundamentals, the coefficients on $q+1$ BDG are 0.443, 0.219, 0.049, 0.209, and 0.129, respectively, all at 1% significance level.

The last four columns in Table 2 link BDG to innovation outcomes. The BDG can forecast PA and PG metrics while explaining contemporaneous results. The coefficients on BDG are 0.263, 0.209, 0.179, and 0.129, respectively. The findings imply that a one percent increase in on-chain data can lead to about 0.2 percent increases in firm innovation in terms of patent applications and grants.

Note that firm fundamentals for $q+1$ in practice are only computed and announced after $q+1$, whereas BDG for $q+1$ is available by the end of $q+1$. Therefore, timely observations of BDG allow for nowcasting of firm fundamentals. BDG clearly correlates with contemporaneous operational and financial performances and positively predict future performances. Had we aggregated on-chain data growth at a higher frequency, more real-time BDG would have predictive power for shorter horizons as well.

Table 3 compares BDG with other nowcasters using alternative data. In each column, we add all six nowcasters as additional controls in the regression. We find these nowcasters do not diminish the ability of BDG for nowcasting and forecasting firm fundamentals. To nowcast and forecast firm's ROA, the coefficients of BDG are 0.524 and 0.408 and the t-statistics are 3.24 and 2.99. To nowcast and forecast firms' total asset growth, the coefficients of BDG are 0.179 and 0.153 and the t-statistics are 3.34 and 2.77. To nowcast and forecast firm's sales growth, the coefficients of BDG are 0.050 and 0.037 and the t-statistics are 2.85 and 2.73. Compared to BDG, the forecasting power of the other nowcasters is significant only sporadically. For example, SEAG and CARG only predict ROA, asset growth and sales growth, while SPEG only predicts patent outcomes.

The requirement for all seven nowcasters to be available for the firm significantly reduces the sample size in Table 3. We thus also horse race BDG against alternative nowcaster, one at a time, in larger samples, and reach very a similar conclusion. Again, BDG's incremental predictive power always remains. Overall, the results indicate that BDG indeed contains valuable information about firm fundamentals not captured elsewhere.

3.2 BDG and Earnings Surprises

While information on blockchain data is typically not available to the public in real time, it can be released via future earnings announcements. In this subsection, we examine whether the BDG can nowcast and forecast future earnings surprises. We use standardized unexpected earnings (SUE), defined as actual earnings in the current quarter minus earnings 4 quarters ago, scaled by stock price in the current quarter, following Livnat and Mendenhall (2006), as proxy for earnings surprise. We conduct panel data regressions of the quarterly SUE (for fiscal quarters $q+1$ and $q+2$, which are announced in quarters $q+2$ and $q+3$) on the BDG in quarter $q+1$ and control variables of Panel A of Table 1 in quarter q. BDG's predictive power of earnings surprises and market reaction during earnings announcements in the next two quarters.

We also examine whether BDG can nowcast and forecast earnings announcement abnormal returns (CAR). CAR is the cumulative abnormal returns over the three-day window surrounding the earnings announcement. Abnormal return is calculated as the raw daily return minus the daily return on size and market-to-book matched portfolio as in Livnat and Mendenhall (2006). We conduct panel data regressions of the quarterly CAR (corresponding to announcements of quarter $q+1$ and $q+2$ earnings) on BDG in quarter $q+1$ and control variables of Panel A of Table 1 in quarter q. For panel data regressions, we also control for the industry and year-quarter fixed effects.

Again, we winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a mean of zero and standard deviation of one to reduce the effect of outliers. Standard errors are double clustered by industry and by year-quarter. If BDG contains nowcasting and forecasting information about SUE or CAR, we should expect the slope coefficient to be positive and significant.

Consistent with our expectation, for quarter $q+1$ SUE, Table 4 shows that the coefficient on the BDG is 0.251 with a t-statistic of 4.19 when accounting for past SUE, control variables, and the industry and year-quarter fixed effects. For quarter $q+2$ SUE, the coefficient on the BDG is 0.187 with a t-statistic of 3.16 after controls. Moreover, consistent with Bernard and Thomas (1989), the lagged SUE at quarter q is strongly positively correlated with the future SUE. In Column 3 and

4, we find that BDG can forecast CAR in the next two quarters. The coefficients on the BDG are 1.839 (t-statistic = 2.72) and 1.437 (t-statistic = 2.04).

We also examine BDG's SUE nowcasting and forecasting after adding other nowcasters as controls in Table 5. In each column, we add all six nowcasters as additional controls in the regression. We find that these nowcasters cannot significantly change our BDG nowcasting and forecasting power to earnings surprises. To nowcast and forecast firm's SUE, the coefficients of BDG are 0.114 and 0.088 and the t-statistics are 2.56 and 2.08. To forecast firms' next two quarter CARs, the coefficients of BDG are 1.518 and 1.308 and the t-statistics are 2.23 and 1.99.

4 Asset Pricing Implications and Return Forecasts

We examine the link between BDG and the cross-section of future stock returns using portfolio-level and firm-level regression analyses. If BDG improves fundamentals, it should be recognized by the market and serves as a predictor for future equity returns. Those who observe BDG in a timely manner therefore can generate significant trading profits.

4.1 Univariate Portfolio Sorts Using BDG

To construct the long-short portfolio, at the end of each quarter from 2016/Q2 to 2021/Q3, individual stocks are sorted into quintile portfolios based on their BDGs in that quarter and are held for the next quarter. We then compute the value-weighted and equal-weighted average monthly excess return of each quintile portfolio. To examine the cross-sectional relation between the BDG and the future stock returns, we form a long-short portfolio that takes a long position in the highest quintile of BDG and a short position in the lowest quintile of BDG.

In Table 6, we report the average monthly excess returns of each quintile portfolio and the long-short portfolio (in excess of the one-month deposit interest rate). We also report the abnormal returns (alphas) estimated with various factor models, including the China q-factor model based on Hou et al. (2015), China five-factor model based on Fama and French (2015), the LSY3 and LSY 4 factor models in Liu et al. (2019a). Controlling for these factors helps to ensure that the BDG indeed contains incremental predictive power beyond these well-known factor models. We also report average excess returns in earnings announcement months and average excess returns in non-earnings announcement months.

In general, the excess returns and alphas of five quintile portfolios increase monotonically from quintile 1 to quintile 5. The long-short portfolio that buys 20% of the stocks with the highest BDG (quintile 5) and short-sells 20% of the stocks with the lowest BDG (quintile 1) earns a value-weighted (equal-weighted) average return of 0.880% (1.462%) per month with a t-statistic of 3.44 (6.08), translating into an annualized return of 10.560% (17.544%). Controlling for the factors does not change the magnitude and statistical significance of the return spreads on the BDG-sorted portfolios for most of the factor models. The alpha is from 0.688% (HXZ) to 0.659% (LSY4) per month and the corresponding t-statistic is from 3.26 to 3.34 for the value-weighted portfolio. Finally, the significant relation between BDG and future returns is largely coming from the short leg of the arbitrage portfolio as the economic magnitude and statistical significance are larger among the stocks in the short leg than those in the long leg. This implies that lower BDG firms are overvalued relative to firms with higher BDG, perhaps due to the short selling limitation in China. In earnings announcement months, the value-weighted (equal-weighted) long-short excess returns are 0.580% (1.092%). In non-earnings announcement months, the value-weighted (equal-weighted) long-short excess returns are 0.300% (0.370%). The excess returns in earnings announcement months are about 2-3 times larger than the excess returns in non-earnings announcement months.

4.2 Fama-MacBeth Cross-Sectional Regression

We conduct firm-level Fama-MacBeth cross-sectional regressions to test if BDG predicts the cross-section of monthly returns in the next quarter. The test allows us to examine the incremental predictive power of BDG by controlling for other known return predictors. Each month, we run a cross-sectional regression of stock returns in that month on the last quarter BDG as well as a number of control variables, including lagged size, book-to-market, ROA, leverage, PPE growth, intangible growth, earnings surprise, short-term return reversal, price momentum, idiosyncratic volatility, illiquidity, turnover ratio, analyst coverage, and institutional ownership. To minimize the effect of outliers, all independent variables are winsorized at the 1st and 99th percentiles. We also control for the industry and region fixed effects following the CSRC industry classification and China province classification. The stock-level cross-sectional regressions are run each month and the standard errors of the average slope coefficients are corrected for heteroskedasticity and autocorrelation following Newey and West (1987).[10]

Panel A of Table 7 reports the Fama-MacBeth cross-sectional regressions' results. In Column 1, we include only BDG in the cross-sectional regressions. Consistent with the portfolio results,

[10]To compute standard errors, we use the Newey-West adjustment with three lags.

we find a positive and significant relation between the BDG and one-month-ahead returns. The average slope coefficient on the BDG ratio is 0.407 with a t-statistic of 4.49. In Column 2, we further control other well-known return predictors in the cross-sectional regressions. We find a positive and significant relation between the BDG and one-month-ahead returns controlling for a large number of predictors. The BDG retains significant predictive power, and the magnitude of the average slope coefficient decreases only slightly to 0.388, suggesting that the information embedded in BDG is almost orthogonal to that in other known return predictors. The slope coefficients on the control variables are consistent with prior literature: market capitalization (SIZE), short term reversal (STR), and idiosyncratic volatility (IVOL) are negatively correlated with the future return, and ROA, earnings surprise (SUE), and institutional ownership (IO) are positively related to the next month's return.

In Column 3, we include INDRET, which is computed as the value-weighted CSRC industry portfolio returns, as a control variable in our main regression to further control for the industry effect. Specifically, we adjust the dependent variable by subtracting the firm's value-weighted CSRC industry return INDRET from the firm's current month return. Doing so allows us to tease out the return predictive power from the BDG rather than the one-month industry momentum effect. The coefficient of the BDG remains similar controlling for the industry return directly. In Column 4, we further control for the geographic momentum that are shown to affect stock returns systematically. Specifically, we use RET-GEORET, which is the difference between the firm's return and the corresponding province portfolio returns. We replace the firm's raw return with this geographic-adjusted return as the dependent variable and run the same monthly cross-sectional regressions. Again, the magnitude of the slope coefficient on BDG becomes slightly weaker, but remains highly significant.

Panel B of Table 7 reports Fama-MacBeth cross-sectional regressions of BDG and other nowcasters. The nowcasters cannot significantly change the predictive power of BDG. After adding all six nowcasters, the predictive coefficient of BDG becomes 0.282 and the corresponding t-statistic is 2.78. The sample size is much smaller due to the requirement that both BDG and other nowcasters have to be non-missing. Overall, these results indicate that BDG provides value-relevant information and incremental predictive power over other well-known return predictors, nowcasters, and regression specifications.

4.3 Observability and Pricing Efficiency using BDG

While the predictive power of BDG on firm fundamentals and future stock returns suggests the tangible impact of blockchains, whether an investor can use BDG to trade profitably depends on the observability of on-chain data. A majority of the firms in our sample adopt permissioned blockchains, which means only members of the network can observe the information. Members of permissioned blockchains are not insiders of specific companies. Therefore unlike the insider trading documented in Chang and Da (2022), they can legally use BDG to construct profitable trading strategies, as we have shown. To the extent that data on blockchain adoption and data growth are commercially available in a timely manner, BDG serves as a powerful predictor for trading by general market participants.

Particularly interesting to us is to provide a "lower bound" on how BDG can be used by equity market investors to design profitable strategies to arbitrage, which can help impound fundamental information into asset prices. To this end, we examine a special subsample of firms that adopted open blockchains (as defined by the data provider to be blockchains that one can join with a fee), which corresponds to about 10% of the full sample. Table 8 reports the findings. The long portfolio using the highest BDG-quintile open-blockchain firms and the long-short BDG-sorted portfolio generate significant excess returns of 62 and 55 basis points and sharpe ratios of 0.72 and 0.61, respectively. The equal-weighted portfolios generate even higher Sharpe ratios and almost 1% of montly excess returns. The returns remain significant even after controling for various risk-factors.

The trading profits with publicly observable information imply that arbitrageurs, in addition to permissioned members, are going to impound BDG into market valuations quickly in the market. The transparency of blockchain data potentially allows for a faster incorporation of fundamentals into asset prices.

5 Economic Channels Underlying On-Chain Data's Impact

We now discuss channels through which BDG can improve firm fundamentals and thus future valuations of the firms. Through a number of use cases, we first list the channels that firm executives had in mind when adopting blockchains. We then show that the empirical patterns from heterogeneity analysis are consistent with the channels. Finally, we note that endogeneity can arise in our setting because firms decide whether to use blockchain services or grow on-chain data. We use both instrumental variable and difference-in-difference methods to address this concern and provide suggestive evidence for causal linkages.

5.1 Real-Life Usage Cases and the Blockchain Data Edge

We describe four examples using firms in our sample. We anonymize the firm names per the data usage agreement.

Suppliers and customers in a permissioned blockchain. Our data include one of the largest auto glass manufacturers in China, with 25% of the global auto glass market and 70% of the auto glass market in China. Its customers include several Chinese listed auto companies, such as SAIC, FAW Jiefang, Geely Automobile, etc.

The company updates inventory and purchase of raw materials on the blockchain every day, including silica sand, soda ash, feldspar, dolomite, limestone, thenardite, etc.; it also updates on the blockchain inventory and sales volume of automotive glass products, including windshield glass, sunroof glass, thermal insulation glass, sound insulation glass, dimming glass, etc. In addition, the company updates auto glass production line investment and profit data (including factory location, construction time, production glass type, investment amount, sales, etc.) and daily news of factory operations (including the number of employees, hourly wages, total hours worked, electricity consumption, etc.) on the blockchain.

The automobile company that purchases glass from the company also updates the inventory of remaining auto glass and the usage of auto glass on the blockchain every day, as well as their inventory and sales of cars. In addition, the three auto companies update the information on the 4S store for selling cars (including location, time, customer volume, sales models, etc.) and the daily news of the store (including the number of customers, number of employees, hourly wages, operating costs, etc.) on the blockchain.

Banks and borrower companies. Many automobile companies in our sample disclose their car sales on the blockchain every day, including the number of small cars, the number of medium-sized cars and the number of large cars. This information is observable to the banks that they borrow from.

For example, on March 31, 2021, the sales and inventory of small cars reported on-chain by a company were 5713 and 31238, the sales and inventory of medium-sized cars were 2414 and 14,561, and the sales and inventory of large cars were 531 and 788. The book balance was 523 million yuan and the number of employees was 78,300. The bank provided different loan products to the company on the blockchain, including 1) 1-year term, 3.7% loan interest rate and monthly repayment of equal principal and interest. 2) 5-year term, 4.65% loan interest rate, monthly repayment of equal principal and interest. 3) 10-year term, 5.8% loan interest rate, monthly equal principal repayment.

On June 30, 2021, the sales and inventory of small cars by the same company were 6,831 and 28,534, the sales and inventory of medium-sized cars were 3,152 and 11,629, and the sales and inventory of large cars were 613 and 583. The book balance was 569 million yuan and the number of employees was 79,100. The bank provided different loan products to the company on the blockchain, including (i) one-year term, 3.65% loan interest rate and monthly repayment of equal principal and interest, (ii) 5-year term, 4.6% loan interest rate, monthly repayment of equal principal and interest, and (iii) 10-year term, 5.7% loan interest rate, monthly equal principal repayment.

We see that with the increase of car sales and the decrease of inventory, the bank provided more favorable loan interest rates and conditions. In general, borrower companies disclose operational, financial, employee, marketing, and sales information to the bank so that the bank can provide specific financial products to companies based on real-time updated firm information. Because the firm information is timely, real, transparent, and complete, banks believe the target company is much safer than other off-chain companies and would like to give lower credit rate and longer credit maturity to companies. In addition, banks can increase their profitability and reduce their delinquency and default rates due to more prompt information processing than other off-chain banks.

At the same time, many banks disclose various of financing products and condition (e.g. rate, maturity, repayment methods) on-chain to help managers and executives at target companies compare financing products from different banks and select the most suitable product for reducing costs and improving performance (short-term operating profit, long-term sustainable development, ESG score, etc.).

Companies and institutional investors. One electronic consumer company sells products such as mobile phones, televisions, notebooks, and desktops. The company discloses the sales volumes and inventory of these electronic products on the blockchain every day. For example, on September 15, 2020, the sales volume of mobile phones was 5,281, the sales volume of TVs was 1,032, the sales volume of laptops was 2,718, and the sales volume of desktops was 875. The book balance of the company was 179 million yuan and the number of employees was 23,400. On September 15, 2020, a hedge fund held a 1.98% stake in the company.

On December 15, 2020, the sales volumes for mobile phones, TVs, notebooks, and desktop computers were 7,213 units, 2,013 units, 2,419 units, and 907 units respectively. The book balance of the company was 233 million yuan and the number of employees was 22,800. On December 15, 2020, the hedge fund had increased its stake in the company to 3.51%. The increase in the company's sales of electronic products seems to have led hedge funds to increase their holdings in the electronics consumer company.

More generally, companies would disclose information on-chain to the funds that invest in them. Funds also disclose positions and transaction information in a timely manner to increase investor trust and to offer a menu of products to potential investors.

