

Data-driven Value-enhancing Strategies: How to Increase Firm Value Using Data Science

Hyoung-Goo Kang^a, Ga-Young Jang^{b,*}, Moonkyung Choi^c

^a Associate Professor, Department of Finance, Hanyang University Business School, Korea

^b Adjunct Professor, Department of Finance, Hanyang University Business School, Korea

^c Ph.D. Asset Management Team, The Ministry of Employment and Labor (MOEL)

ABSTRACT

This paper proposes how to design and implement data-driven strategies by investigating how a firm can increase its value using data science. Drawing on prior studies on architectural innovation, a behavioral theory of the firm, and the knowledge-based view of the firm as well as the analysis of field observations, the paper shows how data science is abused in dealing with meso-level data while it is underused in using macro-level and alternative data to accomplish machine-human teaming and risk management. The implications help us understand why some firms are better at drawing value from intangibles such as data, data-science capabilities, and routines and how to evaluate such capabilities.

Keywords: Architectural Innovation, A Behavioral Theory of the Firm, Data-driven Strategy, Knowledge-based View

I. Introduction

When a firm's value is primarily dependent on its data and data-science capabilities, the firm's business model is called 'data-driven' (Schaefer et al., 2017; Sorescu, 2017). The valuation gap between the firms that have a data-driven model and the firms that do not continues to increase (Brynjolfsson et al., 2011). A data-driven value-enhancing strategy aims to increase a target firm's value by maximizing

the value of data and data-science capabilities while integrating them as the firm's strategic resources, leading to the competitive advantage of the firm (Barney, 1986; Conner, 1991; Conner and Prahalad, 1996; Dierickx and Cool, 1989).

The recent outbreak of COVID-19 has cemented the dominance of big technology companies that integrate data into their business model. For instance, the value of large technology firms increased throughout the crisis. As of May 2020, technology firms such

*Corresponding Author. E-mail: kathy.g.jang@gmail.com

as Apple, Microsoft, Alphabet, Amazon, and Facebook accounted for 20% of the S&P 500, becoming the true survivors of the crisis. This phenomenon could indicate that the success of a business would depend on the utilization of data, and that the success will be further reinforced upon the occurrence of other similar events that entail uncertainty (e.g., climate change, global power struggle, artificial intelligence, and inequalities). Data will continue to be an important intangible asset or a strategic resource of a company that determines the value of a business (Pongpisutsopa et al., 2020).

Other examples can be found in the financial sector. Although both banks and fintech companies compete in the same strategic group (Cool and Schendel, 1988; Dess and Davis, 1984; Porter, 1985), their valuation ratios differ significantly and so do their abilities to handle data. If traditional financial institutions like banks can handle data as fintech firms do, they may also have a good chance of increasing their firm value. Hence, data-driven value-enhancing strategies may be particularly necessary for the banking sector which has suffered from low valuations for a long time. Then, how should one design and implement such strategies? This paper addresses this important question, which the existing literature has largely overlooked.

In the financial sector, private equity firms regard value-enhancing as an important goal of their 'investment model' (Pomerance and McCarthy, 2018).¹⁾ Venture capitalists often call value-enhancing 'company-making' (Chesbrough, 2002; Hisrich and Jankowicz, 1990). Value-enhancing is hence the value proposition of private equity firms and venture

capitals. Considering this, it is no wonder that they have recently become interested in value-enhancing strategies based on data. According to our interviews with the experts in the field, several private equities regard data-driven value-enhancing at the core of their investment model and advertise its scheme as a fundraising theme.²⁾ Similarly, a leading venture capitalist has started applying data science even for the traditional deal-sourcing and risk management roles. For instance, an expert in one large private equity firm said, "We are interested in firms whose data capabilities are underestimated. Putting data experts in the top management team of such firms would increase firm value." Another said, "... Instilling data-driven and evidence-based routines in organizational processes is a value-enhancing strategy." In addition, a venture capitalist said, "Data scientists can become good venture capitalists because we use data science for deal sourcing as well as for value-increasing." Another VC said, "We are looking for startups with data so we can increase their value." It is not a secret that sophisticated private equities and venture capitalists are aggressively hiring data scientists. Similarly, such data-driven strategies would matter for other organizations as they care about the returns to their stakeholders. Data-driven value-enhancing strategies would also be beneficial for investors and the national economy in line with the fourth industrial revolution.

Meanwhile, successful data-driven strategies require appropriate guidelines. Nonetheless, to the authors' knowledge, there are no prior studies or practical guidelines on this matter yet. This study aims to fill the gap in the literature and practice.

1) A similar term, *value creation*, is more often used by private equity firms, but in this paper, we use the term, *value-enhancing*, instead to highlight the data-driven strategy's role in increasing the preexisting value of a firm.

2) Although the paper primarily draws from the existing literature to make suggestions, it also uses interviews to find evidence from the field. A brief description of the interviews is included in <Appendix A>.

Furthermore, we critically evaluate a number of different value-enhancing strategies and suggest what a company lacking data-science capabilities can do to implement the strategy to increase firm value.

The remainder of this paper is structured as follows. Section 2 discusses a common misconception about data-driven value-enhancing strategies and where data science tends to be abused or misleading. Section 3 discusses where data science tends to be underused. Section 4 proposes frameworks to develop and deploy data-science capabilities. The final section discusses the limitations and concludes.

II. Abusing Meso-Level Data

The company, GAP Inc., is famous for its innovative use of data.³⁾ It is a global clothing and accessories retailer based in the United States and owns other brands such as Old Navy and Banana Republic. However, the company has stagnated since the 2000s. During this period, fashion brands such as ZARA and Uniqlo overtook GAP in the market. While in trouble, the company appointed Mr. Art Peck as the CEO of GAP Inc., who, right after being appointed, dismissed a number of creative directors who were considered the quintessential assets in the fashion industry as they predicted fashion trends. At that time, the design themes of fashion houses used to be based on their predictions.

Nevertheless, Art Peck insisted that GAP should rely less on the intuitions of creative directors or designers, and more on using data for decision making. Mr. Peck attempted to play a version of ‘Money Ball’ (Lewis, 2004) in the fashion industry.

3) The following article presents further stories about GAP’s data strategy, <https://hbr.org/podcast/2018/11/could-big-data-replace-the-creative-director-at-the-gap>

The fashion industry was aghast and against Mr. Peck’s view. There were worries that creativity might disappear, and the sector would be distorted. Such reactions might be natural because the stakeholders who were comfortable with their existing status would resist any disruptive changes (Choi, 2019). For example, while incumbents verbally advocate innovation, many of them turn against it when innovation takes place. Innovators are often criticized and dismissed for being naïve and not understanding the inherent nature of an organization, industry, and practice.

