Investing in Innovation

April 2022

Kai Wu

Founder & Chief Investment Officer

kai@sparklinecapital.com

Executive Summary

Innovation stocks have become a popular investment theme but are facing scrutiny. We build a half-century backtest of innovation that invests in a rotating portfolio of technologies trending in patent data. Innovation has delivered positive long-term returns distinct from growth and other traditional factors. While innovation is prone to speculative bubbles, this can be mitigated using valuation and other intangible pillars.

Introduction

The Death of Innovation

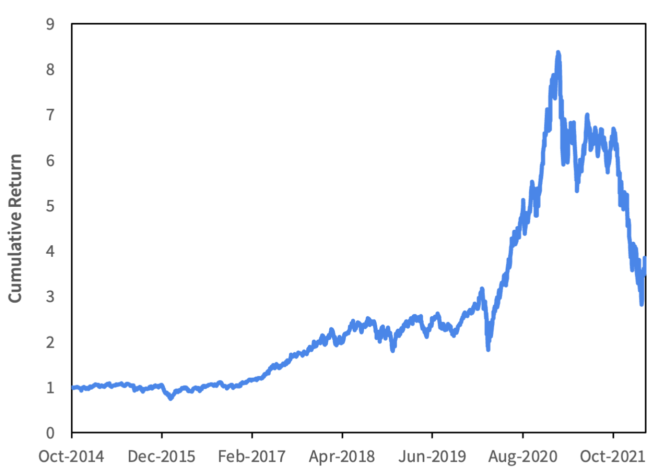
Over the past few years, investors have developed a growing appetite for so-called "innovation stocks." This has drawn a wave of disruptive startups, such as Coinbase and Zoom, to the public markets. It has also catapulted innovative firms, such as Tesla, to dominate news cycles and trading volumes.

Furthermore, this demand had led fund managers to launch dozens of thematic products focused on innovation. These funds package innovative stocks with compelling narratives, such as artificial intelligence, blockchain, and metaverse.

While early investors in innovation stocks were rewarded with phenomenal returns, the past year witnessed a sharp reversal. The flagship ARK Innovation ETF (ARKK), having fallen 60% from its highs, has become a lightning rod in a swirling debate over the future of innovation investing.

Exhibit 1

ARKK ETF Performance



Source: ARK Invest, Sparkline. As of 3/31/2022.

Innovation's apostles argue that volatility is an inherent feature of disruption and this discomfort is the reason its followers should expect to be rewarded. Amazon's 2,000X lifetime returns were only available to the true believers able to hold through the crucible of the dot-com crash.

On the other hand, a rising chorus of skeptics contend that innovation investing is nothing but a rebrand of growth investing, the discredited pied piper luring naive investors into speculative bubbles. They warn the recent selloff is just the start of a dot-com style bear market in innovation.

Defining Disruption

The debate rages in the financial media, taking on an almost religious fervor. However, despite the commotion, not a single rigorous empirical study exists on the long-term performance of innovation investing.

One reason is the lack of consensus on how to even define innovation investing. The dispersion in the top holdings of five representative innovation ETFs underscores this point. Tesla and Block are the only stocks in all five ETFs, and the average position-level correlation is a mere 11%.

Exhibit 2

Innovation ETF Top Holdings

ARKK	KOMP	DTEC	XT	GINN
ARK Innovation	*S&P Kensho New Economies*	*ALPS Disruptive Technologies*	*iShares Exponential Technologies*	*Goldman Sachs Innovate Equity*
Tesla	Teledyne	Cutera	CF Industries	Tesla
Teladoc	Elbit	Pagseguro	Junshi Bio	NVIDIA
Coinbase	Vonage	Dexcom	SQM	Amazon
Roku	iRhythm	Tesla	Exelixis	Apple
Zoom	Bruker	StoneCo	Mandiant	Microsoft
Block	Leidos	Block	Abbvie	Alphabet
Exact Sciences	Toyota	Nemetschek	Subsea	Meta
Unity	Box	AeroVironment	Bristol Myers	Alibaba
Twilio	Lockheed	DiaSorin	Jazz Pharma	Intel
Spotify	ABB	Crowdstrike	Citrix	J&J
Total weight of top 10 positions				
59%	**9%**	**12%**	**8%**	**17%**

Source: ARK Invest, S&P, ALPS, iShares, Goldman Sachs, Sparkline. As of 3/31/2022.

Moreover, the earliest innovation track records (and indices) only go back to 2013, a period dominated by a historic bull market. In order to truly understand innovation investing, we must study a deeper and more representative sample.

The Arc of Innovation

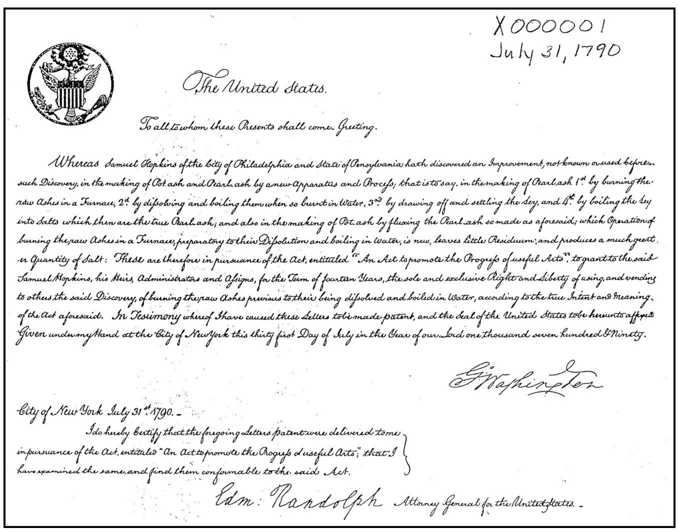
Annals of Innovation

> "There is only one thing stronger than all the armies of the world: and that is an idea whose time has come."
>
> Victor Hugo

We first introduced data from the U.S. Patent and Trademark Office (USPTO) in Investing in the Intangible Economy (Oct 2020). Patents are both a bellwether of technological trends and a measure of firm-level innovation. The USPTO was established in 1836 but maintains records dating back to 1790. The first patent was signed by George Washington!

Exhibit 3

Patent Zero



X000001
July 31, 1790

The United States.

To all to whom these Presents shall come Greeting.