Startups and venture capital. In the process of fundraising, one start-up coffee company in our sample was automatically uploading coffee sales to its enterprise blockchain it has. For example, on February 1, 2019, the company disclosed data on the blockchain: espresso 233 cups, Americano 152 cups, Machiatto 77 cups, latte 317 cups, mocha 561 cups and cappuccino 445 cups. It received valuations quotes from two VCs: one of which was 15 million RMB and the other was 18 million.

On February 1, 2020, the coffee company disclosed data on the blockchain: Espresso 622 cups, Americano 612 cups, machiato 434 cups, latte 1285 cups, mocha 802 cups and cappuccino 590 cups. It received valuation quotes from three VC companies, which were 35 million, 45 million, and 50 million.

The increase in coffee sales seems to have helped convince the VCs to increase their valuation of the company. In general, startups can disclose operational, financial, employee, marketing, and sales information to VCs. The blockchain can increase trading transparency, investor trust, and information symmetry across companies, VCs, and limited partners (LPs). The transparency can potentially reduce informational frictions and increase firm valuation.

VC funds may also disclose information on-chain to the LPs in a timely manner. On-chain VCs raise more capital from LPs due to greater trust, have more successful capital exits due to improved transparency from startups, and exhibit higher profitability and lower risks due to stronger and more prompt information processing abilities compared with off-chain VCs.

Trust, transparency, and cost saving. Based on the aforementioned cases and interviews and brochures offered publicly by the blockchain data service platform, putting data on-chain can benefit a firm along the following dimensions:

1. Reducing information asymmetry and signaling: greater accessibility to relevant parties or even to the public, coupled with more timely or even real-time reporting, imply that information asymmetry among various transaction parties and players in relationship financing are significantly reduced. One might argue that firms may choose to report on-chain more pieces of good news than bad news, rendering on-chain data growth more correlated with positive shocks. In practice, this is difficult to do because the uploading of information is typically programmed based on the variable category and data consistency would need to be maintained. For example, if a factory's electricity usage is inconsistent with its production report, the firm may be inspected by auditors and regulators.

2. Mitigating moral hazard through algorithmic trust: whereas retroactive manipulation of disclosure is common, blockchain's immutability prevents ad hoc or strategic revisions of records, such as in the case of ESG metrics (Berg et al., 2020). More continuous disclosure also prevents timing-based manipulations due to managerial myopia or short-termism.

   Once transaction data are recorded on chain, it is also very difficult for hackers to compromise them. In any industry where protecting sensitive data is crucial — financial services, government, healthcare — blockchains have the opportunity to transform the way critical information is shared and communicated.

3. Saving monitoring and intermediary costs: the greater transparency and trust discussed above imply that costs associated with monitoring the managers and the governing bodies as well as learning about the firms can be saved. The distributed and synchronized information gets rid of conventional arbitragers and intermediaries, thus reducing accounting and auditing frictions too.

   If a company deals with products that are traded through a complex supply chain, it is likely familiar with how hard it can be to trace an item back to its origin. When exchanges of goods are recorded on a blockchain, the company ends up with an audit trail that shows where an asset came from and every stop it made on its journey. This historical transaction data can help to verify the authenticity of assets and prevent fraud while improving traceability. Note that privacy and propriety of data can be preserved in these processes through the use of encryption methods, zero-knowledge proof schemes, and privacy-preserving multiple-party computation Cao et al. (2018); Hastings et al. (2021).

5.2 Heterogeneity Analyses

Given that on-chain data can help reduce information asymmetry, mitigate agency issues, and save monitoring and intermediary costs, we would expect firms that traditionally suffer the most from these issues to benefit the most from blockchain adoption and BDG. These firms tend to be private and small firms with less institutional investors and analyst coverage. Large firms with high institutional ownership or analyst coverage enjoy a more transparent disclosure practice and information production efforts by institutions and analysts also diminish the incremental value of BDG. SOEs are more trusted because of government backing and have less need for blockchains. Finally, the use of blockchains represents a new digital dimension that a firm can gain an edge over its rivals, so its benefits are likely greater in industries that are more competitive.

To test these, we first divide the observations into subsamples of large versus small, high versus low institutional ownership, high versus low analyst coverage, state-owned enterprises (SOEs) versus private firms, and competitive versus non-competitive industries. We obtain the enterprise type (state-owned and non-state-owned) from the CSMAR. Non-state-owned firms use more blockchain technology services and have more blockchain data than state-owned firms, as state-owned enterprises have sufficient financial strength to build their own private blockchains rather than using open blockchains. We use Herfindahl–Hirschman index (HHI) to proxy for the competition in an industry, where HHI is computed as the sum of squares of each firm market-cap shares (in percentages) in the industry.

Table 9 reports the results which are consistent with our conjecture. We find BDG's nowcasting and forecasting power remains within each subsample, but it is lower among large firms, firms which have higher institutional ownership and analyst coverage, and firms in more competitive industries. A multi-variate regression in Table 10 further confirms that BDG's impact is bigger for small, low-IO, low-coverage, and private firms, as indicated by all the coefficients on the interaction terms being significant and positive. Without the benefit of greater transparency and mitigated agency issues, we would not expect that small, low-IO, low-coverage, and private firms to systematically have greater positive shocks than others.

5.3 Identifying the Impact of Blockchain on Firm-level Outcomes

We propose both an IV strategy and a difference-in-difference strategy that aim to identify the impact of on-chain data growth and blockchain adoption on firm-level outcomes. We find evidence consistent with the aforementioned channels and observations such as that the effects are larger for less transparent firms.

Note that it would be fine if firms that perform better also generate more data on-chain (reverse causality), but blockchain data generation can be observed more in real time so it is still a good predictor of firm performance (which is instead observed with a delay).

Instrumental variable (IV) approach. We first adopt the instrumental variable approach in Chen et al. (2021b). Based on the idea of knowledge spillover, we use the development of blockchain technology in the firm's industry to construct the IV. Specifically, we use the number of firms in the firm's industry that are leaders in developing and providing blockchain services (Blockchain Service Industry), i.e., those firms on the "List of Companies with Blockchain Digital Services" (hereafter, the List) maintained by China's Cyber Security and Digitization Committee. Due to knowledge spillover within the industry (e.g., Kim and Valentine, 2021), the firms in the same industry as those on the list are more likely to have the ability to use blockchain services. Hence Blockchain Service Industry satisfies the relevance criterion. At the same time, when the number of blockchain service leaders in an industry is unlikely to directly affect fundamentals and earnings performance of another firm, except through blockchain-related channels, the exclusion restriction is thus plausibly satisfied.

Using this IV and the control variables, we estimate the determinant model for the blockchain data growth and report the results in Table 11. The coefficients on Blockchain Service Industry are significant and positive. We then use the control function approach to address endogeneity by including the residuals estimated from the first-stage regression, to forecast fundamentals, earnings, and returns. We find that the prediction coefficients on BDG are significantly positive, consistent with our earlier regression results. Note that in the 2nd stage regression we include industry fixed effect and city-quarter fixed effect, in order to address the concern that the exclusion restriction is violated because firms in the same industry or have headquarters in the same city tend to be affected by industry blockchain leaders through non-blockchain related chanels. For example, industry blockchain investment could be correlated with industry performance, i.e., industries investing in blockchains may be the ones anticipating health growth anyway or they may have been extremely profitable and therefore hold excess funding to experiment on new technologies.

However, in the above IV specification, we cannot include industry-quarter fixed effect as that would drive away all variations. We therefore specify a complementary IV: the natural logarithm of the number of companies in the focal firm's city (based on headquarter locations) that are included in the "List of Companies with Blockchain Digital Service," which we refer to as the Blockchain Service Region (BSR) instrument. Tabel 12 reports the results with region fixed effect and industry-quarter fixed effects in the second stage. While neither IV is perfect, they together provide plausible evidence that BDG may cause firm fundamentals and stock returns to be bigger.

Difference-in-difference tests. In order to address the endogeneity concern of blockchain data growth on fundamentals, earnings, and returns, we conduct additional difference-in-difference tests. Specifically, we compare performance outcomes of a firm before and after it uses blockchain services, benchmarked against a control group of peer firms in the same industry and headquarter city with similar characteristics but do not use blockchain services. We examine four quarters before and four quarters after a firm's adoption of blockchain, where event quarter zero is the quarter when the firm first uses the blockchain services. The dummy variable Treat equals one when the firm uses blockchain services, otherwise zero. The control firms do not use blockchain services from our data provider's platform. For each treatment firm, we match control firms in the same industry and region using the propensity score matching method based on each characteristic of fundamentals, earnings, and returns. The dummy variable Post equals one when the firm begin to use blockchain services, otherwise zero.

Table 13 reports the diff-in-diff results of fundamentals, earnings, and returns. In Panel A-C, the coefficients for the Post term are statistically indistinguishable from zero, as expected. The coefficients for the Treat term is positive and significant at the 10% level, implying that firms adopting blockchains in our sample tend to have better trends in fundamentals and valuation metrics. Most importantly, the coefficients for the interaction term Treat*Post are economically large and statistically significant mostly at the 1% level. The results suggest that blockchain adoption and storing data on-chain indeed lead to improvements in firm performance. Note that firms in the control group may have adopted blockchain through other platform services, which means that our findings are likely underestimate of the effects.

6 Further Analyses and Discussion

Our analyses thus far focus on Chinese firms for several reasons: China has arguably the biggest market for FinTech innovations and applications, in addition to being the second largest economy.[11] Moreover, it is one of the few places where cryptocurrencies are banned and yet developing and applying the blockchain technology is highly encouraged. This is important in isolating the fundamental value creation of blockchains as distributed database systems from the market sentiment and speculation surrounding cryptocurrencies.

That said, there are significant heterogeneity across the firms within China and one may question the findings' external validity beyond the Chinese setting. Additional linkages among the firms likely have non-trivial interactions with blockchain adoption and usage. We therefore conduct additional analyses for robustness and extension of our study.

6.1 Robustness in Subsamples

We conduct various robustness tests of our findings: First, we study BDG's nowcasting and forecasting power within the manufacturing industry versus the sample of non-manufacturing industries. The manufacturing industry account for 56% of all sample firms and the non-manufacturing industries account for the remaining 44%. Second, we study BDG's nowcasting and forecasting within the sample of the 5 largest provinces vs. the sample of other provinces' firms. The 5 largest provinces are Guangdong, Zhejiang, Shanghai, Jiangsu, and Beijing, accounting for 60% of all sample firms in total. Finally, we study BDG's nowcasting and forecasting power during the pre-Covid (from April 2016 to December 2019) and post-Covid (from January 2020 to September 2021) episodes.

Table 14 shows these subsample results. In Panel A, we find that the BDG in different subsamples can nowcast and forecast firm fundamentals. Panel B shows that BDG can nowcast and forecast earnings surprise and CAR in different subsamples. In both panels, the coefficients on BDG are larger among firms in manufacturing industries, in Top5 provinces, among private enterprises, and during the post-Covid sample period. Panel C shows that BDG can generate economically and statistically significant returns and alphas in different subsamples. Again, the excess returns and alphas are larger among firms in manufacturing industries, in Top5 provinces, among private enterprises, and during the post-Covid sample period.

[11] For exmaple, Ant has the most number of blockchain patents and has released its first carbon offset records on blockchain Gkritsi (2021).

6.2 International Evidence

We further examine BDG nowcasting and forecasting power to firm fundamentals, earnings surprise, and long-short excess returns and alphas in several other countries. We obtain BDG information of firms in Indonesia, Malaysia, South Korea, and Thailand. Table 15 reports our findings in the international setting. Panel A examines BDG nowcasting and forecasting power to firm fundamentals. Again, the coefficients on BDG in quarter $q+1$ for firm fundamentals (ROA, asset growth, sales growth) in quarter q+1 are significant after accounting for various control variables and adding the industry and year-quarter fixed effects in the four countries. As for firms in Indonesia, Malaysia, and Thailand, the coefficients of BDG in quarter $q+1$ for firm fundamentals (ROA, asset growth, and sales growth) in quarter $q+2$ is significant. But for firms in South Korea, the coefficients become insignificant.[12]

Panel B examines BDG nowcasting and forecasting power of earnings surprises in Indonesia, Malaysia, South Korea, and Thailand. The coefficients of BDG in quarter $q+1$ for earnings surprise (SUE and CAR) in quarter $q+1$ are significant after accounting for the control variables and the industry and year-quarter fixed effects in the four countries. For firms in Indonesia, Malaysia, and Thailand, the coefficients of BDG in quarter $q+1$ for earnings surprise (SUE and CAR) in quarter $q+2$ are significant. But for firms in South Korea, the coefficients become insignificant.

Panel C reports the long-short value-weighted excess returns and alphas of BDG in the four countries. The returns in Indonesia, Malaysia, South Korea, and Thailand are 0.322%, 0.571%, 0.306%, and 0.444%, respectively. The Fama and French (2018) six-factor alphas of Indonesia, Malaysia, and Thailand are significant, it is insignificant for South Korea.

Overall, we find strong evidence for BDG's nowcasting and forecasting power in international markets. The only exception entails South Korea, where BDG's predictive power becomes significant only economically but not statistically. This observation is likely due to the fact that firm disclosure in South Korea is already more transparent and timely in recent years, through government programs such as the Open Data Portal (https://www.data.go.kr/en/index.do), not to mention that it has the most developed financial markets among these countries, which means high market efficiency and the lack of preditable returns at intermediate and long horizons.

[12] The patent data of the four countries are not obtainable.

6.3 Other Nowcasters and Forecasters Related to the Digital Economy.

One might be concerned the a firm's adoption of blockchains is correlated broadly with digitization and automation. Firms with on-chain data may simultaneously use industrial robots and IoTs, hire more STEM employees, or cloud storage and computation, which may be the drivers for fundamentals and are correlated with blockchain usage. To address the concern of omitted variables, we merge our data set with another proprietary dataset containing information of the firms' usage of these variables related to the digital economy. After the merge, we have more than 2900 firm-quarter observations. We regress firm outcomes on BDG and other controls including these digital economy variables.

Table 16 reports the results. We observe that IoT usage growth ($IOTG_q$), growth in the number of industrial robots ($ROBG_q$), and growth in the number of STEM employees ($STEMG_q$) have some predictive power on firm fundamentals. Cloud data growth (CDG_q), turns out to be a powerful predictor for firm fundamentals and stock returns, consistent with the fact that data are by-products of economic activities and could help firms to improve their products and services as well. That said, on-chain data brings a greater degree of openness than cloud-based data and BDG has incremental predictive power over other nowcasters for all firm fundamentals except for earnings surprise in the next quarter as measured by CAR.

6.4 On-chain Data by Category

The on-chain data used in our study come labeled into seven categories by the data provider: Operation, Financials, Human Resources, Marketing, Supply Chain, and others. This categorization allows us to further attribute the predictive power of BDG and analyze the channels that BDG could operate through. To this end, we define the corresponding BDG variables for each data category and examine their nowcasting and forecasting power on firm fundamentals and stock returns.

We report the findings in Table 17. On-chain data growth in the Operation and Financials category drive most of the BDG impact discussed earlier. Since information on firm operations and financials have to be disclosed even without blockchains, whereas disclosure on other categories such as Marketing and IT is more discretionary, we are assured that the timeliness and transparency of blockchain-based disclosure do play an important role.

6.5 On-chain Data in Ecosystems

On-chain data in supply-chain ecosystems. BDG may have implications beyond individual firms because many participants of blockchain ecosystems may have access to on-chain information and benefit from it. As an illustration, we test whether customers' BDG is value-relevant and informative of the suppliers' fundamentals and vice versa. We repeat the fundamental and return predictive exercises in the supply chain. The results are reported in Table 18. For each supplier, we aggregate the BDG of its customers; for each customer, we aggregate the BDG of its suppliers. We then use the resulting values as predictors and nowcasters.

Indeed, we find that customers' BDG can nowcast and forecast firm fundamentals and earnings surprises of suppliers, and vice versa. The forecasting power, if anything, is as good as using the BDG of the firm in question. What this reflects is the interoperability and community-based nature of blockchain platforms.

Blockchain peer data growth. In blockchains, all the nodes constitute the ecosystem. Similar to what we observe in supply chains, we expect that when the ecosystem grows, a firm could also improve. We define blockchain peers to a focal firm as other firms in the same blockchain network. We then construct equal weidhgted blockchain peer data growth (BPDG) as a new independent variable and tests its nowcasting and forecasting power of the focal firm's fundamentals and stock returns. The results are reported in Table 19. We document that BPDG positively predicts focal firms' fundamentals in the current quarter and the next quarter. BPDG also predicts earnings surprises with coefficients comparable to BDG. Finally, BPDG-sorted long-short value-weighted portfolios generate monthly risk-adjusted alpha of over 50 basis points.