Was Mr. Peck’s attempt successful? From the financial perspective, it is hard to say because GAP’s stock performance had been lukewarm compared to Inditex⁴⁾ and SPDR from 2016/01/01 to 2020/06/08. As a result, Art Peck resigned in November 2019, taking responsibility for the company’s poor performance. Then, what made the data-driven strategy of Art Peck unsuccessful?

We argue that the mixed performance of Art Peck’s data-driven strategy is attributable to the practice of abusing meso-level data. Let us first explain what ‘meso-level’ means. In general, a meso-level falls between the micro and macro levels.⁵⁾ A meso-level is often concerned with more local factors (e.g., communities, organizations) than a macro-level (e.g., legal, regulatory, economic), but is not as narrow or specific as a micro-level (e.g., individuals). Meso-level data arise at the group or organization level. A group or a firm’s daily, weekly, and monthly financial performance or behavioral data are examples of

4) “Inditex is one of the world’s largest fashion retailers, with eight brands (Zara, Pull&Bear, Massimo Dutti, Bershka, Stradivarius, Oysho, Zara Home, and Uterqüe) selling in 202 markets through its online platform or its over-7,000 stores in 96 markets.” <https://www.inditex.com/en/about-us/who-we-are>

5) According to the description of “Level of Analysis” by Wikipedia.

meso-level data. With few exceptions (e.g., stock prices), it is expensive to generate meso-level data in real-time because one needs to aggregate micro-level data first. This explains why meso-level data is often used for establishing short- or mid-term strategies.

Given the nature of meso-level data, it is hard to apply sophisticated techniques (e.g., deep learning) to such data. Even if one collects daily data for ten years, the data is insufficient to develop any deep learning model. Imagine a problem of predicting a fashion trend of a customer segment for the next season using the historical data about the segment's fashion trend for the last ten years. It would be difficult to develop an advanced machine-learning model. Similarly, imagine predicting the daily returns of KOSPI 200 using 10 years of data such as hosts of financial, accounting, and even unstructured big data. In such cases with short time-series data ($T = 252 \times 10$), one should be highly selective and careful to include a large number of independent variables in estimating models. To avoid abusing the data, then, we argue that one would need decent and academically grounded theories to complement the limitations of the stock-market data.

In reality, data science is abused in many cases. The reason is simple. Meso-level data appears easy to analyze. Anyone who learned basic statistics can create a simple algorithm. If one learns machine learning in a class, one can develop a simple trading strategy using daily stock returns or algorithms to predict future trends and make simulated profits. Indeed, such tasks are often given as homework assignments in university classes. Furthermore, meso-level data is sometimes easier to obtain than micro-level data, which is frequently subject to privacy issues. Since it appears easy at first glance to obtain and analyze meso-level data, it is no wonder that many firms and data consultants advertise their

'data-driven' services in formulating strategies, trading algorithms, predictions, and intelligence that require producing meso-level insights for short- or mid-term forecasting horizons (Ka and Kim, 2014). However, these approaches have serious problems that have not been discussed in the literature yet.

Let us assume we are faced with the problem of predicting the performance of a business strategy (e.g., marketing a particular fashion style). Suppose that the size of the data is $N = T \times K$, where T and K represent the number of observations (length) and the number of feature variables (feature size), respectively. Data is called 'big data' when N is large. There are two ways to make big data (i.e., increase N) by increasing T or increasing K . Academic researchers tend to highlight T , the length of a dataset. However, interestingly, industry practitioners tend to focus more on K . Some possible reasons would be that it is easier to increase K than T or that it is convenient to develop a marketing message with K , such as "we have a new dataset to predict the stock market." or "only our firm has access to the feature and will use it to develop new products." For example, when predicting the performance of a company, one can easily increase K by adding more variables such as corporate accounting data, crawled text data, macroeconomic data, and other unstructured data (Kang et al., 2019).

Yet, increasing K can complicate the problem rather than solve it because of the curse of dimensionality. As we collect more features (K) to predict the performance of a firm, the length of data, T , should increase in proportion to K^2 . In addition, if we try millions of tossing coins simultaneously, one of them would fit the trend perfectly ($R^2 = 100\%$). This is the curse of dimensionality.

Another problem is that T is often difficult to increase (e.g., the length of time is limited). However,

one can determine the distribution of a variable only when T is large. For example, if one coin is tossed millions of times, the probabilities of getting heads and tails will converge to the actual probability. Then, how do we solve this problem that is inherent in meso-level data? Would a firm need to hire world-class experts in artificial intelligence or machine learning to solve this problem to predict the future trend or customer behavior?

In this situation, academic theories or business acumen would matter more than technology, i.e., “strategy, not technology, drives digital transformation” (Kane et al., 2015). Meso-level analysis does not require a world-class AI expert but calls for the joint work of an experienced field expert and an econometrician who can deal with identification problems (Angrist and Pischke, 2008; Roberts and Whited, 2013). Thus, theories compensate for the lack of data. Intuitively speaking, if we can observe everything, why does one need a theory? Whether a strategy works or not is a problem of causality, which in turn calls for addressing identification problems. Simply collecting a large number of data points or using a highly complex model does not automatically guarantee the identification of causality because one needs specific techniques (e.g., “Mostly Harmless Econometrics”, Angrist and Pischke, 2008) to solve identification problems.

For this reason, a firm needs to be careful when working with consulting firms or data scientists who conduct deep learning without grounded theories. Often, their actions may be no more than data mining and, thus, lead to false results. Special skills or experiences are necessary to build and test hypotheses in the framework of ‘mostly harmless econometrics.’ In this regard, we present a four-step strategy for a firm interested in hiring a data consultant to develop ‘intelligence’ on meso-level data analysis as follows:

Step 1. Ask about the data structure (i.e., N , T , K at $N = T \times K$) and the approach to address the curse of dimensionality.

Step 2. Ask what theories or frameworks are used to design strategies, which is equivalent to hypothesis formulation in academia (e.g., a firm’s brand equity would increase in a customer segment if the firm conducted strategy X).

Step 3. Ask what kind of identification strategies are used to test their hypotheses.

Step 4. Avoid working with a data consultant who does not present plausible answers to the previous questions.

It is a simple yet useful guideline for the firm to avoid hiring the wrong personnel for the job.