Whereas Samuel Hopkins of the City of Philadelphia and State of Pensylvania hath discovered an Improvement, not known or used before such Discovery, in the making of Pot ash and Pearl ash by a new Apparatus and Process; that is to say, in the making of Pearl ash 1st. by burning the raw Ashes in a Furnace, 2d. by dissolving and boiling them when so burnt in Water, 3d. by drawing off and settling the Ley, and 4th. by boiling the Ley into Salts which then are the true Pearl ash; and also in the making of Pot ash by fluxing the Pearl ash so made as aforesaid; which Operation of burning the raw Ashes in a Furnace, preparatory to their Dissolution and boiling in Water, is new, leaves little Residuum, and produces a much greater Quantity of Salt: These are therefore in pursuance of the Act, entituled "An Act to promote the Progress of useful Arts", to grant to the said Samuel Hopkins, his Heirs, Administrators and Assigns, for the Term of fourteen Years, the sole and exclusive Right and Liberty of using, and vending to others the said Discovery, of burning the raw Ashes previous to their being dissolved and boiled in Water, according to the true Intent and Meaning of the Act aforesaid. In Testimony whereof I have caused these Letters to be made patent, and the Seal of the United States to be hereunto affixed. Given under my Hand at the City of New York this thirty first Day of July in the Year of our Lord one thousand seven hundred & Ninety.

G. Washington

City of New York July 31st. 1790.

I do hereby Certify that the foregoing Letters patent were delivered to me in pursuance of the Act, entituled "An Act to promote the Progress of useful Arts"; that I have examined the same, and find them conformable to the said Act.

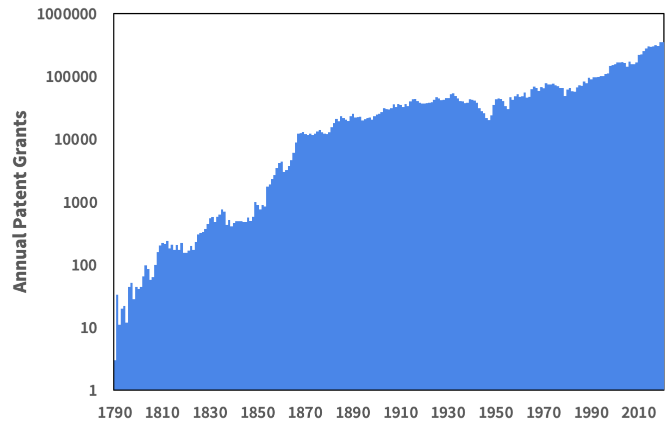
Edm: Randolph Attorney General for the United States.

Source: Wikipedia.

While patent data has limitations (e.g., many firms prefer to rely on trade secrets), its meticulously maintained history allows us to explore the arc of innovation over centuries. Since 1790, patent activity has grown exponentially. The number of annual patent grants was 41 in 1800, 24,000 in 1900, and 350,000 in 2020. Since 1800, patent activity has increased at a steady 4.2% compound annual growth rate.

Exhibit 4

Exponential Innovation

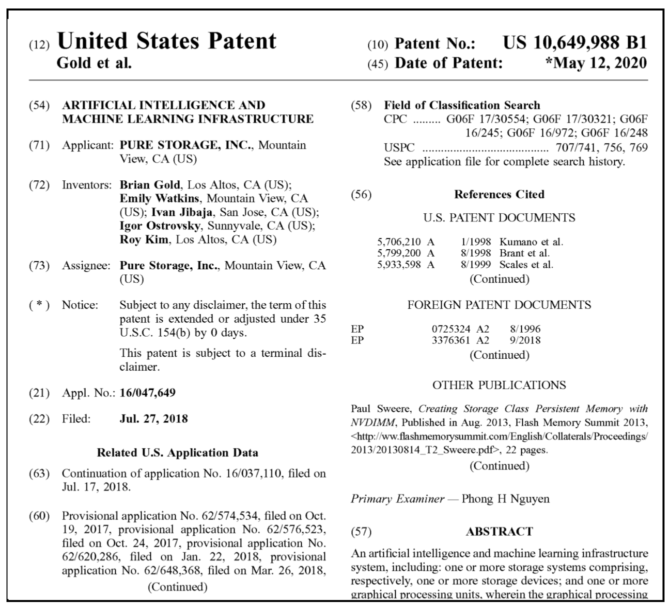


Source: USPTO, Sparkline. As of 12/31/2020.

The structure of patents is now highly standardized. Key fields include title, abstract, assignee (i.e., owner), citations, classifications, descriptions, claims and publication date.

Exhibit 5

Patent Example



(12) **United States Patent**
Gold et al.

(10) **Patent No.: US 10,649,988 B1**
(45) **Date of Patent: *May 12, 2020**

(54) **ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING INFRASTRUCTURE**

(71) Applicant: **PURE STORAGE, INC.**, Mountain View, CA (US)

(72) Inventors: **Brian Gold**, Los Altos, CA (US); **Emily Watkins**, Mountain View, CA (US); **Ivan Jibaja**, San Jose, CA (US); **Igor Ostrovsky**, Sunnyvale, CA (US); **Roy Kim**, Los Altos, CA (US)

(73) Assignee: **Pure Storage, Inc.**, Mountain View, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **16/047,649**

(22) Filed: **Jul. 27, 2018**

Related U.S. Application Data

(63) Continuation of application No. 16/037,110, filed on Jul. 17, 2018.

(60) Provisional application No. 62/574,534, filed on Oct. 19, 2017, provisional application No. 62/576,523, filed on Oct. 24, 2017, provisional application No. 62/620,286, filed on Jan. 22, 2018, provisional application No. 62/648,368, filed on Mar. 26, 2018,
(Continued)

(58) **Field of Classification Search**
CPC G06F 17/30554; G06F 17/30321; G06F 16/245; G06F 16/972; G06F 16/248
USPC 707/741, 756, 769
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,706,210	A	1/1998	Kumano et al.
5,799,200	A	8/1998	Brant et al.
5,933,598	A	8/1999	Scales et al.

(Continued)

FOREIGN PATENT DOCUMENTS

EP	0725324	A2	8/1996
EP	3376361	A2	9/2018

(Continued)

OTHER PUBLICATIONS

Paul Sweere, *Creating Storage Class Persistent Memory with NVDIMM*, Published in Aug. 2013, Flash Memory Summit 2013, <http://ww.flashmemorysummit.com/English/Collaterals/Proceedings/2013/20130814_T2_Sweere.pdf>, 22 pages.
(Continued)

Primary Examiner — Phong H Nguyen

(57) **ABSTRACT**

An artificial intelligence and machine learning infrastructure system, including: one or more storage systems comprising, respectively, one or more storage devices; and one or more graphical processing units, wherein the graphical processing

Source: Google, Sparkline.

Our first objective is to classify related patents to find the technology clusters that are gaining traction over time. In theory, we should be able to see in the historical record the fossils of past technological revolutions – railroads, steel, electricity, cars, computers, and the internet.