7 Conclusions

Despite the explosive growth of cryptocurrencies and decentralized finance, whether the underlying technology adds significant value and will thus sustain broad adoption remains unclear. Using proprietary data on firm-level blockchain records from 2015 to 2021, we conduct the first large-sample study linking blockchains to firm fundamentals and asset valuation in China where cryptocurrencies are completely banned. Because information on firm fundamentals is released with a delay, BDG also serves as a powerful nowcaster. We find that year-over-year quarterly blockchain data growth (BDG) contains value-relevant information for predicting assets growth, sales growth, ROA, standardized unexpected earnings (SUE), and innovation outcomes measured through patents. BDG also forecasts stock returns, especially around future earnings announcements, with a long-short BDG-sorted portfolio generating a 10.56% risk-adjusted return annually. BDG is also superior compared to other nowcasters such as online product search, parking lot occupancy, and customer ratings. The results are driven by on-chain data concerning firm financials and operations and are robust across industries, regions, and in international samples. In our sample, we do not observe declines in BDG's predictive power. How firms change their reporting on-chain or through conventional channels in response to such predictability constitutes interesting future research.[13]

We further discuss the underlying economic channels through which blockchains as distributed ledgers and databases helps the firms to better build trust, security, transparency, and efficiency for data storage, disclosure, and information exchange, which ultimately improves business profits and market valuations. We propose strategies that aim at identifying the impact of blockchains on firm-level outcomes and find evidence consistent with real-life use cases and heterogeneity analyses that reveal firms with greater information asymmetry, lower disclosure quality, and less public trust benefit more from blockchain adoption and on-chain data growth. As blockchain technology advances and adoption grows, the value generation and investment relevance of on-chain data will only increase over time. While payment functionality likely amplifies the benefits of the technology, we can, as the evidence shows, have blockchains without crypto.

References

Abadi, Joseph and Markus Brunnermeier, "Blockchain economics," Technical Report, National Bureau of Economic Research 2018.

Agarwal, S., W. Qian, and X. Zou, "Disaggregated sales and stock returns.," *Management Science.*, 2021.

Babina, Tania, Greg Buchak, and Will Gornall, "Customer Data Access and Fintech Entry: Early Evidence from Open Banking," *Available at SSRN*, 2022.

Barwick, Panle Jia, Yongheng Deng, Xinfei Huang, Shanjun Li, and Teng Li, "COVID-19 Pandemic and Unemployment: Evidence from Mobile Phone Data in China 1," *Statistics*, 2020, *134*, 135.

[13] It is possible that firms manipulate their reports and a signal jamming equilibrium ensues. That said, firms often provide auditors access to their blockchain data, and misreporting may be difficult to do.

Basu, Soumya, David Easley, Maureen O'Hara, and Emin Sirer, "StableFees: A Predictable Fee Market for Cryptocurrencie," *Available at SSRN 3318327*, 2019.

Bech, Morten L and Rodney Garratt, "Central bank cryptocurrencies," *BIS Quarterly Review September*, 2017.

Berg, Florian, Kornelia Fabisik, and Zacharias Sautner, "Rewriting history II: The (un) predictable past of ESG ratings," *European Corporate Governance Institute–Finance Working Paper*, 2020, *708* (2020), 10–2139.

Berg, Tobias, Andreas Fuster, and Manju Puri, "FinTech Lending," Technical Report, National Bureau of Economic Research 2021.

Bernard, Victor L. and Jacob K. Thomas, "Post-earnings-announcement drift: delayed price response or risk premium?," *Journal of Accounting Research*, 1989, *27*, 1–36.

Biais, Bruno, Christophe Bisiere, Matthieu Bouvard, and Catherine Casamatta, "The blockchain folk theorem," *Review of Financial Studies*, 2019, *32* (5), 1662–1715.

Bruce, Andrew, "The Near-Term Future Of Blockchain: Tracking Carbon Offsets," *Forbes*, 2021, *16 Aug.*

Cahill, Daniel, Dirk G Baur, Zhangxin Frank Liu, and Joey W Yang, "I am a blockchain too: How does the market respond to companies' interest in blockchain?," *Journal of Banking & Finance*, 2020, *113*, 105740.

Cao, Sean, Lin William Cong, and Baozhong Yang, "Financial reporting and blockchains: Audit pricing, misstatements, and regulation," *Working Paper*, 2018.

_ , _ , Meng Han, Qixuan Hou, and Baozhong Yang, "Blockchain architecture for auditing automation and trust building in public markets," *Computer*, 2020, *53* (7), 20–28.

Capponi, Agostino and Ruizhe Jia, "The Adoption of Blockchain-based Decentralized Exchanges: A Market Microstructure Analysis of the Automated Market Maker," *Available at SSRN 3805095*, 2021.

Chang, Ran and Zhi Da, "Nowcasting firms' fundamentals: Evidence from the cloud," *Working Paper*, 2022.

Chen, Mark A, Shuting Sophia Hu, Joanna Xiaoyu Wang, and Qinxi Wu, "The Effects of Technological Disruption on Suppliers and Customers: Evidence from Blockchain Laws," *Available at SSRN 3915895*, 2021.

Chen, X., Q. Cheng, and T. Luo, "The economic value of blockchain applications: Early evidence from asset-backed securities," 2021.

Cheng, Stephanie F, Gus De Franco, Haibo Jiang, and Pengkai Lin, "Riding the blockchain mania: Public firms' speculative 8-K disclosures," *Management Science*, 2019, *65* (12), 5901–5913.

Chiu, Jonathan, Seyed Mohammadreza Davoodalhosseini, Janet Hua Jiang, and Yu Zhu, "Bank market power and central bank digital currency: Theory and quantitative assessment," *Available at SSRN 3331135*, 2019.

Chiu, Tzu-Ting, "Blockchain Adoption and Investment Sensitivity to Stock Price." PhD dissertation, Norwegian School of Economics 2021.

Cong, Lin William and Simon Mayer, "The Coming Battle of Digital Currencies," *Available at SSRN 3992815*, 2021.

__ **and Zhiguo He**, "Blockchain disruption and smart contracts," *Review of Financial Studies*, 2019, *32* (5), 1754–1797.

__ **, Beibei Li, and Qingquan Tony Zhang**, "Alternative data in fintech and business intelligence," in "The Palgrave Handbook of FinTech and Blockchain," Springer, 2021, pp. 217–242.

__ **, Danxia Xie, and Longtian Zhang**, "Knowledge accumulation, privacy, and growth in a data economy," *Management Science*, 2021, *67* (10), 6480–6492.

__ **, George Andrew Karolyi, Ke Tang, and Weiyi Zhao**, "Value Premium, Network Adoption, and Factor Pricing of Crypto Assets," *Network Adoption, and Factor Pricing of Crypto Assets (December 2021)*, 2021.

__ **, Wayne R Landsman, Edward L Maydew, and Daniel Rabetti**, "Tax-Loss Harvesting with Cryptocurrencies," *Available at SSRN 4033617.*

__ **, Wenshi Wei, Danxia Xie, and Longtian Zhang**, "Endogenous growth under multiple uses of data," *Journal of Economic Dynamics and Control*, 2022, p. 104395.

__ **, Xi Li, Ke Tang, and Yang Yang**, "Crypto wash trading," *arXiv preprint arXiv:2108.10984*, 2021.

__ **, Ye Li, and Neng Wang**, "Token-based platform finance," *Journal of Financial Economics*, 2021.

__ **, __ , and __** , "Tokenomics: Dynamic adoption and valuation," *Review of Financial Studies*, 2021, *34* (3), 1105–1155.

__ **, Zhiguo He, and Jiasun Li**, "Decentralized mining in centralized pools," *Review of Financial Studies*, 2021, *34* (3), 1191–1235.

__ **, Zhiheng He, and Ke Tang**, "Staking, Token Pricing, and Crypto Carry," *Working Paper*, 2022.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, "In search of attention," *Journal of Finance*, 2011, *66*, 1461–1499.

Easley, David, Maureen O'Hara, and Soumya Basu, "From mining to markets: The evolution of bitcoin transaction fees," *Journal of Financial Economics*, 2019, *134* (1), 91–109.

Fama, Eugene F. and James D. MacBeth, "Risk, return, and equilibrium: Empirical tests," *Journal of Political Economy*, 1973, *81*, 607–636.

__ **and Kenneth R. French**, "Common risk factors in the returns on stocks and bonds.," *Journal of Financial Economics*, 1993, *33* (1), 3–56.

__ **and** __ , "Dissecting anomalies," *Journal of Finance*, 2008, *63* (4), 1653–1678.

__ **and** __ , "A five-factor asset pricing model," *Journal of Financial Economics*, 2015, *116* (1), 1–22.

Gans, Joshua S, Hanna Halaburda et al., "Some economics of private digital currency," *Economic analysis of the digital economy*, 2015, pp. 257–276.

Gkritsi, Eliza, "China Releases First Carbon Offset on Ant Group Blockchain: Report," *CoinDesk*, 2021, *14 Sept.*

Goldstein, I. and L. Yang, "Good disclosure, bad disclosure.," *Journal of Financial Economics*, 2019, *131* (1), 118–138.

Green, T. C., R. Huang, Q. Wen, and D Zhou, "Crowdsourced employer reviews and stock returns," *Journal of Financial Economics*, 2019, *134* (1), 236–251.

Griffin, John M and Amin Shams, "Is Bitcoin really untethered?," *Journal of Finance*, 2020, *75* (4), 1913–1964.

Guo, Feng, Stephanie Walton, Patrick R Wheeler, and Yiyang Zhang, "Early Disruptors: Examining the Determinants and Consequences of Blockchain Early Adoption," *Journal of Information Systems*, 2021, *35* (2), 219–242.

Harvey, Campbell R, Ashwin Ramachandran, and Joey Santoro, *DeFi and the Future of Finance*, John Wiley & Sons, 2021.

Hastings, Marcella, Brett Hemenway Falk, and Gerry Tsoukalas, "Privacy-preserving network analytics," *Available at SSRN 3680000*, 2021.

He, Zhiguo, Jing Huang, and Jidong Zhou, "Open banking: credit market competition when borrowers own the data," Technical Report, National Bureau of Economic Research 2020.

Hirshleifer, David, Po-Hsuan Hsu, and Dongmei Li, "Innovative efficiency and stock returns," *Journal of Financial Economics*, 2013, *107*, 632–654.

_ , _ , and _ , "Innovative originality, profitability, and stock returns," *Review of Financial Studies*, 2018, *31* (7), 2553–2605.

Hou, Kewei, Xue Chen, and Zhang Lu, "Digesting anomalies: An investment approach," *Review of Financial Studies*, 2015, *28*, 650–705.

Howell, Sabrina T, Marina Niessner, and David Yermack, "Initial coin offerings: Financing growth with cryptocurrency token sales," *Review of Financial Studies*, 2020, *33* (9), 3925–3974.

Huang, J, "The customer knows best: The investment value of consumer opinions," *Journal of Financial Economics*, 2018, *128* (1), 164–182.

Huberman, Gur, Jacob D Leshno, and Ciamac Moallemi, "Monopoly without a monopolist: An economic analysis of the bitcoin payment system," *Review of Economic Studies*, 2021, *88* (6), 3011–3040.

Iyengar, Garud, Fahad Saleh, Jay Sethuraman, and Wenjun Wang, "Blockchain adoption in a supply chain with market power," *Available at SSRN 3950580*, 2021.

Jegadeesh, Narasimhan and Sheridan Titman, "Returns to buying winners and selling losers: Implications for stock market efficiency," *Journal of Finance*, 1993, *48*, 65–91.

John, Kose, Thomas J Rivera, and Fahad Saleh, "Economic implications of scaling blockchains: Why the consensus protocol matters," *Available at SSRN 3750467*, 2020.

Jones, Charles I and Christopher Tonetti, "Nonrivalry and the Economics of Data," *American Economic Review*, 2020, *110* (9), 2819–58.

Katona, Z., M. Painter, P. N. Patatoukas, and J Zeng, "On the capital market consequences of alternative data: Evidence from outer space," *In 9th Miami Behavioral Finance Conference*, 2018.

Kim, Jinhwan and Kristen Valentine, "The innovation consequences of mandatory patent disclosures," *Journal of Accounting and Economics*, 2021, *71* (2-3), 101381.

Lakhani, Karim R and M Iansiti, "The truth about blockchain," *Harvard Business Review*, 2017, *95* (1), 119–127.

Lehar, Alfred and Christine A Parlour, "Miner collusion and the bitcoin protocol," *Available at SSRN 3559894*, 2020.

Li, Tao, Donghwa Shin, and Baolian Wang, "Cryptocurrency pump-and-dump schemes," *Available at SSRN 3267041*, 2021.

Liu, Jianan, Robert F. Stambaugh, and Yu Yuan, "Size and value in China," *Journal of Financial Economics*, 2019, *134* (1), 48–69.

Liu, Yukun, Aleh Tsyvinski, and Xi Wu, "Common risk factors in cryptocurrency," Technical Report, National Bureau of Economic Research 2019.

Liu, Zhuang, Michael Sockin, and Wei Xiong, "Data privacy and temptation," Technical Report, National Bureau of Economic Research 2020.

Livnat, Joshua and Richard R. Mendenhall, "Comparing the post-earnings announcement drift for surprises calculated from analyst and time series forecasts," *Journal of Accounting Research*, 2006, *44*, 177–205.

Lyandres, Evgeny, Berardino Palazzo, and Daniel Rabetti, "Do tokens behave like securities? An anatomy of initial coin offerings," *SSRN Electronic Journal*, 2019.

Malinova, Katya and Andreas Park, "Tokenomics: when tokens beat equity," *Available at SSRN 3286825*, 2018.

Newey, Whitney K. and Kenneth D. West, "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, 1987, *55*, 703–708.

Pawczuk, Linda, Rob Massey, and David Schatsky, "Breaking Blockchain Open: Deloitte's 2018 Global Blockchain Survey," *Deloitte Consulting: Denver, CO, USA*, 2018.

Rajgopal, S., M. Venkatachalam, and S Kotha, "The value relevance of network advantages: The case of e–commerce firms.," *Journal of Accounting Research*, 2003, *41* (1), 135–162.

Saleh, Fahad, "Blockchain without waste: Proof-of-stake," *Review of Financial Studies*, 2021, *34* (3), 1156–1190.

Tang, Huan, "The value of privacy: Evidence from online borrowers," *Available at SSRN 3880119*, 2019.

Zhu, C., "Big data as a governance mechanism.," *Review of Financial Studies*, 2019, *32* (5), 2021–2061.

Figure 1: Average Blockchain Data Size and Number of Sample Firms

Panel A shows the number of firms in our sample by quarter. Panel B shows the average quarterly on-chain data size (in terabytes) for each firm. Panel C shows the on-chain data as a percentage of each firm's overall "cloud-based" data (including blockchain data).

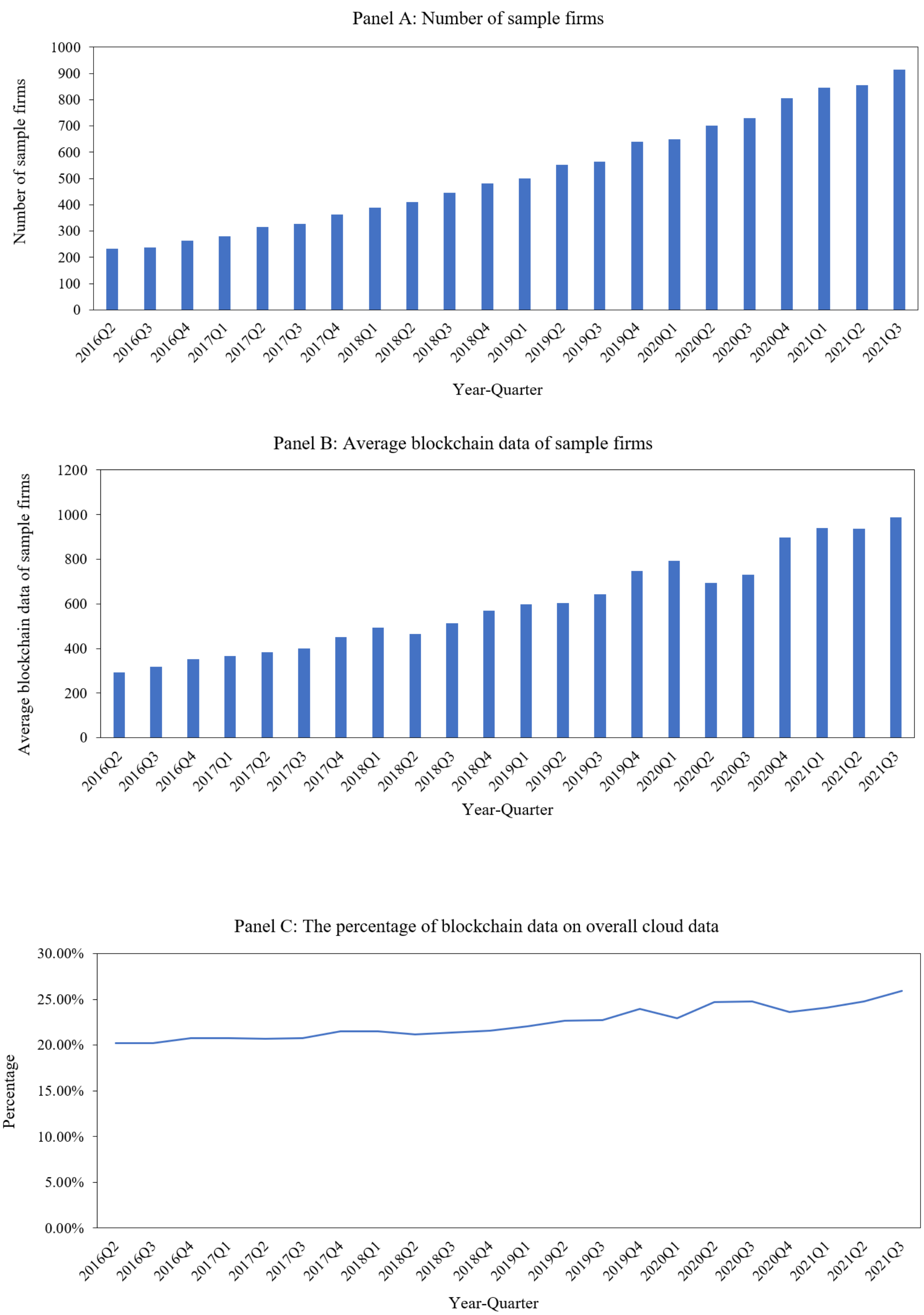

Table 1: Summary Statistics of Variables

Our sample consists of all publicly listed firms on Shanghai and Shenzhen stock exchanges from Q2 2016 to Q3 2021, excluding stocks having become public within the past 12 months or having less than 15 days of trading records during the most recent month.