III. Macro-Level Data and Scenario Planning

Scenario analysis is essential to designing mid- to long-term strategies. The inputs to scenario planning are views about macroeconomic or socio-political variables (Ramirez and Wilkinson, 2016). To formulate the views, one needs to collect and analyze macro-level data. However, logically speaking, if it is difficult to apply machine learning to meso-level data, it would be more difficult with macro-level data to forecast macro-level events. For example, suppose that we are interested in oil prices in the future. A 20-year data, which includes only 5,000 observations (20 years \times 252 business days), would make it difficult to practice deep learning, which requires the estimates of thousands of parameters. While it is possible to apply deep learning to transaction data in seconds or real-time data, such data are not commonly used or practical for establishing mid- to long-term

strategies.

Despite the above limitations, data-driven algorithms can be surprisingly helpful for formulating mid- to long-term strategies (Kang et al., 2019). According to behavioral economics, people tend to exhibit behavioral biases in mid- to long-term decision-making. The longer the horizon, the larger the biases. Planning fallacies, hyperbolic discounting, and optimism bias are examples of such biases. In particular, humans tend to make decisions using mental shortcuts on instant events, which is called availability heuristics. In conclusion, while it is difficult for machines to make mid- to long-term decisions, it is even more difficult for humans to do so.

To solve this, machine-human complementarity may be useful (Kleinberg et al., 2018; Lyons et al., 2019; Walliser et al., 2019). For example, data-driven algorithms in combination with human intuition can enhance decision-making and lead to data-driven value-enhancing. In their book called, *Zero to One*, Thiel and Masters (2014) explain this relationship as follows:

...men and machines are good at fundamentally different things. People have intentionality - we form plans and make decisions in complicated situations. We're less good at making sense of enormous amounts of data. Computers are exactly the opposite: they excel at efficient data processing, but they struggle to make basic judgments that would be simple for any human (p. 143).

In summary, machine algorithms can assist and double-check the decisions of humans and organizations by incorporating behavioral theories (Cyert and March, 1963) and behavioral economics in data-based algorithms. Furthermore, an organization can continuously update the model with practical information from its industry and the growing body

of literature, which will in turn lead to enhanced absorptive capacity (Cohen and Levinthal, 1990) and learning in the organization (March, 1991). This iterative routine will be the basis of data-driven value-increasing.

The area where data can create large value is enterprise risk management (ERM) (Kang et al., 2019). Risk management is regarded as one of the most important elements in a decision-making process, especially for financial and human-resource decision-making. However, we argue that data science is underused in ERM processes. In particular, firms tend to ignore Knightian uncertainty (Keynes, 1921; Knight, 1921) although the era of Knightian uncertainty (which includes pandemic, climate change, geopolitical competition, wealth gap, the future of capitalism, artificial intelligence, etc.) has arrived (Kang et al., 2018; Kim et al., 2022). According to Keynes (1921), Knightian uncertainty arises when the decision maker cannot quantify the uncertainty in decision making. For instance, Knightian uncertainty can arise when a state is too complex or ambiguous so that even its probability distribution cannot be specified. Keynes (1937) explains the Knightian uncertainty as follows:

By "uncertain" knowledge, let me explain, I do not mean merely to distinguish what is known for certain from what is only probable. The game of roulette is not subject, in this sense, to uncertainty... Or, again, the expectation of life is only slightly uncertain. Even the weather is only moderately uncertain. The sense in which I am using the term is that in which the prospect of a European war is uncertain, or the price of copper and the rate of interest twenty years hence, or the obsolescence of a new invention, or the position of private wealth owners in the social system in 1970. About these matters there is no scientific basis on which to form any calculable probability whatever. We

simply do not know. Nevertheless, the necessity for action and for decision compels us as practical men to do our best to overlook this awkward fact and to behave exactly as we should if we had behind us a good Benthamite calculation of a series of prospective advantages and disadvantages, each multiplied by its appropriate probability, waiting to be summed. (pp. 213-214)

Intuitively, risk managers should collect information from the media, industry sources, and communities, to forecast Knightian uncertainty. However, in reality, managers do not have enough time to read the Financial Times thoroughly, for example. This is where a machine can help humans. A machine can collect unstructured data and formulate an early warning system.

Furthermore, it would constitute a routine to develop a data-driven culture (Waller, 2020) to address Knightian uncertainty by attempting to quantify it. The Ministry of Employment and Labor (MOEL) of Korea is a good example of adopting an early warning system (EWS) based on unstructured data. MOEL's asset management team has been using the EWS system since early 2019, and they successfully predicted the serious economic harm associated with the COVID-19 risk as early as January 2020, two months ahead of the global economic shocks.

IV. The Framework for Data-driven Value-enhancing Strategies

This section presents general frameworks and guidelines regarding how to increase firm value and find investment opportunities grounded on data and data science capabilities.⁶⁾ To address the issues, we suggest three data-driven value-enhancing strategies for different organizational goals summarized in <Table 1>.

To use data to accomplish value-enhancing and innovation, a firm needs to consider at least [1] what innovations to pursue (data-driven innovation strategy), [2] how to change its organization to implement the strategy (data-driven organizational strategy), and [3] how to keep the changes sustainable to be competitive in the market (data-driven organizational learning for sustainable competitive advantage). In the context of value-enhancing strategies, we specifically highlight architectural innovation (Henderson and Clark, 1990) as a data-driven innovation strategy, a Behavioral Theory of the Firm (BTF) (Cyert and March, 1963) as a data-driven organizational strategy, and the Knowledge-based View of the firm (KBV) (Cohen and Levinthal, 1990; Kogut and Zander, 1996) to build data-driven organizational learning processes suitable for a sustainable competitive advantage. The following subsections discuss the three approaches,

<Table 1> Data-driven Value-enhancing Strategies for Difference Organizational Goals

| No. | Considerations | Approaches | Key Literature |
|-----|-----------------------------------|---|--|
| 1 | Innovation Strategy | Architectural Innovation | Henderson and Clark (1990) |
| 2 | Organizational Change | A Behavioral Theory of the Firm (BTF) | Cyert and March (1963) |
| 3 | Sustainable Competitive Advantage | Knowledge-based View (KBV); Absorptive Capacity (AC) | Cohen and Levinthal (1990); Kogut and Zander (1996) |

6) Kessler (2019) calls data “the world’s most valuable resource” in his article, *Data Protection in the Wake of the GDPR: California’s Solution for Protecting*.

architectural innovation, BTF, and KBV, respectively.⁷⁾

4.1. Architectural Innovation

Despite their expertise in their field, managers in low-valuation industries (e.g., banks) tend to lack a data science background (e.g., coding) and thus have difficulty in understanding and applying technological innovations such as unstructured data and deep learning. A problem is that the more successful they have been, the less likely they are to adopt innovations, which in turn prompts their downfall, i.e., the innovator's dilemma (Christensen, 2013). Furthermore, the academic literature grows too fast and is too complex for the managers to understand the implications to apply them systematically.