The USPTO categorizes patents into technology groupings using a hierarchical classification system. Unfortunately, this mapping is not very accurate. A recent study found that 78% of patents mentioning "machine learning" are missing from the Machine Learning category (G06N 20). Furthermore, maintaining this fixed structure requires frequent manual revisions, which often considerably lag new inventions.

We will instead extract topics directly from patent text using natural language processing (NLP). We use the embeddings from Investment Management in the Machine Learning Age (Jun 2019). We train our model on textual data in the title, abstract, and description. We include bigrams and trigrams to capture phrases such as "internal combustion engine."

We retrain our model each month using a rolling window. This allows us to create a point-in-time view of the evolving patent corpus, automatically categorizing new technologies as they arise in real time.

Our embeddings capture the semantic relationship between terms. The next exhibit shows examples for five technologies at different points in time. The results are intuitive. For example, given the concept "blockchain," the model finds related terms such as "distributed ledger" and "private key."

Exhibit 6
Patent Embeddings

Blockchain	Polymerase Chain Reaction	Semiconductor	Internal Combution Engine	Loom
2021	2011	1981	1931	1891
distributed ledger	pcr	transistor	compressor	weft
cryptographic	assay	integrated circuit	gas turbine	saw set
cryptocurrency	quantitation	cathode ray tube	carburetor	feed cutter
private key	amplification	amorphous silicon	centrifugal pump	metallic cartridge
token	mass spectromry	plasma etching	refrigerant	warp
ledger	genotyping	doping	impeller	shuttle loom
trusted	thermal cycling	resister	throttle	child carriage
digital certificate	sample preparation	dielectric	blower	shedding
private key	nucleic acid	crystal	thermostatic	releasing
smart contract	ligation	diode	pneumatic	stop motion

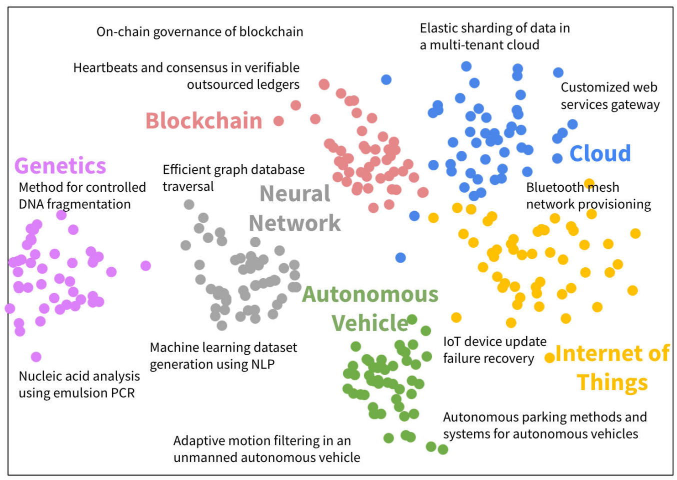
Source: USPTO, Sparkline. As of 3/31/2022.

Technological revolutions are not driven by a single, isolated breakthrough. Rather, they arise from the recombinant force of dozens of interconnected advances. Our model groups semantically-related patents into broader technological constellations. These emergent clusters often serve as the backbone of major technological revolutions.

The results are similar to those produced using citation network analysis. However, since we do not know how many citations a patent earns until many years in the future, this approach is not useful for real-time investment strategies.

Fortunately, our method allows us to link clusters of similar technologies based on their textual content. The next exhibit provides an example. We first seed the visualization with a few patents from seven different categories (e.g., US1108544 "On-chain governance of blockchain"). We then have the model find patents that are nearby in embedding space.

Exhibit 7
Technology Map



Source: USPTO, Sparkline. As of 3/31/2022.

The algorithm successfully identifies related patents. For example, it surfaces the blockchain patent "Heartbeats and consensus in verifiable outsourced ledgers." It also embeds relatedness across clusters. For example, "cloud computing" is much closer to "internet of things" than to "genetics."

Technological Revolutions

> "A technological revolution can be defined as a powerful and highly visible cluster of new technologies, products and industries, capable of bringing about an upheaval in the whole fabric of the economy and propelling a long-term upsurge of development."
>
> 🇻🇪 Carlota Perez (2002)

In Technological Revolutions and Financial Capital (2002), the economist Carlota Perez proposed an influential model connecting innovation and capital cycles. She argued that history has been shaped by five great waves of innovation. In each surge, a new cheap input (e.g., steam, oil, chips) unlocks a host of world-changing technologies, industries and infrastructure (e.g., canals, railways, fiber).

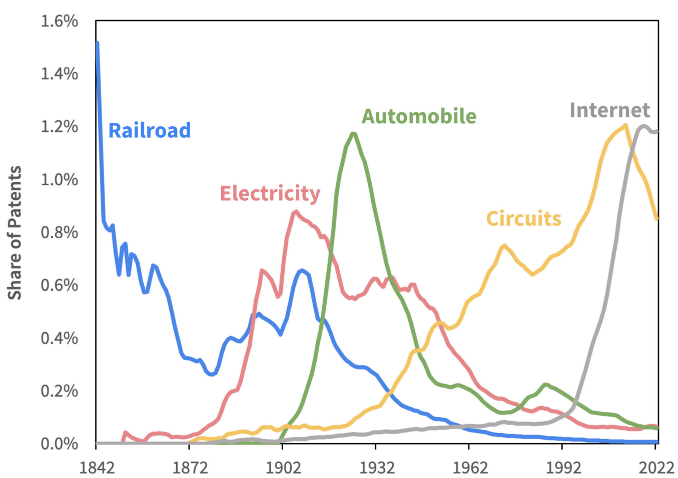
Exhibit 8
Technological Revolutions

	Technologies and Industries	Infrastructure	Big Bang	Year
The Industrial Revolution	Mechanised cotton industry; wrought iron; machinery	Canals and waterways; turnpike roads; water power	Arkwright's mill opens in Cromford (UK)	1771
Age of Steam and Railways	Steam engines and machinery; iron and coal mining; railway construction; rolling stock production; steam power	Railways; postal service; telegraph; great ports; city gas	Test of 'Rocket' steam engine for Liverpool-Manchester railway (UK)	1829
Age of Steel, Electricity and Heavy Engineering	Cheap steel; steam engine for steel ships; heavy chemistry and civil engineering; electrical equipment; canned and bottled food; paper and packaging	Worldwide shipping in steel steamships; transcontinental railways; worldwide telegraph; telephone; electrical networks	Carnegie Bessemer steel plant opens in Pittsburgh, PA (US)	1875
Age of Oil, the Automobile and Mass Production	Mass-produced automobiles; cheap oil and fuels; petrochemicals; internal combustion engine; home electrical appliances; refrigerated and frozen foods	Networks of roads, highways, ports and airports; networks of oil ducts; universal electricity; worldwide analog telecom	First Model-T produced by Ford plant in Detroit, MI (US)	1908
Age of Information and Tele-communications	Cheap microelectronics; computers and software; telecommunications; control instruments; computer-aided biotechnology and new materials	Worldwide digital telecom; internet and email; electricity networks; high-speed physical transport by land, air and water	Intel microprocessor announced in Santa Clara, CA (US)	1971

Source: Perez, Sparkline.