Panel A shows firm characteristics. BD is the amount of blockchain data of a firm in a quarter. RET_{t+1} is the one-month-ahead return. SIZE is the firm's market capitalization computed as the logarithm of the market value of the firm's outstanding equity. BM is the logarithm of the firm's book value of equity divided by its market capitalization, where the BM ratio is computed following Fama and French (2008). Firms with negative book values are excluded from the analysis. ROA is the quarterly operating income scaled by lagged assets. LEV is the quarterly sum of long-term debt and short-term borrowing scaled by total assets. Short-term reversal (STR) is the stock's lagged-one monthly return. MOM is the stock's cumulative return from the start of lagged-twelve month to the end of lagged-two month (skipping the STR month), following Jegadeesh and Titman (1993). PPE Growth (PG) is the year-over-year quarterly growth in property, plant, and equipment scaled by total assets. Intangible Growth (IG) is the year-over-year quarterly growth in intangible assets scaled by total assets. TO is the quarterly turnover computed as the number of shares traded divided by the total number of shares outstanding in a quarter. ILLIQ is the quarterly illiquidity measure computed as the absolute daily return divided by daily dollar trading volume, averaged in a quarter. IVOL is the idiosyncratic volatility defined as the standard deviation of daily residuals estimated from the regression of daily excess stock returns on the daily market, size, and value factors of Fama and French (1993) in a quarter. SUE is the standardized unexpected earnings defined as actual earnings in the current quarter minus earnings 4 quarters ago, scaled by stock price in the current quarter following Livnat and Mendenhall (2006). ANA is defined as the number of analysts following the firm in a quarter, and IO is the percentage of tradable shares held by institutional investors in a quarter.

Panel B shows characteristics of firm fundamentals, earnings surprise, innovation performance. AG is the quarterly growth of total assets. SG is the quarterly growth of sales. ROA is the quarterly operating income scaled by lagged assets. SUE is the standardized unexpected earnings defined as actual earnings in the current quarter minus earnings 4 quarters ago, scaled by stock price in the current quarter following Livnat and Mendenhall (2006). CAR is the cumulative abnormal returns over the three-day window surrounding the earnings announcement. Abnormal return is calculated as the raw daily return minus the daily return on size and market-to-book matched portfolio as in Livnat and Mendenhall (2006). PA is the log of one plus quarterly number of patents applied of the firm. PG is the log of one plus quarterly number of patents granted of the firm.

Panel C shows statistics of BDG and other nowcasters. BDG is the annual growth of the amount of blockchain data of a firm in a quarter relative to that in the same quarter last year. SEAG is the year-over-year quarterly growth of search volume for firms' products. APPG is the year-over-year quarterly growth of firms' App visiting volume. CUSG is the year-over-year quarterly growth of customer product ratings of firms. EMPG is the year-over-year quarterly growth of employer ratings of firms. CARG is the year-over-year quarterly growth of number of cars in parking lots of firms. SPEG is the year-over-year quarterly growth of credit card spending to products and services of firms. All variables are winsorized at the 1st and 99th percentiles to mitigate the effect of outliers.

Panel D shows blockchain data by category as a percentage of total on-chain data. 5% percentile, median, and 95% percentile of each variable are shown too.

Panel A: Firm characteristics					
	Mean	SD	P5	P50	P95
BD	599.122	6.037	281.981	540.041	1155.915
SIZE	22.457	0.954	21.628	22.495	23.229
BM	0.445	0.253	0.099	0.352	0.932
ROA	1.351	1.760	-0.984	1.386	5.117
LEV	0.179	0.168	0.000	0.172	0.545
STR	0.010	0.120	-0.492	0.011	0.914
MOM	0.075	0.582	-0.421	0.076	1.000
PG	0.019	0.012	0.002	0.016	0.042
IG	0.042	0.035	0.004	0.033	0.109
TO	0.466	1.039	0.057	0.202	8.127
ILLIQ	0.154	0.473	0.010	0.041	13.736
IVOL	0.020	0.009	0.011	0.054	0.083
SUE	0.135	2.293	-11.024	0.123	5.447
ANA	7.455	8.522	0.000	4.986	27.672
IO	6.589	9.090	0.001	2.599	27.624

Panel B: Characteristics of firm fundamentals, earnings surprise, innovation performance					
	Mean	SD	P5	P50	P95
AG	0.179	0.319	-0.230	0.122	0.783
SG	0.175	0.291	-0.093	0.106	0.725
ROA	1.455	1.760	-1.041	1.267	5.060
SUE	0.124	2.374	-10.963	0.121	5.761
CAR	0.300	6.715	-10.683	-0.379	13.616
PA	1.176	1.653	0.000	0.000	1.991
PG	1.040	1.744	0.000	0.000	1.468

Panel C: BDG and other nowcasters					
	Mean	SD	P5	P50	P95
BDG	0.091	0.228	-0.143	0.072	0.362
SEAG	0.127	0.506	-0.487	0.103	1.185
APPG	0.081	0.421	-0.339	0.062	0.795
EMPG	0.061	0.678	-0.458	0.041	1.231
CUSG	0.072	0.798	-0.688	0.100	1.323
CARG	0.107	0.451	-0.461	0.090	1.058
SPEG	0.072	0.518	-0.661	0.063	0.995

Panel D: On-chain data by category (as percentages of total on-chain data)					
	Mean	SD	P5	P50	P95
Operation	0.253	0.298	0.023	0.267	0.786
Financials	0.236	0.308	0.024	0.233	0.697
Human Resources	0.100	0.111	0.010	0.096	0.285
Marketing	0.102	0.109	0.010	0.099	0.323
IT	0.091	0.110	0.011	0.097	0.279
Supply Chain	0.120	0.118	0.010	0.124	0.324
Others	0.099	0.131	0.010	0.094	0.285

Table 2: BDG and Firm Fundamentals

This table reports the results on the regressions of firm fundamentals measured in quarter $q + 1$ or quarter $q + 2$ on the BDG in quarter q+1 and other control variables in quarter q. The dependent variables include return on assets (ROA), growth of total assets (AG), growth of sales (SG), the log of one plus quarterly number of patents applied (PA), and the log of one plus quarterly number of patents granted (PG). We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and one standard deviation. The control variables are from Panel A of Table 1. The t-statistics of robust standard errors clustered at the industry and year-quarter levels are reported in parentheses. Coefficients marked with $*$, $**$, and $***$ are significant at the 10%, the 5%, and the 1% level, respectively. The sample period is from second quarter of 2016 to third quarter of 2021.

	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)
BM_q	-0.678**	-0.509*	-0.067***	-0.140***	-0.020	-0.055**	-0.056	-0.089**	-0.012	-0.048
	(-2.24)	(-1.68)	(-3.05)	(-3.24)	(-0.91)	(-2.29)	(-1.42)	(-2.05)	(-0.53)	(-1.57)
ROA_q	4.751	3.644*	0.508***	1.351***	0.433***	0.399**	0.869***	1.426***	0.584***	0.819***
	(1.24)	(1.96)	(4.23)	(3.77)	(3.24)	(2.37)	(6.53)	(6.51)	(4.44)	(3.13)
LEV_q	-1.092***	-1.190***	-0.028	-0.063	-0.026	-0.115**	-0.020	-0.050	-0.025	-0.051
	(-3.92)	(-5.45)	(-0.85)	(-1.60)	(-1.06)	(-2.18)	(-0.47)	(-0.93)	(-0.65)	(-1.37)
PG_q	-0.432**	-0.281	0.008	0.046	-0.019	0.028	0.012	0.080	-0.023	0.024
	(-2.13)	(-0.8)	(0.31)	(0.95)	(-0.32)	(0.34)	(0.41)	(1.10)	(-0.46)	(0.40)
IG_q	-0.052	0.466	0.104	0.310	0.050	0.090	0.060	0.226	0.039	0.063
	(-0.06)	(0.69)	(0.94)	(1.24)	(0.46)	(0.52)	(0.52)	(0.87)	(0.28)	(0.27)
SUE_q	-0.006	0.107***	0.017***	0.053***	0.010*	0.015*	0.012**	0.032***	0.006	0.012
	(-0.29)	(3.07)	(3.93)	(5.49)	(1.89)	(1.78)	(2.35)	(4.91)	(1.10)	(1.14)
SIZE	0.105	0.015	-0.006	-0.033	-0.012*	-0.047***	-0.010	-0.046***	-0.016*	-0.068***
	(1.26)	(0.25)	(-0.57)	(-1.53)	(-1.72)	(-3.26)	(-0.82)	(-2.60)	(-1.89)	(-3.86)
STR	0.374*	0.049	-0.020	-0.089**	-0.037**	-0.129***	-0.034	-0.142***	-0.052***	-0.2***
	(1.67)	(0.32)	(-0.79)	(-2.35)	(-2.02)	(-3.39)	(-1.17)	(-2.63)	(-2.95)	(-5.52)
MOM	0.164***	0.154***	0.021***	0.033***	0.025***	0.041***	0.010**	0.022**	0.015**	0.023*
	(3.47)	(3.68)	(2.72)	(2.93)	(3.43)	(3.80)	(2.33)	(2.09)	(2.40)	(1.68)
TO	0.049	0.051	0.006	0.011	0.008	0.015	0.008	0.014	0.013	0.014
	(0.81)	(0.90)	(0.80)	(0.70)	(0.69)	(0.81)	(1.06)	(1.17)	(0.99)	(1.09)
ILLIQ	2.575	2.536	0.545	0.754	0.233	0.338	0.239	0.464	0.201	0.241
	(0.14)	(0.17)	(0.42)	(0.47)	(0.21)	(0.18)	(0.32)	(0.31)	(0.13)	(0.15)
IVOL	-3.883***	-2.156***	0.082	0.416***	-0.101	0.142	0.079*	0.531***	-0.139**	0.212*
	(-7.34)	(-4.61)	(1.41)	(3.53)	(-1.59)	(0.97)	(1.81)	(3.82)	(-2.14)	(1.76)
ANA	0.002	0.003	0.000	0.000	-0.001**	-0.001*	0.000	0.000	-0.001***	-0.001*
	(0.67)	(0.78)	(-0.71)	(-0.45)	(-2.42)	(-1.71)	(-1.08)	(-0.60)	(-4.49)	(-1.88)
IO	0.012***	0.009***	0.001***	0.002***	0.001***	0.002***	0.001***	0.001*	0.001	0.001**
	(3.43)	(3.91)	(4.05)	(3.71)	(2.66)	(3.56)	(2.62)	(1.81)	(1.52)	(2.05)
AG_q			0.294***	0.158***						
			(4.83)	(3.92)						
SG_q					0.609***	0.401***				
					(10.85)	(5.44)				
PA_q							0.322***	0.164***		
							(4.84)	(3.42)		
PG_q									0.197***	0.112***
									(3.00)	(2.90)
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

Table 3: Nowcasting and Forecasting Firm Fundamentals

This table reports the results on the regressions of firm fundamentals measured in quarter q+1 or quarter q+2 on the BDG and nowcasters in quarter q+1 and other control variables in quarter q. The dependent variables include return on assets (ROA), growth of total assets (AG), growth of sales (SG), the log of one plus quarterly number of patents applied (PA), and the log of one plus quarterly number of patents granted (PG). The nowcasters include the year-over-year quarterly growth of search volume for firms' products ($SEAG_q$), the year-over-year quarterly growth of firms' App visiting volume ($APPG_q$), the year-over-year quarterly growth of customer product ratings of firms ($CUSG_q$), the year-over-year quarterly growth of employer ratings of firms ($EMPG_q$), the year-over-year quarterly growth of number of cars in parking lots of firms ($CARG_q$), and the year-over-year quarterly growth of credit card spending to products and services of firms ($SPEG_q$). We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and one standard deviation. The control variables are from Panel A of Table 1. The t-statistics of robust standard errors clustered at the industry and year-quarter levels are reported in parentheses. Coefficients marked with $*$, $**$, and $***$ are significant at the 10%, the 5%, and the 1% level, respectively. The sample period is from second quarter of 2016 to third quarter of 2021.

	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}	0.524***	0.408***	0.179***	0.153***	0.05***	0.037***	0.234***	0.18***	0.106***	0.081**
	(3.24)	(2.99)	(3.34)	(2.77)	(2.85)	(2.73)	(2.76)	(2.95)	(3.29)	(2.37)
$SEAG_{q+1}$	0.23***	0.217**	0.126***	0.088**	0.027**	0.022*	0.006	0.004	0.012	0.012
	(2.68)	(2.34)	(3.23)	(2.04)	(2.10)	(1.95)	(0.05)	(0.03)	(0.24)	(0.22)
$APPG_{q+1}$	0.139	0.106	0.059	0.056	0.013	0.010	0.062	0.046	0.033	0.027
	(0.99)	(0.69)	(1.36)	(0.88)	(1.30)	(0.72)	(0.94)	(0.74)	(0.61)	(0.53)
$EMPG_{q+1}$	0.202*	0.12	0.075**	0.062*	0.016*	0.015	0.105*	0.085	0.062	0.046
	(1.75)	(1.47)	(2.35)	(1.82)	(1.83)	(1.29)	(1.70)	(1.59)	(1.55)	(1.16)
$CUSG_{q+1}$	0.130	0.099	0.047	0.037	0.018	0.013	0.038	0.035	0.014	0.015
	(1.11)	(0.77)	(1.68)	(1.08)	(1.23)	(1.02)	(0.73)	(0.62)	(0.64)	(0.46)
$CARG_{q+1}$	0.295***	0.239**	0.132***	0.107**	0.029***	0.018**	0.071	0.063	0.048	0.032
	(2.59)	(2.40)	(2.86)	(2.41)	(2.80)	(2.38)	(1.40)	(1.21)	(1.58)	(0.90)
$SPEG_{q+1}$	0.064	0.047	0.027	0.023	0.006	0.005	0.154**	0.107**	0.063**	0.067*
	(0.61)	(0.57)	(0.57)	(0.52)	(0.64)	(0.38)	(2.58)	(2.09)	(2.11)	(1.80)
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	3643	3518	3308	3194	3855	3722	2602	2513	2602	2513
Adj. R2	0.68	0.55	0.61	0.49	0.43	0.34	0.35	0.25	0.26	0.25

Table 4: BDG and Earnings Surprise

This table reports the results on the regressions of earnings surprise measured in quarter $q+1$ or quarter $q+2$ on the BDG in quarter $q+1$ and other control variables in quarter q. The dependent variables include standardized unexpected earnings (SUE) and earnings announcement abnormal returns (CAR). We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and one standard deviation. The control variables are from Panel A of Table 1. The t-statistics of robust standard errors clustered at the industry and year-quarter levels are reported in parentheses. Coefficients marked with $*$, $**$, and $***$ are significant at the 10%, the 5%, and the 1% level, respectively. The sample period is from second quarter of 2016 to third quarter of 2021.

	SUE_{q+1}	SUE_{q+2}	CAR_{q+1}	CAR_{q+2}
BDG_{q+1}	0.251***	0.187***	1.839***	1.437**
	(4.19)	(3.16)	(2.72)	(2.04)
BM_q	-0.163	-0.029	0.019	0.149
	(-0.40)	(-0.10)	(0.08)	(0.45)
ROA_q	5.824*	6.139***	-6.949***	0.86
	(1.77)	(3.38)	(-2.63)	(0.38)
LEV_q	-0.025	-0.056	-0.601	-0.783*
	(-0.10)	(-0.34)	(-1.52)	(-1.92)
PG_q	-0.349	-0.002	-0.071	0.253
	(-1.01)	(-0.01)	(-0.15)	(0.47)
IG_q	1.672	0.964	-0.614	0.365
	(1.52)	(1.29)	(-0.29)	(0.21)
SUE_q	0.307***	0.282***	-0.122	0.058
	(3.90)	(4.34)	(-1.47)	(1.47)
SIZE	-0.018	-0.158	-0.35**	-0.379
	(-0.36)	(-1.52)	(-2.41)	(-1.61)
STR	-0.079	-0.57**	-0.911**	-1.202**
	(-0.34)	(-2.01)	(-2.24)	(-2.15)
MOM	0.195**	0.151**	-0.037	0.004
	(2.41)	(2.57)	(-0.37)	(0.07)
TO	0.049	0.042	-0.011	0.001
	(0.82)	(0.71)	(-0.10)	(0.01)
ILLIQ	3.603	4.542	-3.609	0.743
	(0.15)	(0.30)	(-0.22)	(0.03)
IVOL	-2.656***	-0.014	-0.454	1.922*
	(-5.44)	(-0.02)	(-0.68)	(1.89)
ANA	-0.009***	-0.007**	0.001	0.001
	(-3.03)	(-2.46)	(0.05)	(0.10)
IO	0.004**	0.006*	0.010**	0.002
	(1.99)	(1.87)	(2.02)	(0.46)
Industry FE	Y	Y	Y	Y
Year-Quarter FE	Y	Y	Y	Y
N	10922	10806	11152	11033
Adj. R2	0.41	0.35	0.07	0.07

Table 5: Nowcasting and Forecasting Earnings Surprise

This table reports the results on the regressions of earnings surprise measured in quarter $q+1$ or quarter $q+2$ on the BDG and nowcasters in quarter q+1 and other control variables in quarter q. The dependent variables include standardized unexpected earnings (SUE) and earnings announcement abnormal returns (CAR). The nowcasters are defined in Table 3. We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and one standard deviation. The control variables are from Panel A of Table 1. The t-statistics of robust standard errors clustered at the industry and year-quarter levels are reported in parentheses. Coefficients marked with $*$, $**$, and $***$ are significant at the 10%, the 5%, and the 1% level, respectively. The sample period is from second quarter of 2016 to third quarter of 2021.