On the other hand, data scientists who lack field experience could mindlessly undertake data mining. For example, data scientists may repeat trials and errors without truly understanding the value proposition of the fields and the contexts of the sectors where they apply their data capabilities. This naturally leads to spurious results. Not surprisingly, managers are already becoming skeptical about artificial intelligence (Conkle, 2020). In particular, the problem may be more serious in areas such as competitive strategies and investment decision-making, where it is important to combine data science with academic theories and practical experiences from related fields such as economics and business management in order to resolve the identification problems (e.g., causality vs. correlation).

Neither data scientists nor field experts alone can solve this problem. Perhaps it is only the data scientists with sufficient industry experience, knowledge of aca-

demical theories, and practice who can solve the problem. However, only a handful of large technology companies or hedge funds can afford such experts. Alternatively, if a company can provide a service that solves the problem of combining data science expertise and field expertise regarding the Fourth Industrial Revolution with easy UI/UX (like iPhones), it will be a tremendous success.⁸⁾ However, such a service is yet to exist. As such, the problem of combining field expertise with data science remains an important challenge for both small startups and large institutions. Given these challenges, what kind of data-driven innovations should a firm develop to enhance the firm value using data science?

Henderson and Clark (1990) classify innovation into four categories as shown in Panel A in <Table 2>. The strategy for pursuing innovation has two axes: the concept and the relationship between concepts. The concept axis is about either reinforcing or overturning concepts. The relationship axis is about either maintaining or changing the relationship between concepts. In the end, four types of innovations are derived: incremental innovation, modular innovation, architectural innovation, and radical innovation.

Among these innovations, which one is the appropriate model to solve the problems of traditional firms with low valuations that are keen on data-driven value-enhancing strategies? Panel B in <Table 2> describes our suggestions. First, radical innovation (i.e., overturning concepts and changing their relationship) is overly expensive and time-consuming for most organizations except for technology giants (e.g., big techs) or large hedge funds. Concerning data, radical innovation requires developing innovative artifi-

7) We attach specific guidelines to implement data-driven value-enhancing strategies for organizations following the existing literature in <Appendix B>.

8) This is the vision of a fintech startup, Handa Partners (<http://www.handapartners.com>).

<Table 2> Innovation Frameworks

| <i>Panel A: Henderson & Clark (1990)'s Classification</i> | | | |
|---|---|---|---|
| | | Core Concept | |
| | | Reinforced | Overtured |
| The linkage between core concepts and components | Unchanged | Incremental | Modular |
| | Changed | Architectural | Radical |
| <i>Panel B: Innovation Types and Data Application</i> | | | |
| Innovation Type | Description | Data Application | Target Area |
| Incremental | Conventional meaning and connection | Catch up on data capabilities | Catch-up projects |
| Modular | Relationship unchanged, but updated and enhanced information | Generate new data or different interpretations of existing data | Fintech startups |
| Architectural | Constant core meaning; different relationships between meanings | Reconfigure existing system | Large, but traditional financial institutions |
| Radical | New architecture; new concepts | Design a new system of knowledge and technologies | Big techs |

cial-intelligence techniques while simultaneously connecting new techniques in a creative manner, which not many firms are capable of or need at their current stage.

Second, incremental innovation, reinforcing concepts and maintaining their relationships, is what most companies are already attempting to do, but this strategy is vulnerable to the innovator's dilemma (Christensen, 2013). The theory of the innovator's dilemma warns of a situation in which a firm eagerly pursues the innovation targets that the current market desires, but eventually fails because of the very attempt. Hence, incremental innovation is possibly appropriate for firms that lack significant data capabilities compared to industry peers. They can start from small incremental strategies to learn from peers and possibly catch up to them if opportunities are open. However, incremental innovation itself would not achieve significant data-driven value-enhancing and, again, is vulnerable to the innovator's dilemma.

Third, modular innovation, which is overturning concepts while maintaining their relationships, fo-

cuses on collecting new data that are not normally used or creating new data science methods given existing frameworks or business relationships. However, collecting innovative data is appropriate only for those that already integrate data collection with their business model. For instance, Facebook collects precious individual-level data as part of its business model. Users 'pay' their data to enjoy Facebook. Purchasing innovative data is infeasible either because they are usually the core resources of leading big technology firms, and therefore those firms have no reason to trade the data. Even if a firm purchases precious data, its value would decrease significantly once the data is transferred to other organizations. Suppose that even if a firm obtained Facebook's data, it would be very difficult to use the data as effectively as Facebook because the data is an integral part of Facebook's business model and a core strategic resource of the firm. A core strategic resource is by definition socially complex, ambiguous, and not significantly transferrable (Barney, 1986) with "the characteristics of the asset accumulation process:

time compression diseconomies, asset mass efficiencies, inter-connectedness, asset erosion, and causal ambiguity” (Dierickx and Cool, 1989).

Furthermore, it is unnecessary to develop new artificial intelligence technologies for most companies because researchers in academia are consistently developing new machine learning models, and they often post their findings on Github or Gitlab for free. Just following them would be enough for most firms. In fact, since the speed of knowledge accumulation is so fast, it would be hard for most firms to follow the speed, let alone overtake it with new technologies. Therefore, we recommend pursuing modular innovations for startups grounded in universities or research institutes. If a startup succeeds, a larger firm can easily import its product as a “new module” in its existing business architecture. This can be done conveniently because modular innovation by definition does not require changing existing relationships between concepts.

This leaves us with only one choice, architectural innovation, for most traditional firms suffering from low valuations. It is relatively easy to customize and apply the existing models rather than developing a new machine learning model. Even undergraduate students can download and experiment with recent working papers and their codes posted on GitHub. In the end, it is important to grasp the specific business questions of the industry, and then combine existing resources to solve the questions. This constitutes an architectural innovation, reinforcing concepts while changing their relationship. Hiring a world-class artificial intelligence expert would not help to accomplish architectural innovation because she would be “pigeonholed within a company” (Waller, 2020).

However, even architectural innovation would not be easy if an organization lacks planning. The biggest challenge here is not technology. Rather, the hard

problem is organizational challenges and strategy. For instance, some practical data scientists revealed during our interviews that most traditional financial firms are not properly utilizing “even the data they already own, let alone drawing the knowledge from academic or open sources”. Again, “challenges reside in organizational issues” as several data scientists in banks put it. This issue will be described in the next subsection.