Amazingly, we see these technological revolutions play out in the patent data. The next exhibit uses our NLP model to cluster patents into broad categories (e.g., the automobile) and plots the share of patents in each.

Exhibit 9
Technological Revolutions



Source: USPTO, Sparkline. As of 12/31/2021.

Railroads were the technological marvel of the early 1800s, with a secondary peak at the turn of the 20th century as railways were retrofitted with steel. The Age of Steel also coincided with the mass electrification of the economy. In the 1920s, the automobile revolutionized our infrastructure. In the late-1950s, circuitry ushered in the Age of Information, which accelerated further with the internet in the late-1990s.

We can drill down into the individual technologies powering each revolution (e.g., "internal combustion engine" instead of "automobile"). Rather than absolute levels, we focus on the rate of change in order to find "trending technologies." Exhibit 10 shows top trending technologies by decade.

Exhibit 10
Greatest Hits by Decade

1850s	harvester	loom	plow	sewing machine	steam
1860s	cultivator	stove	lamp	corn planter	plow
1870s	car coupling	furnace	fire escape	spring	stove
1880s	railway car	electric	car coupling	furnace	telephone
1890s	electric railway	bicycle	electric	gas engine	type writer
1900s	internal combustion	acetylene gas	rotary engine	turbine	loom
1910s	airplane	internal combustion	dynamo electric	phonograph	safety razor
1920s	washing machine	typewriter	carburetor	spark plug	radio
1930s	airplane	air conditioner	refrigerator	circuit breaker	amplifier
1940s	polymer	electronic	television	plastic	catalyst
1950s	automobile	television	conveyor	refrigerator	television
1960s	transistor	nuclear reactor	computer	polyester	radar
1970s	laser	integrated circuit	sensor	optical fiber	solid state
1980s	magnetic recording	optical fiber	semiconductor	solar	digital camera
1990s	superconducting	semiconductor	fabrication	database	air bag
2000s	wireless network	semiconductor	mobile communication	memory	server
2010s	portable electronic	touch	OLED	cloud	lithography

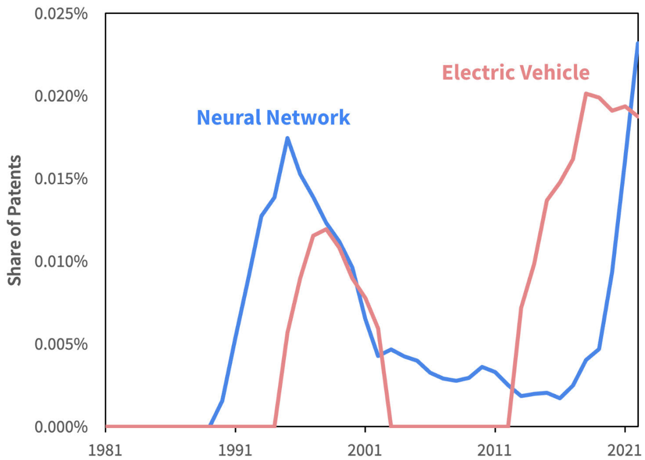
Source: USPTO, Sparkline. As of 12/31/2019.

We can now see the underlying technologies driving each great wave. For instance, the Information Age was unlocked by several key innovations spread across decades, such as the transistor, computer, integrated circuit, magnetic memory, and optical fiber.

Technologies appear to trend in wavelike patterns. A few initial breakthroughs are followed by a flurry of innovations building on these foundations. This activity continues to grow until reaching a peak and then subsiding as the next technological wave takes its place at the crest of innovation.

Of course, innovation also has false starts. Electric vehicles were actually the best-selling cars of the early 1900s before being supplanted by the internal combustion engine. They reemerged in the 1990s with the GM EV1, but interest again faded until Tesla. Neural networks were similarly ahead of their time in the 1990s. In both cases, adoption was gated by the cost and quality of key components (i.e., batteries and chips). Only after decades of progress (e.g., Moore's Law) are they finally now realizing their potential.

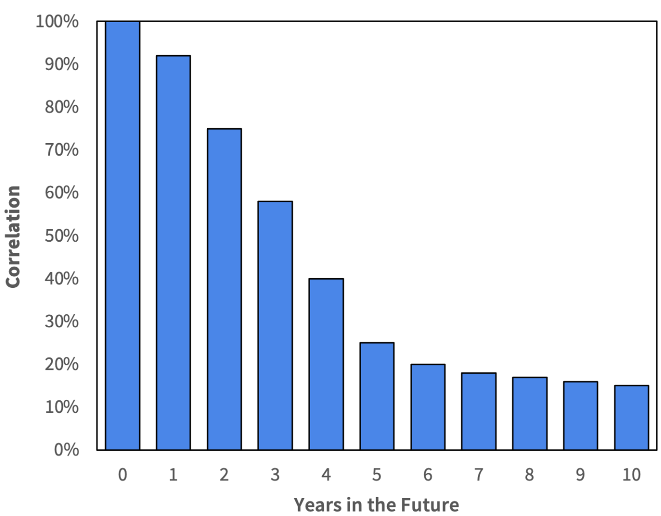
Exhibit 11
False Starts 🏃



Source: USPTO, Sparkline. As of 12/31/2021.

Fortunately, false starts are more the exception than the rule. The arc of innovation bends upward. Technology tends to trend rather than mean-revert. We can demonstrate this by calculating the serial correlation of patent growth rates over various lags.

Exhibit 12
Do Technologies Trend?



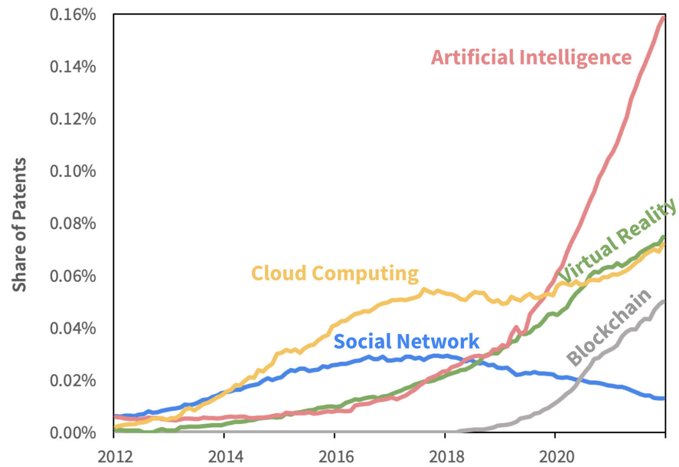
Source: USPTO, Sparkline. As of 12/31/2021.