	SUE_{q+1}	SUE_{q+2}	CAR_{q+1}	CAR_{q+2}
BDG_{q+1}	0.114**	0.088**	1.518**	1.308**
	(2.56)	(2.08)	(2.23)	(1.99)
$SEAG_{q+1}$	0.079**	0.060	0.642**	0.603
	(2.04)	(1.62)	(2.04)	(1.21)
$APPG_{q+1}$	0.038	0.037	0.410	0.275
	(0.92)	(0.64)	(0.67)	(0.53)
$EMPG_{q+1}$	0.054*	0.045	0.528	0.513
	(1.74)	(1.41)	(1.45)	(1.02)
$CUSG_{q+1}$	0.035	0.031	0.397	0.328
	(0.91)	(0.88)	(0.88)	(0.73)
$CARG_{q+1}$	0.104**	0.065*	0.846*	0.702
	(2.25)	(1.85)	(1.90)	(1.54)
$SPEG_{q+1}$	0.016	0.011	0.170	0.132
	(0.48)	(0.35)	(0.41)	(0.25)
Controls	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Year-Quarter FE	Y	Y	Y	Y
N	3564	3441	3806	3674
Adj. R2	0.49	0.41	0.14	0.12

Table 6: Univariate Portfolio Analysis

Panel A reports the average monthly excess returns and alphas on the value-weighted portfolios of stocks sorted by the BDG. Panel B reports the average monthly excess returns and alphas on the equal-weighted portfolios of stocks sorted by the BDG. At each month t from April 2016 to September 2021, individual stocks of companies are sorted into quintiles based on BDG at quarter $q-1$, and are held for the next one quarter. P1 is the portfolio of stocks with the lowest BDG and P5 is the portfolio of stocks with the highest BDG. L/S is a zero-cost portfolio that buys stocks in quintile 5 (highest BDG) and sells stocks in quintile 1 (lowest BDG). All returns and alphas are expressed in percentage. Excess return is the raw return of the portfolio over the risk-free rate. SR is annualized Sharpe ratio for each portfolio. Alpha is the intercept from a time-series regression of monthly excess returns on the factors of alternative models: China q-factor model (HXZ) based on Hou et al. (2015), China five-factor model (FF5) based on Liu et al. (2019a) China three-factor model (LSY3), and Liu et al. (2019a) China four-factor model (LSY4). EA represents average excess returns in earnings announcement months. Non-EA represents average excess returns in non-earnings announcement months. Newey and West (1987) adjusted t-statistics are given in parentheses. Coefficients marked with $*$, $**$, and $***$ are significant at the 10%, the 5%, and the 1% level, respectively. The sample period is from April 2016 to September 2021.

Panel A: Returns and alphas on value-weighted BDG-sorted quintile portfolios								
Rank	Excess	SR	HXZ	FF5	LSY3	LSY4	EA	Non-EA
P1	0.054	0.061	-0.556**	-0.546**	-0.506***	-0.520**	0.036	0.018
	(0.11)		(-2.33)	(-2.31)	(-2.94)	(-2.54)	(0.08)	(0.05)
P2	0.256	0.269	-0.402**	-0.383*	-0.291**	-0.473**	0.177	0.079
	(0.56)		(-2.57)	(-1.66)	(-2.31)	(-2.35)	(0.45)	(0.23)
P3	0.417	0.448	-0.238	-0.168	-0.285*	-0.272	0.279	0.139
	(1.62)		(-0.57)	(-1.40)	(-1.71)	(-0.74)	(1.24)	(0.70)
P4	0.840**	0.895	-0.116	-0.043	0.088	-0.194	0.552	0.288
	(2.13)		(-0.36)	(-0.16)	(0.43)	(-0.90)	(1.61)	(0.94)
P5	0.934***	1.093	0.132*	0.213	0.158	0.138	0.616***	0.318*
	(3.89)		(1.80)	(1.54)	(1.59)	(0.89)	(2.95)	(1.71)
L/S	0.88***	1.023	0.688***	0.759***	0.664***	0.659***	0.580***	0.300*
	(3.44)		(3.26)	(4.20)	(3.89)	(3.34)	(2.61)	(1.72)
Panel B: Returns and alphas on equal-weighted BDG-sorted quintile portfolios								
Rank	Excess	SR	HXZ	FF5	LSY3	LSY4	EA	Non-EA
P1	0.083	0.095	-0.593***	-0.787***	-0.760***	-0.646***	0.062	0.021
	(0.26)		(-5.05)	(-3.16)	(-4.47)	(-3.10)	(0.22)	(0.09)
P2	0.346***	0.388	-0.167***	-0.668***	-0.270**	-0.593*	0.253**	0.093
	(2.88)		(-4.07)	(-2.85)	(-2.17)	(-1.72)	(2.40)	(1.07)
P3	0.496**	0.520	0.056*	-0.614*	-0.075	-0.463	0.354*	0.143
	(2.40)		(1.77)	(-1.66)	(-1.45)	(-1.13)	(1.95)	(0.93)
P4	0.559***	0.622	0.192	-0.236	0.005	-0.090	0.405***	0.155
	(3.85)		(1.14)	(-0.03)	(0.13)	(-0.82)	(3.17)	(1.45)
P5	1.545***	1.551	0.261*	0.349	0.175	0.186	1.154***	0.391**
	(6.07)		(1.84)	(1.63)	(1.49)	(0.85)	(5.14)	(2.14)
L/S	1.462***	1.504	0.854***	1.137***	0.935***	0.831***	1.092***	0.370**
	(6.08)		(4.31)	(5.57)	(4.22)	(4.26)	(5.15)	(2.15)

Table 7: Fama-MacBeth Cross-Sectional Regressions

This table reports the Fama and MacBeth (1973) cross-sectional regression results. Panel A reports the Fama and MacBeth (1973) cross-sectional regressions of BDG. The BDG and other accounting variables in quarter $q-1$ are matched to stock returns in month t. The monthly price-based variables are calculated using the last non-missing observations prior to each month. The dependent variable is the firm's future raw return in the first two columns, the firm's future excess return over its value-weighted industry peers' return (Column 3), or the firm's future excess return over its value-weighted region peers' return (Column 4). Panel B reports the Fama and MacBeth (1973) cross-sectional regressions of BDG and other nowcasters. The dependent variable is the firm's future raw return. The nowcasters are defined in Table 3. We control for the industry and region fixed effects following the CSRC industry classification and China's province classification. All returns are expressed in percentage. The BDG and other firm-specific characteristics are defined in Panel A of Table 1. All explanatory variables are generated using the last non-missing available observation for each quarter $q-1$. Cross-sectional regressions are run every calendar month, and the time-series standard errors are corrected for heteroskedasticity and autocorrelation. Newey and West (1987) adjusted t-statistics are reported in parentheses. Coefficients marked with *, **, and *** are significant at the 10%, the 5%, and the 1% level, respectively. The sample period is from April 2016 to September 2021.

Panel A: Fama and MacBeth (1973) cross-sectional regressions of BDG

Independent Variables	RET	RET	RET-INDRET	RET-REGRET
BDG	0.407***	0.388***	0.477***	0.357***
	(4.49)	(4.53)	(3.73)	(3.12)
SIZE		-0.560**	-0.579**	-0.533**
		(-2.18)	(-2.11)	(-2.10)
BM		0.302	0.326	0.311
		(0.62)	(0.47)	(0.39)
STR		-2.271**	-1.739***	-1.910***
		(-2.42)	(-2.65)	(-2.70)
MOM		-0.178	-0.172	-0.196
		(-0.93)	(-1.36)	(-0.97)
ROA		12.006***	12.956***	12.139**
		(3.10)	(2.83)	(2.44)
LEV		-0.456	-0.529	-0.511
		(-1.13)	(-1.25)	(-1.44)
PG		-0.490	-0.574	-0.600
		(-0.48)	(-0.50)	(-0.75)
IG		0.550	0.533	0.678
		(0.78)	(0.66)	(0.77)
TO		-0.060	-0.061	-0.072
		(-0.25)	(-0.19)	(-0.20)
ILLIQ		9.371	9.703	6.69
		(0.29)	(0.26)	(0.24)
IVOL		-2.843**	-3.212*	-2.789**
		(-2.21)	(-1.94)	(-2.56)
SUE		0.106**	0.097***	0.117***
		(2.56)	(2.77)	(2.61)
ANA		-0.005	-0.003	-0.004
		(-0.37)	(-0.31)	(-0.37)
IO		0.022**	0.018***	0.023**
		(2.30)	(2.63)	(2.16)
Industry FE	Y	Y	N	Y
Region FE	Y	Y	Y	N
N	34,489	33454	33454	33454
Adj. R2	0.06	0.09	0.07	0.07

Panel B: Fama and MacBeth (1973) cross-sectional regressions of BDG and other nowcasters							
	RET	RET	RET	RET	RET	RET	RET
BDG	0.346***	0.376***	0.343***	0.318***	0.295***	0.365***	0.282***
	(3.03)	(3.85)	(3.50)	(3.52)	(2.61)	(3.21)	(2.78)
SEAG	0.166**						0.126**
	(2.34)						(2.05)
APPG		0.107					0.062
		(1.03)					(0.93)
EMPG			0.127				0.149
			(1.33)				(1.38)
CUSG				0.133			0.102
				(0.90)			(0.93)
CARG					0.244**		0.160**
					(2.57)		(2.07)
SPEG						0.054	0.039
						(0.51)	(0.49)
Controls	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y	Y	Y
N	16727	21745	18400	20073	15054	23418	10383
Adj. R2	0.12	0.08	0.11	0.10	0.09	0.09	0.15

Table 8: BDG-Sorted Portfolios of Firms Adopting Open Blockchains

Panel A reports the average monthly excess returns and alphas on the value-weighted portfolios of stocks sorted by the BDG. Panel B reports the average monthly excess returns and alphas on the equal-weighted portfolios of stocks sorted by the BDG. At each month t from April 2016 to September 2021, individual stocks of companies are sorted into quintiles based on BDG at quarter q-1, and are held for the next one quarter. P1 is the portfolio of stocks with the lowest BDG and P5 is the portfolio of stocks with the highest BDG. L/S is a zero-cost portfolio that buys stocks in quintile 5 (highest BDG) and sells stocks in quintile 1 (lowest BDG). All returns and alphas are expressed in percentage. Excess return is the raw return of the portfolio over the risk-free rate. SR is annualized Sharpe ratio for each portfolio. Alpha is the intercept from a time-series regression of monthly excess returns on the factors of alternative models: China q-factor model (HXZ) based on Hou et al. (2015), China five-factor model (FF5) based on Fama and French (2015), Liu et al. (2019a) China three-factor model (LSY3), and Liu et al. (2019a) China four-factor model (LSY4). EA represents average excess returns in earnings announcement months. Non-EA represents average excess returns in non-earnings announcement months. Newey and West (1987) adjusted t-statistics are given in parentheses. Coefficients marked with *, **, and *** are significant at the 10%, the 5%, and the 1% level, respectively. The sample period is from April 2016 to September 2021.

Panel A: Returns and alphas on value-weighted BDG-sorted quintile portfolios								
Rank	Excess	SR	HXZ	FF5	LSY3	LSY4	EA	Non-EA
P1	0.075	0.090	-0.369	-0.365	-0.331**	-0.331	0.049	0.025
	(0.15)		(-1.53)	(-1.44)	(-1.99)	(-1.64)	(0.11)	(0.06)
P2	0.174	0.194	-0.254	-0.248	-0.184	-0.329	0.124	0.055
	(0.39)		(-1.65)	(-1.02)	(-1.57)	(-1.51)	(0.29)	(0.16)
P3	0.292	0.345	-0.146	-0.116	-0.188	-0.185	0.195	0.093
	(1.09)		(-0.36)	(-0.90)	(-1.09)	(-0.50)	(0.75)	(0.49)
P4	0.536	0.610	-0.075	-0.030	0.056	-0.135	0.359	0.194
	(1.45)		(-0.23)	(-0.11)	(0.27)	(-0.54)	(0.98)	(0.58)
P5	0.624**	0.722	0.080	0.135	0.106	0.093	0.401*	0.213
	(2.39)		(1.24)	(0.96)	(0.99)	(0.59)	(1.89)	(1.05)
L/S	0.549***	0.613	0.449**	0.499***	0.437***	0.425***	0.352**	0.188
	(2.80)		(2.51)	(3.46)	(3.29)	(2.76)	(1.98)	(1.40)
Panel B: Returns and alphas on equal-weighted BDG-sorted quintile portfolios								
Rank	Excess	SR	HXZ	FF5	LSY3	LSY4	EA	Non-EA
P1	0.114	0.125	-0.367***	-0.420*	-0.481***	-0.398**	0.082	0.029
	(0.35)		(-3.34)	(-1.96)	(-3.10)	(-2.12)	(0.30)	(0.13)
P2	0.234*	0.260	-0.115***	-0.444*	-0.185	-0.371	0.165	0.062
	(1.93)		(-2.59)	(-1.79)	(-1.31)	(-1.13)	(1.61)	(0.74)
P3	0.338	0.390	0.036	-0.408	-0.050	-0.297	0.215	0.090
	(1.62)		(1.22)	(-1.15)	(-0.87)	(-0.70)	(1.25)	(0.56)
P4	0.390***	0.394	0.124	-0.147	0.003	-0.055	0.266**	0.106
	(2.64)		(0.78)	(-0.02)	(0.09)	(-0.54)	(2.18)	(0.91)
P5	0.933***	1.048	0.163	0.233	0.111	0.120	0.764***	0.252
	(4.12)		(1.23)	(1.12)	(0.93)	(0.53)	(3.30)	(1.48)
L/S	0.818***	0.892	0.529***	0.653***	0.593***	0.518***	0.682***	0.223*
	(4.64)		(3.24)	(4.48)	(3.52)	(3.25)	(3.89)	(1.67)

Table 9: BDG Predictability Across Heterogeneous Firms

This table presents results of BDG predictability across different firms. We split the stock sample into two equal subsamples based on the market capitalization (Large/Small), the institutional ownership (High/Low), the analyst coverage (High/Low), or ownership (State/Private), or Herfindahl–Hirschman index (HHI) (High/Low). Panel A reports the results on the regressions of firm fundamentals measured in quarter $q+1$ or quarter $q+2$ on the BDG in quarter $q+1$ and other control variables in quarter q across different firms. The dependent variables include return on assets (ROA), growth of total assets (AG), growth of sales (SG), the log of one plus quarterly number of patents applied (PA), and the log of one plus quarterly number of patents granted (PG). Panel B reports the results on the regressions of earnings surprise measured in quarter $q+1$ or quarter $q+2$ on the BDG in quarter $q+1$ and other control variables in quarter q across different firms. The dependent variables include standardized unexpected earnings (SUE) and earnings announcement abnormal returns (CAR). We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and one standard deviation. The control variables are from Panel A of Table 1. The t-statistics of robust standard errors clustered at the industry and year-quarter levels are reported in parentheses. Panel C reports the average monthly long-short excess returns and long-short alphas on the value-weighted portfolios of stocks sorted by the BDG across different firms. Newey and West (1987) adjusted t-statistics are given in parentheses. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively. The sample period is from April 2016 to September 2021.