In conclusion, to accomplish architectural innovation, low-valued organizations do not need to hire expensive artificial intelligence experts as most people easily assume (Kim et al., 2007). Not every firm needs world-class experts in artificial intelligence and data science. Instead, they need to combine field experts who correctly identify business questions with the researchers who can draw a body of knowledge from available resources to test the experts’ intuition. To test the experts’ intuition, the researchers should be able to address identification problems (Angrist and Pischke, 2008) because developing recommendations inevitably leads to testing hypotheses. If they work together, even minimal levels of technology and data could have a great effect. Therefore, architectural innovations demand that a firm’s field experts and data scientists should creatively connect existing knowledge around specific business problems. This is again in line with the previously mentioned intuition of “strategy, not technology drives digital transformation” (Kane et al., 2015).

We argue that combining business problems and data strategies is important to undertake architectural innovations. Then, how do we characterize problems and identify data strategies? This question is answered in the next section.

4.2. A Behavioral Theory of the Firm (BTF)

The practical aspects of data strategies can vary and depend on the context of a firm. However, the literature seldom analyzes or generalizes such organizational heterogeneity in developing data strategies. Without any framework for understanding the sources for and patterns of diversity, how would it be possible to generalize the diversity and propose or develop a prescriptive data-driven value-enhancing strategy?

This study attempts to present a perspective framework based on a Behavioral Theory of the Firm (BTF) (Cyert and March, 1963). BTF is one of the main classical theories of the Carnegie School⁹⁾ and has become one of the most popular frameworks in the field of behavioral science, strategic management, organization theory, and information management. Among numerous implications of BTF, this paper focuses on the two critical elements of BTF, namely, *Knightian uncertainty* (Keynes, 1921; Knight, 1921) and *stakeholder conflict*. As mentioned earlier, Knightian uncertainty arises when the decision maker cannot quantify the uncertainty in decision-making. The existing literature argues that entrepreneurship is important in the case of high Knightian uncertainty (Knight, 1921; Mazzucato, 2011) and that the creation of shared value is important in the case of high stake-

holder conflict (Porter and Kramer, 2011). Then, depending on the degrees of Knightian uncertainty and stakeholder conflict, we can combine high and low entrepreneurship and shared-value creation. <Table 3> summarizes this intuition derived from prior studies (Kang et al., 2018; Kim et al., 2022).

More specifically, <Table 3> presents a framework consisting of four approaches for implementing data-driven value-enhancing strategies. The approaches vary depending on the contexts of the challenges that organizations face, which are categorized by the degrees of Knightian uncertainty and stakeholder conflict. In order to decide which approach to undertake, a firm should first identify its business challenges. Second, the firm should measure the degrees of Knightian uncertainty and stakeholder conflict for the problem. Third, the firm should construct a strategy combining entrepreneurship and shared value. Fourth, data scientists should draw a body of knowledge and resources and combine them around the formulated strategy to generate architectural innovation.

To accomplish architectural innovation and implement the associated data strategy, a firm's organization itself should be built for it. This is possibly the most difficult challenge as our interviewees noted. The next subsection discusses some of the potential solutions.

<Table 3> A Framework for Data-driven Value-enhancing Strategies

| Category | High Knightian Uncertainty | Low Knightian Uncertainty |
|---------------------------|---|--|
| High Stakeholder Conflict | High entrepreneurship + high shared value = Social entrepreneurship and nonmarket strategies | Low entrepreneurship + high shared value = Shared economy or opportunities in social impact |
| Low Stakeholder Conflict | High entrepreneurship + low shared value = Experiments and explorations | Low entrepreneurship + low shared value = Process innovation, digital twins |

9) "The Carnegie School was a so-called "Freshwater" economics intellectual movement in the 1950s and 1960s based at Carnegie Mellon University and led by Herbert

A. Simon, James March, and Richard Cyert." https://en.wikipedia.org/wiki/Carnegie_School. See also Gavetti et al. (2007) about neo-Carnegie school.

4.3. The Knowledge-Based View (KBV)

The Knowledge-Based View (KBV) (Cohen and Levinthal, 1990; Kogut and Zander, 1996) proposes the combination of tacit knowledge and the ability to utilize this knowledge (which rival companies cannot easily follow) tends to determine the competitive advantage of a firm. Collecting data or building data-science capabilities are not the final, but an intermediate goal of a firm. Generating a sustainable competitive advantage is usually the goal. Hence, data should become a valuable knowledge resource on which a firm can base its strategies to build a competitive advantage over one's rivals.

The Knowledge-Based View (KBV) emphasizes the importance of organizational learning for transforming data into organizational knowledge. Importantly, absorptive capacity (Cohen and Levinthal, 1990) determines the extent of organizational learning such as the ability to collect and apply information to create profit. Existing KBV literature highlights four determinants of absorptive capacity: *prior knowledge*, *incentive structure*, *organizational routine*, and *social network*. Specifically, the performance of organizational learning is decided by [1] what one is studying (prior knowledge) (Eisenhardt and Santos, 2002), [2] whether one is incentivized to study hard (incentive structure) (Kapoor and Lim, 2007), [3] how is one studying (routine) (Grant, 1996; Grant and Baden-Fuller, 1995), [4] from and with whom one is learning (social network) (Yli Renko et al., 2001).

Let us analyze each determinant in more detail. The first is prior knowledge. The fact that prior knowledge determines the absorptive capacity means that what data a company already owns would affect data-collecting and data-processing ability in the future. Many organizations claim that they do not have enough data, but they often do have data that is

useful. Firms just may not know what they have and what they can do with it. Sometimes, their managers intentionally ignore their data possibly because "their lives are already good enough."¹⁰ Many firms do not realize that their prior data could determine the future trajectory of their data capabilities. This is related to another determinant, incentive structure. In addition, the importance of prior knowledge is also related to combinative capabilities (Kogut and Zander, 1996), which is the capability to combine existing knowledge sourced from inside and outside the organization to acquire new skills. This is again in line with the importance of existing data and architectural innovation. We recommend firms realize that they are already generating data in real-time by their people, things they own, and organizational activities (e.g., Internet of Things). While firms complain about the lack of data, aren't they simply intimidated by the size of the data that they already have? Firms do not use 97% of the data they own (Sebastian-Coleman, 2018), and 87% of organizations lack data-science capabilities.¹¹ How to use such data about their clients and organizations will determine the success of data-driven strategies.