Technologies that have experienced rapid growth tend to continue to experience above-average growth the next year. This persistence seems to last around 3-5 years. Of course, this is just an average – some trends last longer and others shorter. The key is that we can predict the future path of innovation simply by extrapolating past trends.

Modern Disruption

So what does this all mean for the modern investor? Let's zoom in on the big trends of the past decade.

Exhibit 13
Current Technological Trends 🌊



Source: USPTO, Sparkline. As of 12/31/2021.

Cloud computing and social networking were the first of our modern technologies to be widely adopted. The cloud is still growing, but social networking has lost momentum. In contrast, A.I., virtual reality, and blockchain only started trending more recently but are enjoying explosive growth.

Exhibit 14
Trending Technologies Today

	# Patents (2019-2021)	Growth Rate (4-year smoothed)
Blockchain	909	397%
Artificial Intelligence	3836	196%
Quantum Computing	499	102%
3D Printing	1221	93%
Internet of Things	910	92%
Virtual Reality	1808	67%
Autonomous Vehicles	2350	58%
Robotics	2554	33%
Cloud Computing	1328	26%

Source: USPTO, Sparkline. As of 12/31/2022.

Blockchain is the hottest technology with 400% growth over the past four years. There is also interesting work being done in A.I., quantum computing, 3D printing, internet of things, virtual reality, autonomous vehicles, and robotics. Cloud computing has become a "mature" technology but is still enjoying robust 26% annual growth.

Our analysis of two centuries of patent data highlights the triumphant march of human knowledge. As investors, we want to benefit from this engine of growth. However, since we cannot invest directly in intellectual property, we will instead buy the companies that create and own it.

Investing in Innovation

The Ladder of Innovation

> "The difficulty lies not so much in developing new ideas as in escaping from old ones."
>
> 🇬🇧 John Maynard Keynes

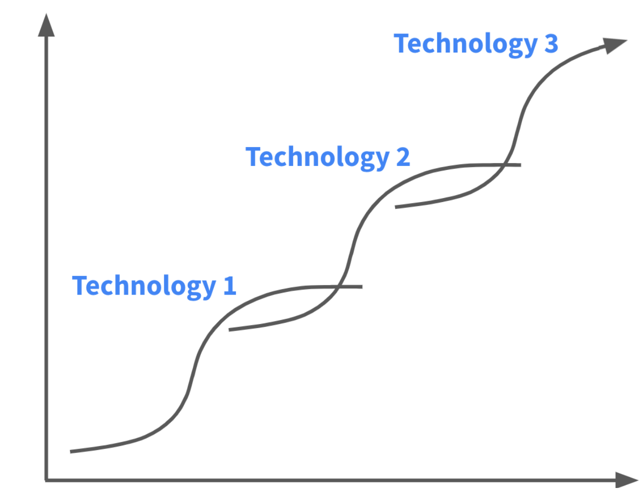
Technological adoption tends to follow "S-curves," in which a few early adopters pave the way for the mass market, followed by a handful of laggards.

Large firms often build a diversified portfolio of IP at varying stages of the S-curve. For example, Alphabet is the holding company for an established search business, a growing cloud segment, and some "moonshots" (e.g., Waymo).

In contrast, we target only early- to mid-stage technologies. As the technologies we hold mature, we rotate into the next set of trending technologies. The goal is to always be owning technologies at the steepest parts of their S-curves. This can be visualized as climbing a ladder of escalating S-curves.

Exhibit 15

Innovation Is a Ladder 🪜



Source: Sparkline.

Unfortunately, it is very rare for a single firm to deftly surf multiple technological waves. Apple is a singular firm as an innovator in technologies ranging from personal computers to smartphones. Most firms become victims of their success and fail to make it to the next wave (e.g., IBM, Intel).

In order to avoid the risk of the disruptors being disrupted, we will not buy and hold a single firm but instead own a portfolio of firms and recycle it as technologies evolve. The rest of this paper will focus on publicly traded stocks, whose liquidity allows us to execute this strategy.

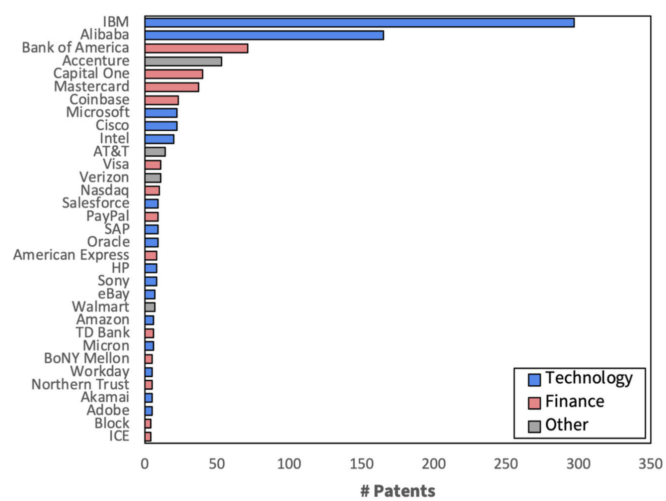
Searching for Innovators

How can we determine which companies have exposure to our target technologies? In other papers, we have looked for clues in various textual sources (e.g., earning calls, 10-Ks, employee resumes, patents). In this article, however, we will use patent data exclusively due to its deep history.

Conveniently, each patent has an assignee (e.g., "Pure Storage, Inc."). We use a simple NLP algorithm to link patents to their corporate owners. We used this method in Value Investor's Guide to Web3 (Jan 2022) to identify the top holders of blockchain patents.

Exhibit 16

Top Blockchain Patent Holders



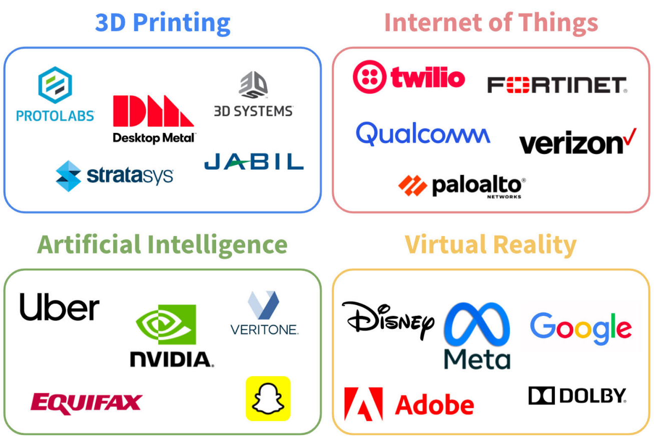
Source: USPTO, Sparkline. Includes public companies with more than 10 crypto job posts in the past year. As of 12/31/2021.