Panel A: nowcasting and forecasting firm fundamentals across different firms

	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}
Large										
BDG_{q+1}	0.540***	0.342**	0.256***	0.157***	0.051***	0.038**	0.235***	0.144***	0.117***	0.111**
	(3.63)	(2.16)	(3.10)	(2.98)	(3.35)	(2.26)	(3.31)	(2.60)	(3.65)	(2.03)
N	5633	5574	5403	5346	5518	5460	4024	3981	4024	3981
Adj. R2	0.55	0.51	0.47	0.39	0.33	0.33	0.23	0.20	0.16	0.13
Small										
BDG_{q+1}	0.935***	0.679***	0.325***	0.325***	0.103***	0.048***	0.368***	0.203***	0.197***	0.160***
	(5.29)	(4.21)	(6.66)	(4.37)	(5.52)	(2.84)	(5.63)	(3.75)	(5.57)	(3.22)
N	5633	5574	5403	5346	5518	5460	4024	3981	4024	3981
Adj. R2	0.61	0.56	0.52	0.43	0.36	0.36	0.26	0.22	0.17	0.14
High IO										
BDG_{q+1}	0.569***	0.463**	0.243***	0.243***	0.068***	0.037**	0.312***	0.156**	0.165***	0.105***
	(4.78)	(2.43)	(4.35)	(3.30)	(4.06)	(2.12)	(3.66)	(2.40)	(3.87)	(3.12)
N	5633	5574	5403	5346	5518	5460	4024	3981	4024	3981
Adj. R2	0.52	0.48	0.44	0.37	0.31	0.31	0.22	0.19	0.15	0.12
Low IO										
BDG_{q+1}	0.829***	0.654***	0.358***	0.239***	0.065***	0.048***	0.376***	0.186***	0.147***	0.141***
	(5.47)	(2.94)	(5.44)	(3.82)	(3.81)	(3.18)	(5.44)	(3.80)	(3.93)	(3.50)
N	5633	5574	5403	5346	5518	5460	4024	3981	4024	3981
Adj. R2	0.64	0.59	0.54	0.45	0.38	0.38	0.27	0.23	0.18	0.15
High coverage										
BDG_{q+1}	0.736***	0.449***	0.242***	0.158**	0.063***	0.036*	0.247***	0.141**	0.146***	0.119***
	(5.09)	(2.64)	(4.43)	(2.46)	(2.99)	(1.83)	(3.53)	(2.43)	(3.88)	(2.58)
N	5633	5574	5403	5346	5518	5460	4024	3981	4024	3981
Adj. R2	0.49	0.46	0.42	0.35	0.29	0.29	0.21	0.18	0.14	0.11
Low coverage										
BDG_{q+1}	0.937***	0.662***	0.412***	0.239***	0.086***	0.063***	0.433***	0.249***	0.183***	0.146***
	(6.66)	(3.59)	(5.77)	(3.74)	(4.82)	(2.74)	(5.65)	(3.33)	(4.65)	(3.15)
N	5633	5574	5403	5346	5518	5460	4024	3981	4024	3981
Adj. R2	0.67	0.62	0.57	0.47	0.40	0.39	0.28	0.24	0.19	0.15

State enterprises										
BDG_{q+1}	0.480***	0.349**	0.243***	0.221***	0.048***	0.029*	0.277***	0.145**	0.162***	0.106***
	(3.49)	(2.46)	(4.47)	(3.03)	(3.86)	(1.67)	(3.51)	(2.27)	(3.05)	(2.70)
N	5070	5016	4863	4811	4966	4914	3621	3583	3621	3583
Adj. R2	0.46	0.43	0.39	0.33	0.28	0.27	0.20	0.17	0.13	0.11
Private enterprises										
BDG_{q+1}	0.821***	0.628***	0.325***	0.258***	0.092***	0.063***	0.350***	0.301***	0.196***	0.179***
	(6.61)	(3.26)	(6.23)	(4.04)	(5.13)	(3.31)	(6.10)	(3.03)	(5.35)	(3.15)
N	6197	6131	5944	5881	6070	6006	4426	4379	4426	4379
Adj. R2	0.70	0.64	0.59	0.49	0.41	0.41	0.30	0.25	0.20	0.16
High HHI										
BDG_{q+1}	0.597***	0.346***	0.282***	0.171**	0.044***	0.038**	0.205***	0.163**	0.140***	0.101**
	(4.36)	(2.92)	(3.56)	(2.42)	(3.54)	(2.38)	(3.18)	(2.48)	(3.89)	(2.49)
N	5633	5574	5403	5346	5518	5460	4024	3981	4024	3981
Adj. R2	0.51	0.47	0.43	0.36	0.30	0.30	0.22	0.18	0.15	0.12
Low HHI										
BDG_{q+1}	0.935***	0.541***	0.441***	0.268***	0.069***	0.060***	0.320***	0.254***	0.218***	0.157***
	(6.68)	(4.48)	(5.44)	(3.70)	(5.42)	(3.64)	(4.87)	(3.80)	(5.96)	(3.81)
N	5633	5574	5403	5346	5518	5460	4024	3981	4024	3981
Adj. R2	0.65	0.60	0.55	0.46	0.39	0.38	0.28	0.23	0.19	0.15
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

Panel B: nowcasting and forecasting earnings surprises across different firms				
	SUE_{q+1}	SUE_{q+2}	CAR_{q+1}	CAR_{q+2}
Large				
BDG_{q+1}	0.151***	0.144*	1.763**	0.942
	(2.98)	(1.87)	(2.40)	(1.57)
N	5461	5403	5576	5517
Adj. R2	0.39	0.33	0.07	0.06
Small				
BDG_{q+1}	0.233***	0.188***	2.911***	1.825***
	(4.14)	(3.65)	(3.12)	(2.69)
N	5461	5403	5576	5517
Adj. R2	0.43	0.37	0.08	0.07
High IO				
BDG_{q+1}	0.186***	0.162**	1.72***	1.284**
	(2.71)	(2.34)	(2.77)	(2.16)
N	5461	5403	5576	5517
Adj. R2	0.37	0.32	0.07	0.06
Low IO				
BDG_{q+1}	0.186***	0.196**	2.603***	1.249**
	(3.12)	(2.46)	(3.27)	(2.06)
N	5461	5403	5576	5517
Adj. R2	0.45	0.39	0.08	0.07
High coverage				
BDG_{q+1}	0.226***	0.143**	1.657***	1.243
	(3.14)	(2.21)	(2.83)	(1.61)
N	5461	5403	5576	5517
Adj. R2	0.35	0.30	0.06	0.06
Low coverage				
BDG_{q+1}	0.268***	0.201***	1.944***	1.425***
	(3.36)	(2.77)	(4.26)	(2.81)
N	5461	5403	5576	5517
Adj. R2	0.47	0.40	0.08	0.08

State enterprises				
BDG_{q+1}	0.138***	0.145*	1.931***	1.132**
	(2.85)	(1.84)	(3.04)	(2.01)
N	4915	4863	5018	4965
Adj. R2	0.33	0.28	0.06	0.05
Private enterprises				
BDG_{q+1}	0.238***	0.173***	2.709***	1.727**
	(3.48)	(2.82)	(3.94)	(2.22)
N	6007	5943	6133	6068
Adj. R2	0.49	0.42	0.09	0.08
High HHI				
BDG_{q+1}	0.196***	0.145***	1.435**	1.121*
	(3.48)	(2.62)	(2.25)	(1.69)
N	5461	5403	5576	5517
Adj. R2	0.36	0.31	0.06	0.06
Low HHI				
BDG_{q+1}	0.307***	0.228***	2.244***	1.754***
	(5.32)	(4.01)	(3.45)	(2.59)
N	5461	5403	5576	5517
Adj. R2	0.46	0.39	0.08	0.07
Controls	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Year-Quarter FE	Y	Y	Y	Y

Panel C: long-short excess returns and alphas across different firms					
Value-weighted	Excess	HXZ	FF5	LSY3	LSY4
Large					
L/S	0.430**	0.453*	0.407**	0.338	0.327*
	(2.32)	(1.93)	(2.40)	(1.48)	(1.71)
Small					
L/S	1.121***	1.011***	0.847***	0.746***	0.763***
	(4.72)	(4.15)	(3.87)	(4.51)	(3.11)
High IO					
L/S	0.533**	0.629***	0.575***	0.505***	0.546**
	(2.54)	(2.66)	(3.35)	(3.02)	(2.06)
Low IO					
L/S	0.868***	0.768***	0.753***	0.668***	0.745***
	(5.22)	(4.53)	(4.21)	(4.09)	(2.83)
High coverage					
L/S	0.621***	0.363**	0.413*	0.437*	0.319*
	(3.04)	(2.25)	(1.96)	(1.77)	(1.66)
Low coverage					
L/S	1.162***	0.724***	0.847***	0.660***	0.782***
	(4.43)	(4.63)	(4.24)	(4.17)	(3.94)
State enterprises					
L/S	0.663***	0.634***	0.470***	0.562***	0.423**
	(2.87)	(2.70)	(2.67)	(2.88)	(2.26)
Private enterprises					
L/S	1.177***	0.854***	0.715***	0.804***	0.718***
	(4.53)	(4.46)	(5.21)	(3.49)	(2.79)
High HHI					
L/S	0.686***	0.536***	0.592***	0.518***	0.514***
	(2.86)	(2.70)	(3.48)	(3.23)	(2.77)
Low HHI					
L/S	1.074***	0.839***	0.926***	0.810***	0.804***
	(4.37)	(4.13)	(5.33)	(4.95)	(4.25)

Table 10: Interaction Effects of BDG and Firm Types

Panel A and B report the results on the regressions of firm fundamentals, and earnings surprise measured in quarter $q+1$ or quarter $q+2$ on the BDG in quarter $q+1$, the interaction terms in quarter q, dummies in quarter q, and other control variables in quarter q. The dependent variables in Panel A include return on assets (ROA), growth of total assets (AG), growth of sales (SG), the log of one plus quarterly number of patents applied (PA), and the log of one plus quarterly number of patents granted (PG). The dependent variables in Panel B include standardized unexpected earnings (SUE) and earnings announcement abnormal returns (CAR). The dummy DSmall equals one when the firm's size is below the median of the sample size. The dummy DLowIO equals one when the firm's institutional ownership is below the median of the sample institutional ownership. The dummy DLowCov equals one when the firm's analyst coverage is below the median of the sample analyst coverage. The dummy DPrivate equals one when the firm is private-owned. The dummy DLowHHI equals one when the firm's HHI is below the median of the sample HHI. We add interaction terms between BDG and these dummies. These dummies are also included into regressions as controls. We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and one standard deviation. The control variables are from Panel A of Table 1. The t-statistics of robust standard errors clustered at the industry and year-quarter levels are reported in parentheses. Panel C reports the Fama and MacBeth (1973) cross-sectional regressions of monthly returns in quarter $q+1$ on BDG in quarter q and the interaction terms in quarter q, dummies in quarter q, and other control variables in quarter q. The BDG and other accounting variables in quarter q are matched to stock returns in month t in quarter $q+1$. The monthly price-based variables are calculated using the last non-missing observations prior to each month. The dependent variables include the firm's future excess returns and abnormal returns adjusted by different factor models. The dummy DSmall equals one when the firm's size is below the median of the sample size. The dummy DLowIO equals one when the firm's institutional ownership is below the median of the sample institutional ownership. The dummy DLowCov equals one when the firm's analyst coverage is below the median of the sample analyst coverage. The dummy DPrivate equals one when the firm is private-owned. The dummy DLowHHI equals one when the firm's HHI is below the median of the sample HHI. We add interaction terms between BDG and these dummies. These dummies are also included into regressions as controls. We control for the industry and region fixed effects following the CSRC industry classification and China province classification. All returns are expressed in percentage. The BDG and other firm-specific characteristics are defined in Panel A of Table 1. All explanatory variables are generated using the last non-missing available observation for each quarter q. Cross-sectional regressions are run every calendar month, and the time-series standard errors are corrected for heteroskedasticity and autocorrelation. Newey and West (1987) adjusted t-statistics are reported in parentheses. Coefficients marked with *, **, and *** are significant at the 10%, the 5%, and the 1% level, respectively. The sample period is from second quarter of 2016 to third quarter of 2021.

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

Panel B: nowcasting and forecasting earnings surprises with interaction terms					
	SUE_{q+1}	SUE_{q+2}	CAR_{q+1}	CAR_{q+2}	
BDG_{q+1}	0.049***	0.043*	0.398*	0.288	
	(2.76)	(1.96)	(1.66)	(1.21)	
BDG_{q+1} * DSmall	0.053***	0.045**	0.665*	0.504	
	(2.62)	(2.11)	(1.80)	(1.63)	
BDG_{q+1} * DLowIO	0.035**	0.052	0.506*	0.236	
	(1.99)	(1.51)	(1.90)	(1.16)	
BDG_{q+1} * DLowCov	0.072**	0.048*	0.474**	0.396	
	(2.21)	(1.72)	(2.54)	(1.58)	
BDG_{q+1} * DPrivate	0.056**	0.043*	0.594**	0.473	
	(2.09)	(1.78)	(2.35)	(1.40)	
BDG_{q+1} * DLowHHI	0.084**	0.056**	0.530***	0.461*	
	(2.55)	(1.98)	(2.89)	(1.84)	
Dummies	Y	Y	Y	Y	
Controls	Y	Y	Y	Y	
Industry FE	Y	Y	Y	Y	
Year-Quarter FE	Y	Y	Y	Y	
N	10922	10806	11152	11033	
Adj. R2	0.50	0.42	0.09	0.08	

Panel C: nowcasting and forecasting returns and alphas with interaction terms					
Value-weighted	Excess	HXZ	FF5	LSY3	LSY4
BDG	0.080***	0.086***	0.093**	0.086*	0.091**
	(2.84)	(2.86)	(2.37)	(1.78)	(1.99)
BDG * DSmall	0.113***	0.087**	0.095**	0.071**	0.070*
	(2.96)	(2.47)	(2.15)	(2.47)	(1.76)
BDG * DLowIO	0.125***	0.156**	0.134**	0.138**	0.126
	(3.01)	(2.56)	(2.35)	(2.50)	(1.61)
BDG * DLowCov	0.130**	0.081***	0.101**	0.085***	0.080**
	(2.49)	(2.73)	(2.49)	(2.61)	(2.56)
BDG * DPrivate	0.145***	0.127**	0.127***	0.130**	0.085
	(2.59)	(2.54)	(3.24)	(2.17)	(1.60)
BDG * DLowHHI	0.170***	0.144***	0.153***	0.151**	0.099*
	(3.01)	(2.98)	(3.79)	(2.56)	(1.89)
Dummies	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y
N	33454	33454	33454	33454	33454
Adj. R2	0.10	0.08	0.08	0.08	0.08

Table 11: BDG predictability using blockchain service industry

This table reports the regression results of the fundamentals, earnings, and returns on the blockchain data growth, after controlling for endogeneity by including the residuals estimated from the first-stage regression (1st stage residual). In the first-stage regression, Panel A estimates the Probit regression of the likelihood of blockchain data growth (BDG) on the instrumental variable Blockchain Service Industry (BSI) and control variables. BSI is the natural logarithm of the number of companies in the focal firm's industry that are included in the "List of Companies with Blockchain Digital Services" maintained by China's Cyber Security and Digitization Committee. Panel B-D shows the second-stage regression of fundamentals, earnings, and returns. We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and one standard deviation. The control variables are from Panel A of Table 1. The z-statistics and t-statistics of robust standard errors clustered at the industry and region * quarter levels are reported in parentheses. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively. The sample period is from April 2016 to September 2021.

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}	3.313***	1.863***	1.431***	0.910**	0.259***	0.230**	0.944***	0.811**	0.816***	0.505**
	(3.92)	(2.81)	(3.61)	(2.10)	(3.44)	(2.31)	(2.84)	(2.46)	(3.80)	(2.30)
1st stage residual	1.391***	1.005**	0.869***	0.454**	0.145***	0.125**	0.636***	0.515*	0.439***	0.339**
	(3.54)	(2.55)	(3.12)	(1.99)	(2.62)	(1.98)	(2.80)	(1.90)	(3.42)	(2.04)
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Region * 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	0.73	0.69	0.65	0.53	0.45	0.43	0.33	0.26	0.20	0.16

Panel C: 2nd stage regression of earnings

	SUE_{q+1}	SUE_{q+2}	CAR_{q+1}	CAR_{q+2}
BDG_{q+1}	1.102***	0.749**	8.070**	5.897
	(3.13)	(2.24)	(2.14)	(1.56)
1st stage residual	0.567***	0.483**	4.497*	3.691
	(2.74)	(2.22)	(1.69)	(1.38)
Controls	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Region * Quarter FE	Y	Y	Y	Y
N	9283	9185	9479	9378
Adj. R2	0.50	0.43	0.09	0.09

Panel D: 2nd stage regression of returns

	$Excess_{q+1}$	$Excess_{q+2}$	$LSY4_{q+1}$	$LSY4_{q+2}$
BDG_{q+1}	1.719***	1.160***	1.331**	1.174**
	(3.34)	(2.72)	(2.57)	(2.04)
1st stage residual	0.992***	0.827**	0.865**	0.646*
	(2.83)	(2.47)	(2.21)	(1.72)
Controls	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Region * Quarter	Y	Y	Y	Y
N	28436	27014	28436	27014
Adj. R2	0.11	0.10	0.09	0.08

Table 12: BDG predictability using blockchain service region

This table reports the regression results of the fundamentals, earnings, and returns on the blockchain data growth, after controlling for endogeneity by including the residuals estimated from the first-stage regression (1st stage residual). In the first-stage regression, Panel A estimates a Probit regression of the likelihood of blockchain data growth (BDG) on the instrumental variable Blockchain Service Region (BSR) and control variables. BSR is the natural logarithm of the number of companies in the focal firm's city (based on headquarter locations) that are included in the "List of Companies with Blockchain Digital Service" maintained by China's Cyber Security and Digitization Committee. Panel B-D shows the second-stage regression of fundamentals, earnings, and returns. We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and one standard deviation. The control variables are from Panel A of Table 1. The z-statistics and t-statistics of robust standard errors clustered at the region and industry * quarter levels are reported in parentheses. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively. The sample period is from April 2016 to September 2021.