The second determinant is incentive structure. A particular problem with incentive structure is that data managers are often uncooperative and hostile in sharing data. This is because data managers are concerned that others may invade their work or discover errors and issues in their practices if the data that they are in charge of is shared. In addition, it is natural that they feel a great threat (or at least burden) of disclosing the data that they have been managing to the experts with advanced degrees or

10) Quotes from the interview.

11) <https://www.gartner.com/en/newsroom/press-releases/2018-12-06-gartner-data-shows-87-percent-of-organizations-have-low-bi-and-analytics-maturity>

publication capabilities. However, this situation should be resolved for any firm to increase the value of its data and capabilities. Therefore, an appropriate incentive structure should be built and shared with data managers so that data managers will openly share their data with experts and cooperate. Since incentives are inseparable from organizational culture, firms need to design an organizational culture about data, e.g., data-driven culture (Subrahmanyam and Jalona, 2020; Waller, 2020).

The third determinant is organizational routines (Dosi et al., 2001). One of the most common obstacles to the application of data science is poor organizational routines around data. For example, there are many cases where the table format and data structure are not systematically managed. If data is not well organized, data value fades and data science is not applicable. In many firms, IT experts decide on data tables and formats, but the problem is that they are not the ultimate users of the data. Ultimate users should decide how to tabulate and structure data. This will make a firm's data analysis more efficient. If users are not capable of such data structuring, a firm should consult academic researchers who have published similar types of problems or data. For instance, given our experiences, firms do not know how to structure panel data efficiently. With the possible help of academic researchers or academic-minded consultants, the front office, not the back office, should design a data dictionary, standard table format, and folder structure in an organization. If data is organized and managed according to the format defined by the front office, cooperation becomes easier. The data can be shared with other departments with APIs (application programming interfaces) in a standard format so that many data scientists can cooperate, and data may even be traded on data exchanges to generate extra revenues. We recom-

mend starting by constructing a 'research dataset' which is free from security, privacy, and other regulatory issues, so that a firm shares it with internal and external researchers. Eventually, the research dataset should be instrumental in overcoming data silos¹²⁾ and needs to evolve into the master data (metadata) of a firm, so that internal and external analysts broadly request access, and then analyze and crosscheck them to form a clear organizational consensus or a point of debate. This will prevent wasting time and energy.

V. Discussion and Conclusion

This paper addresses how to increase firm value using data and data science. To increase the value of a firm, one can increase the value of the firm's tangible or intangible assets. This paper discusses how the managers of a firm can enhance the firm value by enhancing data and data-science capabilities, which are considered the most important intangible assets of high-valued firms in particular today.¹³⁾

More specifically, the paper proposes instructions regarding how to formulate and undertake data-driven strategies to increase firm value, drawing on the existing literature on architectural innovation, a behavioral theory of the firm, and the knowledge-based view of the firm. Furthermore, the field observations and interviews are used to show where and how data science is abused in dealing with meso-level data, whereas it is underused for macro-level data and enterprise risk management to accomplish machine-human teaming.

Lastly, this paper is conceptual and draws insights

12) <https://techcrunch.com/2017/06/23/five-building-blocks-of-a-data-driven-culture/>

13) The Economist reported, "the world's most valuable resource is no longer oil, but data (May 6, 2017)."

from a literature review and qualitative research. The limitations of the study can be described as follows. First, more qualitative case studies would strengthen our implications and particularly help practitioners implement the strategies mentioned in this study. Second, the quantity and boundary conditions of the strategies could be estimated quantitatively in future studies. Third, the model can be extended to an industrial ecosystem, in which crucial competitions take place. For example, one could update the model for platform businesses that are meta-organizations, or hybrids of a market and an organization (Kretschmer et al., 2022). Increasing value economy-wide is a significant mission of any government. We believe that governments can play instrumental roles in implementing data-driven policies and supporting or even leading private sectors so that they can thrive in the fourth industrial revolution based on data-driv-

en value-enhancing strategies.

Acknowledgements

Dear Professor Byounggu Choi, Thank you so much for such great news. We are truly honored to have our manuscript published in APJIS at your discretion.

The camera-ready manuscript is now ready and will be submitted through the system along with other materials (i.e., a copyright agreement and bio). When editing the manuscript, we followed the instruction downloaded from the website. In any case you feel that additional changes are necessary, please kindly let us know so we can address them in time.

Again, thank for your kindness and consideration.

<References>

- [1] Angrist, J. D., and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton university press.
- [2] Barney, J. B. (1986). Strategic factor markets: Expectations, luck, and business strategy. *Management Science*, 32(10), 1231-1241. <https://doi.org/10.1287/mnsc.32.10.1231>
- [3] Brynjolfsson, E., Hitt, L. M., and Kim, H. H. (2011). *Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance?* (SSRN Scholarly Paper ID 1819486). Social Science Research Network. <https://doi.org/10.2139/ssrn.1819486>
- [4] Chesbrough, H. (2002, March 1). Making sense of corporate venture capital. *Harvard Business Review*, March 2002. <https://hbr.org/2002/03/making-sense-of-corporate-venture-capital>
- [5] Choi, B. (2019). The role of firm size and IT capabilities in open and closed innovation. *Asia Pacific Journal of Information Systems*, 29(4), 690-716. <https://doi.org/10.14329/apjis.2019.29.4.690>
- [6] Christensen, C. M. (2013). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business Review Press.
- [7] Cohen, W. M., and Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 128-152. <https://doi.org/10.2307/2393553>
- [8] Conkle, T. (2020). *Council Post: Tempering Your Company's Expectations For AI*. Forbes. <https://www.forbes.com/sites/forbestechcouncil/2020/09/10/tempering-your-companys-expectations-for-ai/>
- [9] Conner, K. R. (1991). A historical comparison of resource-based theory and five schools of thought within industrial organization economics: Do we have a new theory of the firm? *Journal of Management*, 17(1), 121-154. <https://doi.org/10.2307/259111>