Since large firms like IBM tend to have more patents in general, we normalize by the total number of patents. This helps home in on firms like Coinbase, for which blockchain comprises a large share of a small patent portfolio.

We can apply this process to any arbitrary technology. In this paper, we will focus on a rotating list of the top ten trending technologies at each point in time. This recycling allows us to move into new technologies as existing ones mature (i.e., target a "constant maturity").

Our investment universe is the U.S. stock market. The next exhibit shows sample stocks for a few sub-portfolios today.

Exhibit 17
Sample Sub-Portfolios



Source: USPTO, Sparkline. As of 3/31/2022.

We average the weights across the ten sleeves to create a composite innovative stock portfolio. There are currently 212 stocks in the portfolio, which is equally-weighted. The next exhibit shows some representative holdings.

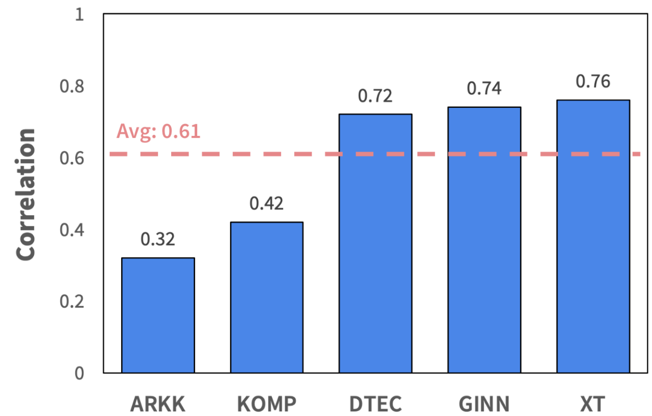
Exhibit 18
Sample Innovation Portfolio

Palantir	Data Analytics	Tesla	Electric Vehicles
Splunk	Data Analytics	General Motors	Automobiles
Pure Storage	Data Storage	Caterpillar	Industrial Machinery
Airbnb	Homesharing	Palo Alto Networks	Cybersecurity
IQVIA	Medtech	ServiceNow	Cloud Computing
Luminar	LIDAR	Meta Platforms	Social Networking
Desktop Metal	3D Printing	Rockwell	Industrial Automation
Proto Labs	3D Printing	Open Text	Data Management
Lam Research	Semiconductor Eq.	Nvidia	GPUs
Uber	Ridesharing	ANSYS	Engineer Simulation

Source: USPTO, Sparkline. As of 3/31/2022.

We validate this portfolio by comparing it to the innovation ETFs from Exhibit 2. The next exhibit shows its correlation to each ETF. The correlation is on average 61% but is dragged down by two outliers. KOMP is unique as it focuses on high price momentum stocks. Meanwhile, ARKK concentrates in a handful of high-conviction names (e.g., 10% in Tesla).

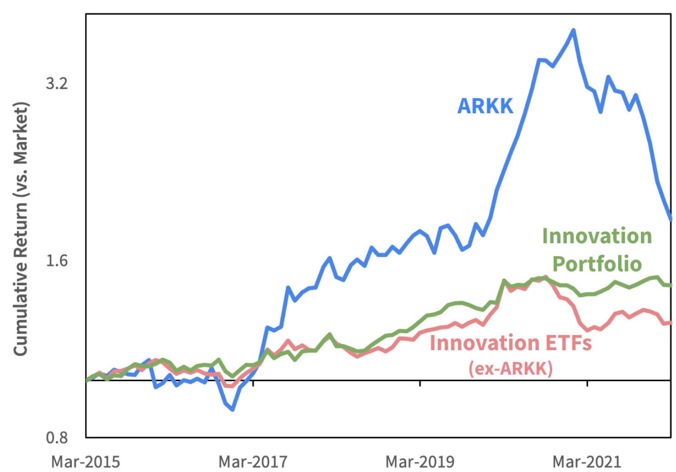
Exhibit 19
Innovation ETF Correlations



Source: ARK Invest, S&P, ALPS, iShares, Goldman Sachs, USPTO, Sparkline. Start date is the inception for each ETF. As of 3/31/2022.

The next exhibit shows each strategy's return relative to the market, defined as an equal-weighted index of U.S. stocks. We split out ARKK, since its high active share introduces lots of idiosyncratic risk (we will return to this point later). Fortunately, an index of the remaining four ETFs maps well onto our innovation portfolio (i.e., 81% correlation).

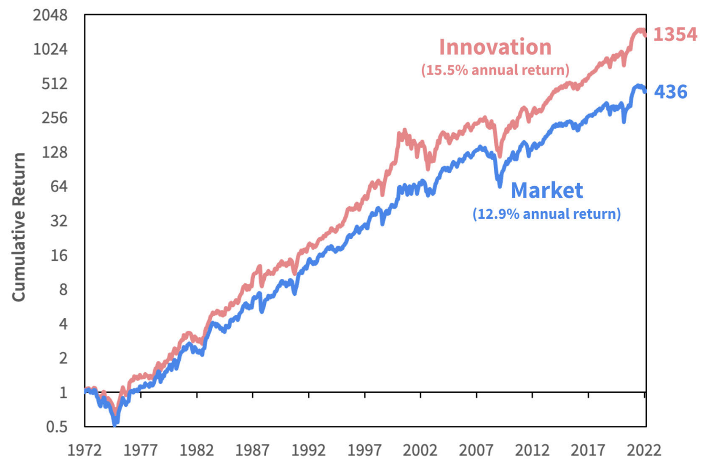
Exhibit 20
ETF Comparison



Source: ARK Invest, S&P, ALPS, iShares, Goldman Sachs, USPTO, Sparkline. Start date is the inception for each ETF. Market is an equal-weighted index of U.S. equities. Innovation holds stocks with high shares of innovative patents. We rebalance monthly and exclude transaction and financing costs. From 3/31/2015 to 3/31/2022. See important backtest disclosure.

Now that we are comfortable that our model captures the "essence of innovation," we extend the strategy back in time. We begin our backtest in 1971, the start of the Information Age according to Perez.