Panel A: 1st stage regression

	BDG_{q+1}	BDG_{q+2}
BSR	0.345***	0.319**
	(2.86)	(2.27)
Controls	Y	Y
N	10140	10032
Adj. R2	0.34	0.29

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}	2.710***	1.653**	1.183***	0.777*	0.212***	0.185**	0.768**	0.705**	0.700***	0.425*
	(3.46)	(2.41)	(3.17)	(1.70)	(3.02)	(2.07)	(2.40)	(2.04)	(3.06)	(1.94)
1st stage residual	1.240***	0.834**	0.774***	0.387*	0.117**	0.110*	0.565**	0.421*	0.355***	0.301*
	(3.10)	(2.16)	(2.70)	(1.75)	(2.12)	(1.70)	(2.45)	(1.67)	(2.89)	(1.79)
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Industry * 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	0.62	0.57	0.56	0.46	0.40	0.38	0.28	0.22	0.16	0.14

Panel C: 2nd stage regression of earnings				
	SUE_{q+1}	SUE_{q+2}	CAR_{q+1}	CAR_{q+2}
BDG_{q+1}	0.983**	0.631**	7.242*	4.759
	(2.57)	(2.00)	(1.76)	(1.35)
1st stage residual	0.461**	0.392*	3.656	3.035
	(2.44)	(1.96)	(1.38)	(1.21)
Controls	Y	Y	Y	Y
Region FE	Y	Y	Y	Y
Industry * Quarter FE	Y	Y	Y	Y
N	9283	9185	9479	9378
Adj. R2	0.45	0.35	0.08	0.07

Panel D: 2nd stage regression of returns				
	$Excess_{q+1}$	$Excess_{q+2}$	$LSY4_{q+1}$	$LSY4_{q+2}$
BDG_{q+1}	1.417***	1.013**	1.066**	1.051*
	(2.72)	(2.21)	(2.11)	(1.71)
1st stage residual	0.824**	0.707**	0.747*	0.532
	(2.36)	(1.99)	(1.78)	(1.37)
Controls	Y	Y	Y	Y
Region FE	Y	Y	Y	Y
Industry * Quarter FE	Y	Y	Y	Y
N	28436	27014	28436	27014
Adj. R2	0.09	0.08	0.07	0.07

Table 13: Difference-in-Difference Tests

This table reports the difference-in-difference tests of fundamentals, earnings, and returns. The sample window is 8 quarters. The first four quarters are when firms do not use blockchain services. The second four quarters are when firms use blockchain services. Panel A-C reports the fundamentals, earnings, and returns, respectively. The dummy variable Treat equals one when the firm uses blockchain service, otherwise zero. The control firms do not use blockchain services from our sample. For each treatment firm, we match control firms in the same industry and use the propensity score matching method based on each characteristic of fundamentals, earnings, or returns. The dummy variable Post equals one when the firm begin to use blockchain services, otherwise zero. The control variables are from Panel A of Table 1. The t-statistics of robust standard errors clustered at the industry * quarter and region * quarter levels are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: difference-in-difference tests of fundamentals

	ROA	AG	SG	PA	PG
Treat * Post	0.131***	0.096***	0.053**	1.123***	1.133**
	(3.52)	(2.80)	(2.15)	(3.09)	(2.27)
Treat	0.092	0.065	0.035	0.722	0.735
	(1.41)	(0.96)	(0.78)	(0.86)	(0.69)
Post	0.055	0.045	0.022	0.502	0.422
	(0.76)	(0.53)	(0.33)	(0.42)	(0.57)
Controls	Y	Y	Y	Y	Y
Industry * Quarter FE	Y	Y	Y	Y	Y
Region * Quarter FE	Y	Y	Y	Y	Y
N	4056	3890	3973	2897	2897
Adj. R2	0.37	0.31	0.21	0.15	0.10

Panel B: difference-in-difference tests of earnings

	SUE	CAR
Treat * Post	0.136***	0.130***
	(5.88)	(4.96)
Treat	0.096**	0.074
	(2.01)	(1.33)
Post	0.061	0.063
	(1.07)	(1.00)
Controls	Y	Y
Industry * Quarter FE	Y	Y
Region * Quarter FE	Y	Y
N	3932	4015
Adj. R2	0.25	0.04

Panel C: difference-in-difference tests of returns		
	Excess	LSY4
Treat * Post	0.105***	0.077***
	(3.76)	(3.17)
Treat	0.063	0.055
	(1.52)	(1.53)
Post	0.039	0.036
	(1.00)	(0.97)
Controls	Y	Y
Industry FE	Y	Y
Year-Quarter FE	Y	Y
N	12044	12044
Adj. R2	0.05	0.04

Table 14: Robustness in Subsamples

This table presents results from the value-weighted portfolios in different stock subsamples. First, the stock subsamples are partitioned by manufacturing industry and Non-manufacturing industries. Second, the stock subsamples are partitioned by the Top5 provinces and Non-Top5 provinces. Third, the stock subsamples are partitioned into before COVID-19 period and after COVID-19 period. Panel A reports the results on the regressions of firm fundamentals measured in quarter q+1 or quarter $q+2$ on the BDG in quarter q+1 and other control variables in quarter q in different stock subsamples. The dependent variables include return on assets (ROA), growth of total assets (AG), growth of sales (SG), the log of one plus quarterly number of patents applied (PA), and the log of one plus quarterly number of patents granted (PG). Panel B reports the results on the regressions of earnings surprise measured in quarter $q+1$ or quarter $q+2$ on the BDG in quarter $q+1$ and other control variables in quarter q in different stock subsamples. The dependent variables include standardized unexpected earnings (SUE) and earnings announcement abnormal returns (CAR). We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and one standard deviation. The control variables are from Panel A of Table 1. The t-statistics of robust standard errors clustered at the industry and year-quarter levels are reported in parentheses. Panel C reports the average monthly long-short excess returns and long-short alphas on the value-weighted portfolios of stocks sorted by the BDG in different stock subsamples. Newey and West (1987) adjusted t-statistics are given in parentheses. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively. The sample period is from April 2016 to September 2021.

Panel A: nowcasting and forecasting firm fundamentals in subsamples										
	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}
Manufacturing industry										
BDG_{q+1}	0.968***	0.625***	0.270***	0.225***	0.071***	0.044**	0.366***	0.237***	0.174***	0.159***
	(4.84)	(2.84)	(4.61)	(3.97)	(3.68)	(2.40)	(4.07)	(3.73)	(4.34)	(2.86)
N	6309	6242	6052	5988	6180	6115	4507	4459	4507	4459
Adj. R2	0.50	0.46	0.42	0.35	0.30	0.30	0.21	0.18	0.14	0.11
Non-Manufacturing industries										
BDG_{q+1}	0.709***	0.431**	0.238***	0.198**	0.057**	0.039*	0.242***	0.170***	0.157***	0.110***
	(4.56)	(2.20)	(3.36)	(2.39)	(2.57)	(1.92)	(4.09)	(2.87)	(3.66)	(2.82)
N	4957	4905	4755	4705	4856	4805	3541	3503	3541	3503
Adj. R2	0.43	0.40	0.36	0.30	0.26	0.25	0.18	0.15	0.12	0.10
Top5 provinces										
BDG_{q+1}	0.885***	0.586***	0.367***	0.245***	0.095***	0.050***	0.418***	0.209***	0.166***	0.141***
	(5.58)	(3.27)	(5.45)	(3.88)	(4.38)	(2.87)	(4.66)	(2.76)	(4.55)	(3.07)
N	6760	6688	6484	6415	6622	6552	4828	4777	4828	4777
Adj. R2	0.52	0.48	0.44	0.37	0.31	0.31	0.22	0.19	0.15	0.12
Non-Top5 provinces										
BDG_{q+1}	0.809***	0.572**	0.299***	0.247***	0.073***	0.049**	0.268***	0.209***	0.156***	0.110***
	(4.02)	(2.55)	(3.64)	(3.25)	(4.31)	(2.50)	(4.35)	(3.13)	(3.16)	(2.94)
N	4507	4459	4323	4277	4415	4368	3219	3185	3219	3185
Adj. R2	0.41	0.37	0.34	0.28	0.24	0.24	0.17	0.15	0.12	0.09
Before COVID-19										
BDG_{q+1}	0.647***	0.576***	0.239***	0.229***	0.057***	0.041***	0.264***	0.197***	0.160***	0.103***
	(6.02)	(2.73)	(5.37)	(3.30)	(3.34)	(2.63)	(4.26)	(3.54)	(3.85)	(2.74)
N	7661	7580	7349	7271	7505	7425	5472	5414	5472	5414
Adj. R2	0.57	0.52	0.48	0.40	0.34	0.34	0.24	0.20	0.16	0.13
After COVID-19										
BDG_{q+1}	0.805***	0.544***	0.320***	0.266***	0.083***	0.049**	0.370***	0.186***	0.185***	0.110***
	(4.81)	(2.92)	(4.10)	(3.36)	(3.47)	(2.41)	(4.15)	(2.81)	(4.05)	(2.76)
N	3605	3567	3458	3421	3532	3494	2575	2548	2575	2548
Adj. R2	0.36	0.33	0.31	0.25	0.21	0.21	0.15	0.13	0.10	0.08
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

Panel B: nowcasting and forecasting earnings surprises in subsamples				
	SUE_{q+1}	SUE_{q+2}	CAR_{q+1}	CAR_{q+2}
Manufacturing industry				
BDG_{q+1}	0.231***	0.203***	2.204***	1.568***
	(3.36)	(2.76)	(3.25)	(2.59)
N	6116	6051	6245	6179
Adj. R2	0.35	0.30	0.06	0.06
Non-Manufacturing industries				
BDG_{q+1}	0.166**	0.172**	1.574***	1.364*
	(2.46)	(2.17)	(2.72)	(1.92)
N	4806	4755	4907	4855
Adj. R2	0.30	0.26	0.05	0.05
Top5 provinces				
BDG_{q+1}	0.284***	0.207**	2.626***	1.642**
	(3.95)	(2.24)	(3.56)	(2.26)
N	6553	6483	6691	6620
Adj. R2	0.37	0.32	0.07	0.06
Non-Top5 provinces				
BDG_{q+1}	0.193***	0.167**	1.985**	1.230**
	(2.63)	(2.28)	(2.54)	(2.08)
N	4369	4322	4461	4413
Adj. R2	0.29	0.25	0.05	0.05
Before COVID-19				
BDG_{q+1}	0.247***	0.177***	1.910***	1.369**
	(3.66)	(2.98)	(3.55)	(2.16)
N	7427	7348	7583	7503
Adj. R2	0.40	0.34	0.07	0.06
After COVID-19				
BDG_{q+1}	0.196***	0.149***	2.442***	1.574*
	(3.32)	(2.76)	(3.57)	(1.83)
N	3495	3458	3568	3531
Adj. R2	0.26	0.22	0.05	0.04
Controls	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Year-Quarter FE	Y	Y	Y	Y

Panel C: long-short excess returns and alphas in subsamples					
Value-weighted	Excess	HXZ	FF5	LSY3	LSY4
Manufacturing industry					
L/S	1.109***	0.655***	0.605***	0.684***	0.635***
	(4.80)	(4.10)	(4.79)	(3.05)	(3.47)
Non-Manufacturing industries					
L/S	0.644***	0.713***	0.588***	0.491***	0.523***
	(3.74)	(3.54)	(3.40)	(2.64)	(2.65)
Top5 provinces					
L/S	1.196***	0.657***	0.924***	0.843***	0.677***
	(4.95)	(4.66)	(4.56)	(3.79)	(2.88)
Non-Top5 provinces					
L/S	0.675***	0.723***	0.669***	0.610***	0.638**
	(3.23)	(3.74)	(3.40)	(3.56)	(2.45)
Before COVID-19					
L/S	0.711***	0.720***	0.653***	0.678***	0.662**
	(3.81)	(3.17)	(3.16)	(3.27)	(2.49)
After COVID-19					
L/S	0.895***	0.744***	0.700***	0.604***	0.773**
	(3.76)	(3.29)	(4.08)	(3.62)	(2.53)

Table 15: International Evidence

This table reports international evidence. Panel A reports the results on the regressions of firm fundamentals measured in quarter $q+1$ or quarter $q+2$ on the BDG in quarter $q+1$ and other control variables in quarter q in Indonesia, Malaysia, South Korea, and Thailand. The dependent variables include return on assets (ROA), growth of total assets (AG), and growth of sales (SG). Panel B reports the results on the regressions of earnings surprise measured in quarter $q+1$ or quarter $q+2$ on the BDG in quarter $q+1$ and other control variables in quarter q in Indonesia, Malaysia, South Korea, and Thailand. The dependent variables include standardized unexpected earnings (SUE) and earnings announcement abnormal returns (CAR). We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and one standard deviation. The control variables are from Panel A of Table 1. The t-statistics of robust standard errors clustered at the industry and year-quarter levels are reported in parentheses. Panel C reports the average monthly long-short excess returns and long-short alphas on the value-weighted portfolios of stocks sorted by the BDG in Indonesia, Malaysia, South Korea, and Thailand. Coefficients marked with *, **, and *** are significant at the 10%, the 5%, and the 1% level, respectively. The sample period is from second quarter of 2016 to third quarter of 2021.

Panel A: nowcasting and forecasting firm fundamentals in the four countries

	ROA_{q+1}	ROA_{q+2}	AG_{q+1}	AG_{q+2}	SG_{q+1}	SG_{q+2}
Indonesia						
BDG_{q+1}	0.429***	0.340**	0.152**	0.144*	0.032**	0.029*
	(2.69)	(2.38)	(2.13)	(1.73)	(2.54)	(1.69)
N	3943	3901	3782	3742	3863	3822
Adj. R2	0.26	0.24	0.22	0.18	0.16	0.15
Malaysia						
BDG_{q+1}	0.694***	0.585***	0.372***	0.291***	0.065***	0.065**
	(3.76)	(3.30)	(3.17)	(2.87)	(3.87)	(2.54)
N	1690	1672	1621	1604	1655	1638
Adj. R2	0.15	0.13	0.12	0.10	0.09	0.09
South Korea						
BDG_{q+1}	0.209**	0.130	0.088*	0.074*	0.020**	0.013
	(2.12)	(1.21)	(1.90)	(1.81)	(2.11)	(1.43)
N	5070	5016	4863	4811	4966	4914
Adj. R2	0.32	0.29	0.27	0.22	0.19	0.19
Thailand						
BDG_{q+1}	0.569**	0.398**	0.276***	0.195*	0.055***	0.045**
	(2.43)	(2.18)	(3.07)	(1.77)	(2.95)	(2.35)
N	2817	2787	2702	2673	2759	2730
Adj. R2	0.20	0.19	0.17	0.14	0.12	0.12
Controls	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y
Year-Quarter FE	Y	Y	Y	Y	Y	Y

Panel B: nowcasting and forecasting earnings surprises in the four countries				
	SUE_{q+1}	SUE_{q+2}	CAR_{q+1}	CAR_{q+2}
Indonesia				
BDG_{q+1}	0.140**	0.095*	1.538***	1.008**
	(2.05)	(1.69)	(2.71)	(1.98)
N	3823	3782	3903	3862
Adj. R2	0.19	0.16	0.03	0.03
Malaysia				
BDG_{q+1}	0.357***	0.258***	2.418***	1.973**
	(4.82)	(3.50)	(3.85)	(2.08)
N	1638	1621	1673	1655
Adj. R2	0.10	0.09	0.02	0.02
South Korea				
BDG_{q+1}	0.089**	0.068	1.068**	0.585
	(2.21)	(1.55)	(2.05)	(1.42)
N	4915	4863	5018	4965
Adj. R2	0.23	0.19	0.04	0.04
Thailand				
BDG_{q+1}	0.220***	0.161***	2.121***	1.060
	(3.10)	(2.62)	(2.73)	(1.62)
N	2730	2701	2788	2758
Adj. R2	0.14	0.12	0.03	0.02
Controls	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Year-Quarter FE	Y	Y	Y	Y

Panel C: long-short excess returns and alphas in the four countries					
Value-weighted	Excess	CAPM	FF3	FF5	FF6
Indonesia					
L/S	0.322***	0.364**	0.388***	0.349*	0.321**
	(3.33)	(2.22)	(2.99)	(1.83)	(2.19)
Malaysia					
L/S	0.571***	0.526***	0.478***	0.376***	0.434**
	(3.95)	(3.65)	(3.69)	(2.94)	(2.39)
South Korea					
L/S	0.306**	0.330**	0.243**	0.285	0.215
	(2.52)	(2.22)	(2.12)	(1.52)	(1.39)
Thailand					
L/S	0.444***	0.324***	0.427***	0.324**	0.339**
	(3.46)	(3.35)	(3.15)	(2.44)	(2.02)

Table 16: Digital Economy Variables

Panel A reports the results on the regressions of firm fundamentals measured in quarter $q+1$ or quarter $q+2$ on the BDG in quarter $q+1$ and other control variables in quarter q across different firms. The dependent variables include return on assets (ROA), growth of total assets (AG), growth of sales (SG), the log of one plus quarterly number of patents applied (PA), and the log of one plus quarterly number of patents granted (PG). Panel B reports the results on the regressions of earnings surprise measured in quarter $q+1$ or quarter $q+2$ on the BDG in quarter $q+1$ and other control variables in quarter q across different firms. The dependent variables include standardized unexpected earnings (SUE) and earnings announcement abnormal returns (CAR). The digital economy variables include the year-over-year quarterly growth of number of IoT chips of firms ($IOTG_q$), the year-over-year quarterly growth of number of industrial robots of firms ($ROBG_q$), the year-over-year quarterly growth of number of STEM (Science, Technology, Engineering, and Mathematics) employees of firms ($STEMG_q$), and the year-over-year quarterly growth of cloud data of firms (CDG_q). We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and one standard deviation. The control variables are from Panel A of Table 1. The t-statistics of robust standard errors clustered at the industry and year-quarter levels are reported in parentheses. Panel C reports the Fama and MacBeth (1973) cross-sectional regressions of monthly returns in quarter $q+1$ on BDG in quarter q and digital economy variables in quarter q, and other control variables in quarter q. The BDG and other accounting variables in quarter q are matched to stock returns in month t in quarter $q+1$. The monthly price-based variables are calculated using the last non-missing observations prior to each month. The dependent variables include the firm's future excess returns and abnormal returns adjusted by different factor models. We control for the industry and region fixed effects following the CSRC industry classification and China province classification. All returns are expressed in percentage. The BDG and other firm-specific characteristics are defined in Panel A of Table 1. All explanatory variables are generated using the last non-missing available observation for each quarter q. Cross-sectional regressions are run every calendar month, and the time-series standard errors are corrected for heteroskedasticity and autocorrelation. Newey and West (1987) adjusted t-statistics are reported in parentheses. Coefficients marked with *, **, and *** are significant at the 10%, the 5%, and the 1% level, respectively. The sample period is from second quarter of 2016 to third quarter of 2021.