- org/10.1177/014920639101700109
- [10] Conner, K. R., and Prahalad, C. K. (1996). A resource-based theory of the firm: Knowledge versus opportunism. *Organization Science*, 7(5), 477-501. <https://doi.org/10.1287/orsc.7.5.477>
- [11] Cool, K., and Schendel, D. (1988). Performance differences among strategic group members. *Strategic Management Journal*, 9(3), 207-223. <https://doi.org/10.1002/smj.4250090302>
- [12] Cyert, R. M., and March, J. G. (1963). A behavioral theory of the firm: Assessment and prospects. *Englewood Cliffs, NJ*, 2(4), 169-187. <https://doi.org/10.5465/19416520.2012.656841>
- [13] Dess, G. G., and Davis, P. S. (1984). Porter's (1980) generic strategies as determinants of strategic group membership and organizational performance. *Academy of Management Journal*, 27(3), 467-488. <http://dx.doi.org/10.2307/256040>
- [14] Dierickx, I., and Cool, K. (1989). Asset Stock Accumulation and Sustainability of Competitive Advantage. *Management Science*, 35(12), 1504-1511. <https://doi.org/10.1287/mnsc.35.12.1504>
- [15] Dosi, G., Nelson, R., and Winter, S. (2001). *The Nature and Dynamics of Organizational Capabilities*. OUP Oxford.
- [16] The Economist. (May 6, 2017). The world's most valuable resource is no longer oil, but data. Retrieved from <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>.
- [17] Eisenhardt, K. M., and Santos, F. M. (2002). Knowledge-based view: A new theory of strategy. *Handbook of Strategy and Management*, 1(1), 139-164.
- [18] Gavetti, G., Levinthal, D., and Ocasio, W. (2007). Perspective—Neo-Carnegie: The Carnegie school's past, present, and reconstructing for the future. *Organization Science*, 18(3), 523-536. <https://doi.org/10.1287/orsc.1070.0277>
- [19] Grant, R. M. (1996). Toward the knowledge-based theory of the firm. *Strategic Management Journal*, 17(S2), 109-122. <https://doi.org/10.1002/smj.4250171110>
- [20] Grant, R. M., and Baden-Fuller, C. (1995). The knowledge-based theory of inter-firm collaboration. *Academy of Management Proceedings*, 1995(1), 17-21. <https://doi.org/10.5465/ambpp.1995.17536229>
- [21] Henderson, R. M., and Clark, K. B. (1990). Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly*, 9-30. <https://doi.org/10.2307/2393549>
- [22] Hisrich, R. D., and Jankowicz, A. D. (1990). Intuition in venture capital decisions: An exploratory study using a new technique. *Journal of Business Venturing*, 5(1), 49-62. [https://doi.org/10.1016/0883-9026\(90\)90026-P](https://doi.org/10.1016/0883-9026(90)90026-P)
- [23] Ka, H. K., and Kim, J. S. (2014). An empirical study on the influencing factors for big data intended adoption: Focusing on the strategic value recognition and TOE framework. *Asia Pacific Journal of Information Systems*, 24(4), 443-472. <https://doi.org/10.14329/apjis.2014.24.4.443>
- [24] Kane, G. C., Palmer, D., Phillips, A. N., Kiron, D., and Buckley, N. (2015). Strategy, not technology, drives digital transformation. *MIT Sloan Management Review and Deloitte University Press*, 14(1-25).
- [25] Kang, H. G., Bae, K., and Ku, B. (2019). Artificial Intelligence and Financial Investment Strategy. *The Korean Economic Review*, 12(3), 93-110.
- [26] Kang, H. G., Woo, W., Burton, R. M., and Mitchell, W. (2018). Constructing MandA valuation: how do merger evaluation methods differ as uncertainty and controversy vary?. *Journal of Organization Design*, 7(1), 1-46. <http://dx.doi.org/10.1186/s41469-017-0025-y>
- [27] Kapoor, R., and Lim, K. (2007). The impact of acquisitions on the productivity of inventors at semiconductor firms: A synthesis of knowledge-based and incentive-based perspectives. *Academy of Management Journal*, 50(5), 1133-1155. <https://doi.org/10.5465/amj.2007.27169706>
- [28] Kessler, J. (2019). Data Protection in the Wake of the GDPR: California's Solution for Protecting "The World's Most Valuable Resource". *Southern California Law Review*, 93(1), 99-128. Retrieved from

- <https://southern-california-law-review.com/2019/11/01/data-protection-in-the-wake-of-the-gdpr-california-solution-for-protecting-the-worlds-most-valuable-resource-note-by-joanna-kessler/>
- [29] Keynes, J. M. (1921). *A Treatise on Probability*. Courier Corporation.
- [30] Keynes, J. M. (1937). The general theory of employment. *The Quarterly Journal of Economics*, 51(2), 209-223. <https://doi.org/10.2307/1882087>
- [31] Kim, E., Jang, G. Y., and Kim, S. H. (2022). How to Apply Artificial Intelligence for Social Innovations. *Applied Artificial Intelligence*, 1-22. <https://doi.org/10.1080/08839514.2022.2031819>
- [32] Kim, Y. J., Nam, K. C., Song, J. K., and Koo, C. M. (2007). An empirical investigation into the role of core-competency orientation and IT outsourcing process management capability. *Asia Pacific Journal of Information Systems*, 17(3), 131-146.
- [33] Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., and Mullainathan, S. (2018). Human decisions and machine predictions. *The Quarterly Journal of Economics*, 133(1), 237-293. <https://doi.org/10.1093/qje/qjx032>
- [34] Knight, F. H. (1921). *Risk, Uncertainty and Profit*. Courier Corporation.
- [35] Kogut, B., and Zander, U. (1996). What firms do? Coordination, identity, and learning. *Organization Science*, 7(5), 502-518. <https://doi.org/10.1287/orsc.7.5.502>
- [36] Kretschmer, T., Leiponen, A., Schilling, M., and Vasudeva, G. (2022). Platform ecosystems as meta organizations: Implications for platform strategies. *Strategic Management Journal*, 43(3), 405-424. <https://doi.org/10.1002/smj.3250>
- [37] Lewis, M. (2004). *Moneyball: The Art of Winning an Unfair Game*. WW Norton & Company.
- [38] Lyons, J. B., Wynne, K. T., Mahoney, S., and Roebke, M. A. (2019). Trust and human-machine teaming: A qualitative study. In *Artificial Intelligence for the Internet of Everything* (pp. 101-116). <https://doi.org/10.1016/B978-0-12-817636-8.00006-5>
- [39] March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1), 71-87. <https://doi.org/10.1287/orsc.2.1.71>
- [40] Mazzucato, M. (2011). The entrepreneurial state. *Soundings*, 49(49), 131-142. <https://doi.org/10.3898/136266211798411183>
- [41] Pomerance, J. M., and McCarthy, D. (2018). Private Equity Investing 2018: Rethinking the Private Equity Investment Model to Build Value. *The Journal of Private Equity*, 22(1), 10-16. <https://doi.org/10.3905/jpe.2018.22.1.010>
- [42] Pongpisutsopa, S., Thammaboosadee, S., and Chuckpaiwong, R. (2020). Factors affecting HR analytics adoption: A systematic review using literature weighted scoring approach. *Asia Pacific Journal of Information Systems*, 30(4), 847-878.
- [43] Porter, M. E. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance* (Illustrated Edition). Free Press.
- [44] Ramirez, R., and Wilkinson, A. (2016). *Strategic Reframing: The Oxford Scenario Planning Approach*. Oxford University Press.
- [45] Roberts, M. R., and Whited, T. M. (2013). Endogeneity in empirical corporate finance. In *Handbook of the Economics of Finance*, 2, 493-572. <https://doi.org/10.1016/B978-0-44-453594-8.00007-0>
- [46] Schaefer, D., Walker, J., and Flynn, J. (2017). A data-driven business model framework for value capture in Industry 4.0. *Advances in Transdisciplinary Engineering*, 6(2017), 245-250. <https://doi.org/10.3233/978-1-61499-792-4-245>
- [47] Sebastian-Coleman, L. (2018). *Navigating the Labyrinth: An Executive Guide to Data Management*. Technics Publications.
- [48] Sorescu, A. (2017). Data-driven business model innovation. *Journal of Product Innovation Management*, 34(5), 691-696. <https://doi.org/10.1111/jpim.12398>
- [49] Subrahmanyam, S. N., and Jalona, S. (2020, February 28). Building a Data-Driven Culture from the Ground Up. *Harvard Business Review*. Retrieved from <https://hbr.org/2020/02/building-a-data-driven-culture-from-the-ground-up>
- [50] Thiel, P. A., and Masters, B. (2014). *Zero to One: Notes on Startups, or How to Build the Future*.