Exhibit 21
ROI: Return on Innovation

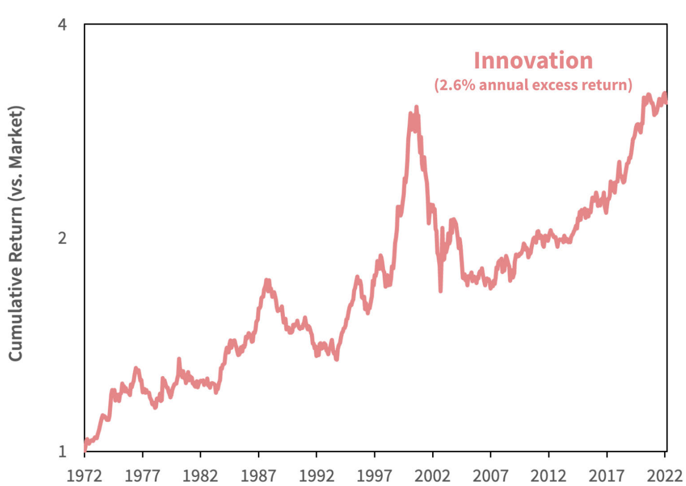


Source: S&P, USPTO, Sparkline. Market is an equal-weighted index of U.S. equities. Innovation holds stocks with high shares of innovative patents. We rebalance monthly and exclude transaction and financing costs. From 12/31/1971 to 3/31/2022. See important backtest disclosure.

Over the past half century, innovation outperformed the market by +2.6% per year. This compounds to a large 3X advantage. In line with the popular narrative, we also find that innovation stocks had greater volatility than the market.

Examining returns relative to the market reveals a consistent uptrend punctuated by a big dislocation in the dot-com bubble. This is not at all surprising, as it is well known that innovative internet stocks got very overvalued in the bubble.

Exhibit 22
Relative Performance



Source: S&P, USPTO, Sparkline. Market is an equal-weighted index of U.S. equities. Innovation holds stocks with high shares of innovative patents. We rebalance monthly and exclude transaction and financing costs. From 12/31/1971 to 3/31/2022. See important backtest disclosure.

Explaining the Premium

We love innovation. Technology lifted the human species from millennia of subsistence to the dominant form of life on earth. And it is the principal driver of economic growth on a long timescale. That being said, investors in innovation do not have a god-given right to excess returns.

In an efficient market, prices accurately incorporate future growth expectations. Even if a disruptive company does ultimately reshape society, its investors will not realize excess returns if this outcome was already priced in.

However, contrary to efficient market theory, innovation stocks *have* outperformed the market. Why might this be? We have two explanations:

1. **Market Inefficiency:** The simplest explanation is that the market is not efficient. Innovation investing requires both analytical and behavioral edges.

   Investors rely heavily on financial statement metrics like net income and sales. In Intangible Value (Jun 2021), we argued that accounting is not well suited to the modern intangible economy. Not only is R&D expensed instead of capitalized, but there is basically no disclosure of the *quality* of R&D. While this information can be gleaned from alternative data (e.g., patents), NLP and other big data tools are required to harvest it.

   Moreover, investors are widely known to be afflicted with short-term bias. For both institutional and behavioral reasons, Wall Street is obsessed with quarterly financial results. This may lead investors to overlook the profound but gradual impact of secular technological shifts. The boiled frog applies a big discount to intangible, long-duration investments.

2. **Risk Premium:** Alternatively, the observed returns could be compensation for unique risks associated with deep technological research.

   While the success of innovations like the internet or cloud computing might seem preordained in hindsight, there were many doubters at the time. Today, Facebook is investing a princely $10 billion per year in the metaverse. The stock's relatively low valuation reflects in part investors' discomfort with Zuckerberg betting the firm on a science fiction novel.

Research in disruptive technology can be boom-or-bust. A broad portfolio of disruptive stocks can help mitigate firm-specific risk (e.g., correctly calling the internet but buying the wrong stocks), but technological risk itself is harder to diversify.

We won't come down conclusively on either side. It's most likely that both factors partially explain the return. Investing in innovation is both challenging and risky. But that's why you get paid!

Innovation as an Asset Class

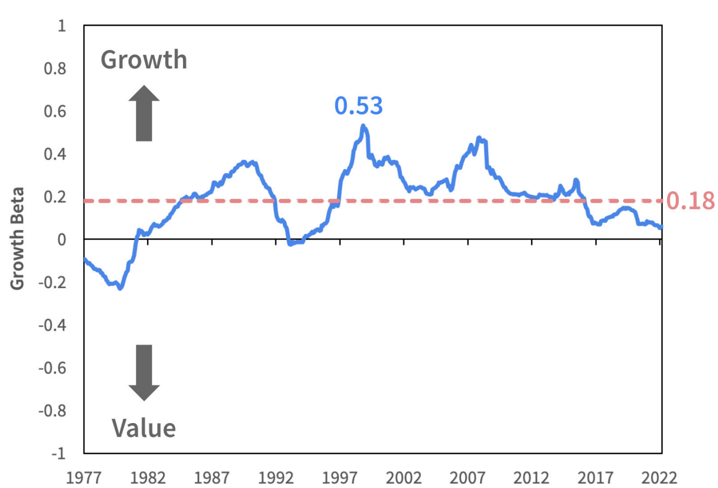
Growth, Tech, and Innovation

One popular way to classify investments is the so-called "style box" framework, which divides stocks into a 2x2 grid on the dimensions of value v. growth and small v. large. Recently, some investors have argued that innovation be considered its own asset class outside of the reductive style box model.

In order for innovation to be its own asset class, however, it must be distinct from existing factors. Since innovation is most commonly considered a form of growth investing (e.g., "deep growth"), we'll start by examining this relationship. Below is innovation's rolling beta to the growth factor.

Exhibit 23

Innovation's Growth Beta



Source: S&P, Ken French, USPTO, Sparkline. Innovation holds stocks with high shares of innovative patents. Rolling 1-year beta of innovation to inverse Fama-French value-growth factor. As of 2/28/2022.

Innovation's growth beta has averaged a mere 0.18. However, it has fluctuated greatly over time. In the dot-com bubble, innovative internet companies became very overvalued and started trading like speculative growth stocks. This led to the boom and bust pattern we saw in the backtest earlier.

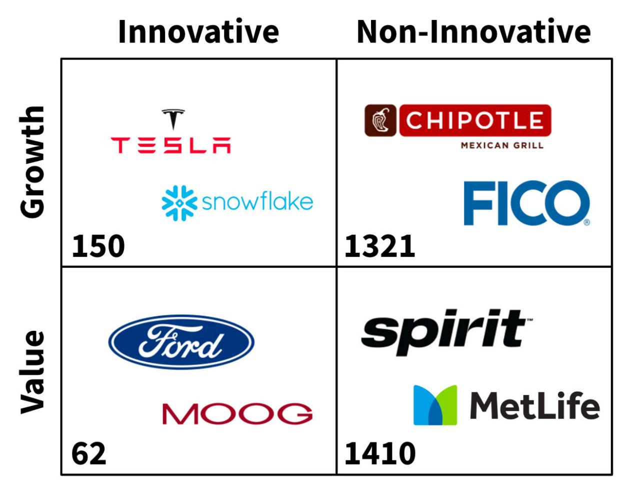
Interestingly, innovation's growth beta has been declining over the past decade. The innovative companies of today are not the same as those of the internet boom. The dot-coms were mostly speculative and unprofitable. In contrast, cutting-edge research now tends to be done by the largest, most profitable firms (e.g., Google, Facebook, Nvidia).