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

Panel B: Nowcasting and forecasting earnings surprises after controlling digital economy variables				
	SUE_{q+1}	SUE_{q+2}	CAR_{q+1}	CAR_{q+2}
BDG_{q+1}	0.096**	0.074*	1.259*	1.097
	(2.13)	(1.68)	(1.80)	(1.64)
$IoTG_{q+1}$	0.303**	0.223*	0.225	0.176
	(2.18)	(1.71)	(1.48)	(1.20)
$ROBG_{q+1}$	0.221**	0.182*	0.266*	0.231
	(2.37)	(1.91)	(1.82)	(1.34)
$STEMG_{q+1}$	0.222**	0.174	0.144**	0.121*
	(2.13)	(1.51)	(2.20)	(1.81)
CDG_{q+1}	0.129**	0.101	1.547**	1.222
	(2.17)	(1.62)	(2.04)	(1.52)
Controls	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Year-Quarter FE	Y	Y	Y	Y
N	4041	3998	4126	4082
Adj. R2	0.48	0.41	0.09	0.08

Panel C: nowcasting and forecasting returns and alphas after controlling digital economy variables					
	Excess	HXZ	FF5	LSY3	LSY4
BDG	0.343***	0.356***	0.185***	0.264***	0.339**
	(2.77)	(3.29)	(2.92)	(3.63)	(2.51)
IoTG	0.175	0.171	0.373	0.513*	2.524
	(1.10)	(1.09)	(1.57)	(1.79)	(1.62)
ROBG	0.036	0.036	0.071*	0.098**	1.705
	(0.83)	(0.78)	(1.73)	(2.14)	(1.45)
STEMG	0.127	0.133	0.147	0.197	1.928
	(1.48)	(1.43)	(0.94)	(1.10)	(1.08)
CDG	0.409***	0.418***	0.259***	0.339***	0.357***
	(3.31)	(3.49)	(3.32)	(4.02)	(2.81)
Controls	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y
N	12378	12378	12378	12378	12378
Adj. R2	0.12	0.12	0.08	0.09	0.10

Table 17: On-chain Data Growth by Category

We test the nowcasting and forecasting effect of BDG in different blockchain categories, including operation, financials, human resources, marketing, IT, supply chain, and others. Panel A and B report the results on the regressions of firm fundamentals, and earnings surprise measured in quarter $q+1$ or quarter $q+2$ on the seven categories of BDG in quarter $q+1$, and other control variables in quarter q. The dependent variables in Panel A include return on assets (ROA), growth of total assets (AG), growth of sales (SG), the log of one plus quarterly number of patents applied (PA), and the log of one plus quarterly number of patents granted (PG). The dependent variables in Panel B include standardized unexpected earnings (SUE) and earnings announcement abnormal returns (CAR). We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and one standard deviation. The control variables are from Panel A of Table 1. The t-statistics of robust standard errors clustered at the industry and year-quarter levels are reported in parentheses. Panel C reports the Fama and MacBeth (1973) cross-sectional regressions of monthly returns in quarter $q+1$ on seven categories of BDG in quarter q and other control variables in quarter q. The seven categories of BDG and other accounting variables in quarter q are matched to stock returns in month t in quarter $q+1$. The monthly price-based variables are calculated using the last non-missing observations prior to each month. The dependent variables include the firm's future excess returns and abnormal returns adjusted by different factor models. We control for the industry and region fixed effects following the CSRC industry classification and China province classification. All returns are expressed in percentage. The seven categories of BDG and other firm-specific characteristics are defined in Panel A of Table 1. All explanatory variables are generated using the last non-missing available observation for each quarter q. Cross-sectional regressions are run every calendar month, and the time-series standard errors are corrected for heteroskedasticity and autocorrelation. Newey and West (1987) adjusted t-statistics are reported in parentheses. Coefficients marked with *, **, and *** are significant at the 10%, the 5%, and the 1% level, respectively. The sample period is from second quarter of 2016 to third quarter of 2021.

Panel A: nowcasting and forecasting firm fundamentals with different blockchain categories										
	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}
$BDG_{q+1}^{Operation}$	0.579***	0.301**	0.267***	0.165*	0.038***	0.035**	0.167***	0.137**	0.120***	0.094**
	(3.46)	(2.39)	(2.98)	(1.91)	(3.11)	(2.02)	(2.61)	(2.10)	(3.22)	(2.09)
$BDG_{q+1}^{Financials}$	0.536***	0.302**	0.253***	0.132*	0.039***	0.029*	0.156***	0.138*	0.119***	0.076*
	(3.40)	(2.35)	(2.68)	(1.92)	(2.60)	(1.73)	(2.60)	(1.85)	(3.29)	(1.80)
$BDG_{q+1}^{Humanresources}$	0.222	0.115	0.098	0.062	0.014	0.014	0.076	0.062	0.052	0.033
	(1.48)	(0.93)	(1.27)	(0.81)	(1.16)	(0.83)	(0.99)	(0.81)	(1.23)	(0.75)
$BDG_{q+1}^{Marketing}$	0.207	0.114	0.107	0.063	0.016	0.014	0.075	0.059	0.047	0.036
	(1.52)	(0.96)	(1.26)	(0.86)	(1.28)	(0.76)	(1.15)	(0.79)	(1.41)	(0.86)
BDG_{q+1}^{IT}	0.180	0.105	0.096	0.052	0.015	0.011	0.067	0.054	0.048	0.030
	(1.27)	(0.90)	(1.04)	(0.78)	(1.13)	(0.77)	(0.90)	(0.72)	(1.20)	(0.75)
$BDG_{q+1}^{Supplychain}$	0.242*	0.145	0.110	0.078	0.019	0.016	0.081	0.074	0.055	0.046
	(1.71)	(1.18)	(1.35)	(0.91)	(1.48)	(0.92)	(1.37)	(0.91)	(1.48)	(0.96)
BDG_{q+1}^{Others}	0.222	0.113	0.097	0.061	0.015	0.013	0.069	0.055	0.049	0.035
	(1.41)	(0.95)	(1.27)	(0.8)	(1.20)	(0.84)	(1.01)	(0.88)	(1.35)	(0.80)
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	5968	5139	5522	5608	5300	5571	3932	3849	3669	3820
Adj. R2	0.59	0.55	0.51	0.42	0.37	0.36	0.26	0.21	0.17	0.14

Panel B: nowcasting and forecasting earnings surprises with different blockchain categories					
	SUE_{q+1}	SUE_{q+2}	CAR_{q+1}	CAR_{q+2}	
$BDG_{q+1}^{Operation}$	0.176***	0.126**	1.331*	0.991	
	(2.98)	(2.00)	(1.81)	(1.41)	
$BDG_{q+1}^{Financials}$	0.167**	0.112**	1.236*	0.851	
	(2.47)	(2.02)	(1.81)	(1.24)	
$BDG_{q+1}^{Humanresources}$	0.071	0.048	0.536	0.398	
	(1.15)	(0.94)	(0.72)	(0.52)	
$BDG_{q+1}^{Marketing}$	0.066	0.055	0.513	0.375	
	(1.17)	(0.82)	(0.83)	(0.61)	
BDG_{q+1}^{IT}	0.061	0.049	0.455	0.379	
	(1.12)	(0.74)	(0.72)	(0.53)	
$BDG_{q+1}^{Supplychain}$	0.088	0.060	0.647	0.509	
	(1.40)	(0.98)	(0.95)	(0.65)	
BDG_{q+1}^{Others}	0.067	0.054	0.458	0.414	
	(1.11)	(0.93)	(0.78)	(0.57)	
Controls	Y	Y	Y	Y	
Industry FE	Y	Y	Y	Y	
Year-Quarter FE	Y	Y	Y	Y	
N	4964	4936	5467	5453	
Adj. R2	0.43	0.36	0.08	0.07	

Panel C: nowcasting and forecasting returns and alphas with different blockchain categories					
	Excess	HXZ	FF5	LSY3	LSY4
$BDG_{q+1}^{Operation}$	0.262***	0.272***	0.351**	0.239**	0.231**
	(3.21)	(3.11)	(2.57)	(2.34)	(2.09)
$BDG_{q+1}^{Financials}$	0.247***	0.242***	0.281***	0.244**	0.218**
	(2.66)	(3.15)	(2.59)	(2.12)	(2.04)
$BDG_{q+1}^{Humanresources}$	0.121	0.109	0.121	0.096	0.090
	(1.29)	(1.21)	(1.09)	(0.91)	(0.92)
$BDG_{q+1}^{Marketing}$	0.108	0.109	0.124	0.101	0.105
	(1.37)	(1.38)	(1.08)	(0.89)	(0.86)
BDG_{q+1}^{IT}	0.096	0.091	0.127	0.094	0.091
	(1.20)	(1.10)	(0.99)	(0.82)	(0.77)
$BDG_{q+1}^{Supplychain}$	0.145	0.124	0.152	0.121	0.121
	(1.49)	(1.39)	(1.13)	(1.11)	(1.04)
BDG_{q+1}^{Others}	0.107	0.103	0.137	0.099	0.105
	(1.20)	(1.21)	(1.00)	(0.84)	(0.88)
Controls	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y
N	15674	16786	16517	17604	17817
Adj. R2	0.09	0.07	0.07	0.07	0.07

Table 18: Predictive Power of BDG in Supply Chain

This table presents results on the predictive power of BDG in the supply chain. Panel A reports the results on the regressions of firm fundamentals measured in quarter $q+1$ or quarter $q+2$ on the BDG predictability of customer-supplier link and supplier-customer link in quarter q+1 and other control variables in quarter q. The dependent variables include return on assets (ROA), growth of total assets (AG), growth of sales (SG), the log of one plus quarterly number of patents applied (PA), and the log of one plus quarterly number of patents granted (PG). Panel B reports the results on the regressions of earnings surprise measured in quarter $q+1$ or quarter $q+2$ on the BDG predictability of customer-supplier link and supplier-customer link in quarter $q+1$ and other control variables in quarter q. The dependent variables include standardized unexpected earnings (SUE) and earnings announcement abnormal returns (CAR). We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and one standard deviation. The control variables are from Panel A of Table 1. The t-statistics of robust standard errors clustered at the industry and year-quarter levels are reported in parentheses. Panel C reports the average monthly long-short excess returns and long-short alphas on the value-weighted portfolios of stocks sorted by the BDG predictability of customer-supplier link and supplier-customer link. Newey and West (1987) adjusted t-statistics are given in parentheses. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively. The sample period is from April 2016 to September 2021.

Panel A: nowcasting and forecasting firm fundamentals using BDG in the supply chain

	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}
Customer - Supplier										
BDG_{q+1}	0.661***	0.393***	0.242***	0.199***	0.049***	0.050**	0.314***	0.186***	0.184***	0.108***
	(4.82)	(2.60)	(4.79)	(3.20)	(3.87)	(2.08)	(3.66)	(2.90)	(3.64)	(3.09)
N	7549	7469	7240	7164	7394	7316	5392	5335	5392	5335
Adj. R2	0.51	0.47	0.43	0.35	0.30	0.30	0.22	0.18	0.14	0.12
Supplier - Customer										
BDG_{q+1}	0.575***	0.436*	0.205***	0.179**	0.048***	0.030*	0.237***	0.142**	0.128***	0.097**
	(3.28)	(1.79)	(4.21)	(2.42)	(3.35)	(1.74)	(3.02)	(2.10)	(3.42)	(2.27)
N	7549	7469	7240	7164	7394	7316	5392	5335	5392	5335
Adj. R2	0.45	0.41	0.38	0.31	0.27	0.26	0.19	0.16	0.13	0.10
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

Panel B: nowcasting and forecasting earnings surprises using BDG in the supply chain					
	SUE_{q+1}	SUE_{q+2}	CAR_{q+1}	CAR_{q+2}	
Customer - Supplier					
BDG_{q+1}	0.177***	0.177**	1.881***	1.267*	
	(3.58)	(2.17)	(2.94)	(1.91)	
N	7317	7240	7472	7392	
Adj. R2	0.36	0.3	0.06	0.06	
Supplier - Customer					
BDG_{q+1}	0.133**	0.139	1.742**	1.155*	
	(2.55)	(1.64)	(2.22)	(1.83)	
N	7317	7240	7472	7392	
Adj. R2	0.32	0.27	0.06	0.05	
Controls	Y	Y	Y	Y	
Industry FE	Y	Y	Y	Y	
Year-Quarter FE	Y	Y	Y	Y	
Panel C: long-short excess returns and alphas using BDG in the supply chain					
Value-weighted	Excess	HXZ	FF5	LSY3	LSY4
Customer - Supplier					
L/S	0.737***	0.652***	0.491***	0.615***	0.600***
	(3.66)	(3.12)	(3.43)	(2.70)	(2.93)
Supplier - Customer					
L/S	0.670***	0.486**	0.568***	0.428**	0.489**
	(2.84)	(2.17)	(2.89)	(2.11)	(2.24)

Table 19: Blockchain Peer Data Growth

This table tests the nowcasting and forecasting effect of blockchain peer data growth (BPDG). Other firms are blockchain peers to the focal firm in the same blockchain. We construct the equal-weighted blockchain peer data growth as the new variable to examine the nowcasting and forecasting effect to firm fundamentals, earnings surprise, and stock returns of the focal firms. Panel A reports the results on the regressions of firm fundamentals measured in quarter $q+1$ or quarter $q+2$ on the BPDG in quarter $q+1$ and other control variables in quarter q across different firms. The dependent variables include return on assets (ROA), growth of total assets (AG), growth of sales (SG), the log of one plus quarterly number of patents applied (PA), and the log of one plus quarterly number of patents granted (PG). Panel B reports the results on the regressions of earnings surprise measured in quarter $q+1$ or quarter $q+2$ on the BPDG in quarter $q+1$ and other control variables in quarter q across different firms. The dependent variables include standardized unexpected earnings (SUE) and earnings announcement abnormal returns (CAR). We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and one standard deviation. The control variables are from Panel A of Table 1. The t-statistics of robust standard errors clustered at the industry and year-quarter levels are reported in parentheses. Panel C reports the average monthly long-short excess returns and long-short alphas on the value-weighted portfolios of stocks sorted by the BDG predictability of customer-supplier link and supplier-customer link. Newey and West (1987) adjusted t-statistics are given in parentheses. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively. The sample period is from April 2016 to September 2021.

Panel A: nowcasting and forecasting firm fundamentals using BPDG in the supply chain

	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}
$BPDG_{q+1}$	0.599***	0.393***	0.290***	0.176***	0.052***	0.051**	0.326***	0.168***	0.149***	0.117***
	(5.61)	(2.75)	(4.67)	(2.94)	(4.25)	(2.32)	(3.75)	(2.86)	(3.26)	(2.88)
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	0.45	0.41	0.38	0.31	0.27	0.26	0.19	0.16	0.13	0.10

Panel B: nowcasting and forecasting earnings surprises using BPDG

	SUE_{q+1}	SUE_{q+2}	CAR_{q+1}	CAR_{q+2}
$BPDG_{q+1}$	0.211***	0.143**	1.769***	1.416**
	(3.81)	(2.01)	(3.51)	(2.11)
Controls	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Year-Quarter FE	Y	Y	Y	Y
N	10922	10806	11152	11033
Adj. R2	0.32	0.27	0.06	0.05

Panel C: long-short excess returns and alphas using BPDG

Value-weighted	Excess	HXZ	FF5	LSY3	LSY4
L/S	0.872***	0.642***	0.514***	0.559***	0.507***
	(2.98)	(2.76)	(2.99)	(3.12)	(2.66)