Broadway Business.

- [51] Waller, D. (2020, February 6). 10 Steps to Creating a Data-Driven Culture. *Harvard Business Review*. Retrieved from <https://hbr.org/2020/02/10-steps-to-creating-a-data-driven-culture>
- [52] Walliser, J. C., de Visser, E. J., Wiese, E., and Shaw, T. H. (2019). Team structure and team building improve human-machine teaming with autonomous agents. *Journal of Cognitive Engineering and Decision Making*, 13(4), 258-278. <https://doi.org/10.1177/15553434198675>
- [53] Yli Renko, H., Autio, E., and Sapienza, H. J. (2001). Social capital, knowledge acquisition, and knowledge exploitation in young technology-based firms. *Strategic Management Journal*, 22(6-7), 587-613. <https://doi.org/10.1002/smj.183>.

<Appendix A> Description of the Interviews with Field Experts

- Interview methods
 - In-person or online meetings for introduction and questioning
 - Follow-up e-mails and phone calls for confirmation of the details

- Interviewees (anonymous)
 - Five quantitative analysts in large brokerage firms
 - Three fund managers in the two largest asset management firms
 - Four managers in the largest digital platforms
 - Five C-level executives in various startups and fintech companies
 - Two C-level managers and three managers in banks
 - Three strategy and IT consultants
 - All interviewees reside in Korea.

- Interview period
 - August 2017 - October 2021
 - Admittedly, most interviews took place during the COVID-19, which may have affected our data and, therefore, the results.

<Appendix B> Guidelines to Implement Data-driven Value-enhancing Strategies for Organizations Following the Existing Literature

| Theory | Organizational guidelines |
|---------------------------------------|---|
| Architectural innovation | <ul style="list-style-type: none"> ● Identify the business problems. ● Hire field experts or practitioners to address the problems. ● Transform the experts' knowledge into an empirical design. ● Combine the existing data and available (open-source) machine-learning codes to implement the empirical design. ● Iterate the above procedure by combining the business model, data, and machine-learning codes to derive architectural innovations. |
| A behavioral theory of the firm (BTF) | <ul style="list-style-type: none"> ● Identify the issues. ● Conceptualize and measure the key organizational variables in BTF (i.e., Knightian uncertainty and stakeholder conflicts) around the issues. ● Select one of four approaches to implement a data-driven value-enhancing strategy (see <Table 3>). |
| The knowledge-based view (KBV) | <ul style="list-style-type: none"> ● Identify prior knowledge (i.e., existing data or data science) as the basis for absorptive capacity (e.g., data collection and processing). ● Design an incentive structure and organizational culture to share data at least internally (e.g., make data managers and IT people less hostile about sharing their data). ● Update organizational routines in using the data (e.g., let the ultimate users decide how to tabulate the data, name folders and features (i.e., data dictionary) as well as APIs formats; make the research dataset easily accessible via APIs; use the data dictionary, research dataset, and APIs to form a clear organizational consensus or a point of debate). |

◆ About the Authors ◆



Hyoung-Goo Kang

Hyoung-Goo Kang is an associate professor of the department of finance at Hanyang University Business School and the deputy head at Hanyang Institute for Entrepreneurship in Seoul, Korea. He is also with the department of computational finance and financial engineering, Hanyang University Blockchain School and Hanyang University Blockchain Research Institute. His research focuses on machine learning (econometrics), financial innovations, innovative institutions, strategic process, asset allocation and the big data approach to behavioral economics. He received his PhD and MA from the Fuqua School of Business at Duke University, his MA from the economics department at the University of Virginia and his BA from the economics department at Seoul National University. He has worked at Lehman Brothers Quantitative Research, Accenture Management Consulting and Republic of Korea Air Force.



Ga-Young Jang

Ga-Young Jang, CFA, received her PhD from Hanyang University Business School. She taught undergraduate students in finance as an adjunct professor at Hanyang University Business School since Fall 2021. Prior to receiving her doctoral degree, she worked as a portfolio manager and a credit analyst for about ten years at various financial institutions. Her research focuses on credit markets, sustainability, and financial innovation. Currently, she works as a researcher and a member of the Diversity, Equity, and Inclusion Committee of the Strategic Management Society (SMS). She received her MA from Seoul National University, and BA from the University of Michigan, Ann Arbor.



Moonkyung Choi

Moonkyung Choi is the Deputy director of asset management team at Ministry of Employment and Labor, Republic of Korea. Dr. Choi currently serves as the member for Alternative investment committee and Performance evaluation committee of National Health Insurance Service, and Alternative Investment Strategy Committee Member of Local Finance Association. Prior to joining MOEL, Dr. Choi was with Korea Fund Rating, where he was team head of Institutional Investment Advisory. Before joining KFR, he was manager of risk management consulting at Korea Asset Pricing. He received the PhD, MA and BA from the Department of Economics at Yonsei University.

Submitted: February 18, 2022; 1st Revision: June 7, 2022; Accepted: August 16, 2022