Innovation and growth are distinct concepts. Value and growth are merely synonyms for cheap and expensive. Innovation, like any other group of stocks, can be cheap or expensive depending on the fashion of the day. While the market tends to assign higher multiples to innovative firms on average, the correlation is weak and unstable.

To hammer this point home, the next exhibit segments U.S. stocks along the dimensions of innovation and growth.

Exhibit 24

Innovation ≠ Growth

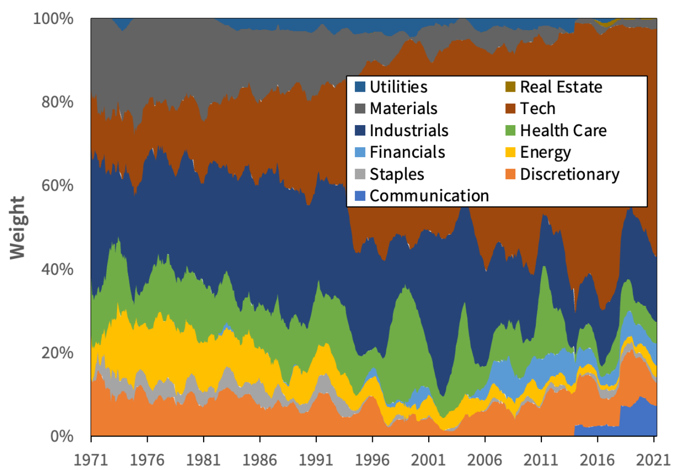


Source: S&P, USPTO, Sparkline. Value vs. growth is defined based on a median threshold on price-to-book value. As of 3/31/2022.

Almost half the universe falls in the off-diagonals. Of our innovative stocks, 30% are considered value. Ford is a value stock with exposure to A.I., robotics, and electric vehicles. Of our growth stocks, 90% are considered non-innovative. High growth and valuations can be achieved from assets other than intellectual property (e.g., brand, regulatory capture).

Finally, we'll test the relationship between innovation and technology. The next exhibit shows the sector composition of the innovation portfolio. Technology has become a larger share of the portfolio over time, but this is also the case for the market in general (e.g., S&P 500 is 27% tech).

Exhibit 25
Sector Composition



Source: S&P, Sparkline. Data before 2010 uses 2010 GICS Structure. As of 3/31/2022.

While many disruptive companies are tech stocks, there are disruptive companies in all sectors. Conversely, only 29% of tech companies are disruptive – many tech companies today are legacy firms. See Value Investing is Short Tech Disruption (Aug 2020) for further exposition on this point.

The Risks of Innovation

In 1993, Fama and French published an influential paper showing that the **market beta**, **size**, and **valuation** of a stock explain its price movements. The notion that size and value are "risk factors" is behind the rise of style boxes and has in many other ways shaped the investment industry.

Over time, researchers expanded the model to include other factors. In 2014, Fama and French proposed adding factors for **profitability** and **investment** (i.e., asset growth rate). Many researchers have also argued to include **momentum** (i.e., 1-year trailing return) in the mix.

While we have so far satisfied ourselves that innovation is distinct from growth and tech, in order to pass the more rigorous scrutiny of quantitative finance nerds, we'll run our result through the Fama-French model (+momentum). 🤓

The following exhibit shows the results of a regression of our innovation portfolio on the Fama-French factors.

Exhibit 26
Fama-French Betas

Factor	Definition	Beta	Standard Error	t-stat
Market	Market - Risk-Free Rate	1.15	0.02	47.9*
Size	Small - Large Cap	0.36	0.05	7.5*
Value	Value - Growth	-0.24	0.04	-5.5*
Profitability	Quality - Junk	-0.37	0.08	-4.7*
Investment	Conservative - Aggressive	0.01	0.08	0.1
Momentum	High - Low Momentum	-0.02	0.05	-0.4
			R-sqr	0.91

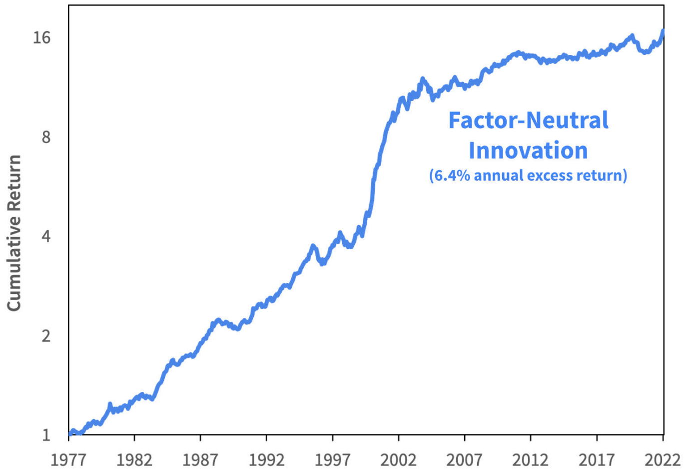
Source: S&P, Ken French, USPTO, Sparkline. Robust standard errors. 1/31/1972 to 2/28/2022. *Significant at the 1% level.

We learn a lot about the properties of innovative stocks. First, our innovation portfolio has a market beta greater than one – these stocks are high-octane. Second, it has a 0.36 beta to size, as it tends to hold smaller stocks. Third, it has a -0.24 beta to value. This aligns with our earlier finding that innovation stocks tend to be more richly valued. Fourth, it has a -0.37 beta to profitability. Innovative firms are less profitable than average, as they invest heavily in the future (and this investment is expensed, not capitalized).

Fifth, the momentum factor is not significant. This dispels the popular myth that trend-following investors chasing hot stories tend to favor innovative companies. It also confirms patent momentum is not subsumed by price momentum. Finally, including industry dummies leads to a 0.4 beta to tech, consistent with our earlier finding (results not shown).

The next exhibit re-runs our backtest after neutralizing the Fama-French factors using trailing 5-year rolling regressions. We isolate innovation's pure alpha (i.e., residual return) by removing the return contribution arising from its exposures to traditional factors.

Exhibit 27
Factor-Neutral Innovation



Source: S&P, USPTO, Ken French, Sparkline. Innovation holds stocks with high shares of innovative patents. We use rolling 5-year regressions on the FF5 model (+momentum) and plot the cumulative residuals. We rebalance monthly and exclude transaction and financing costs. From 12/31/1976 to 2/28/2022. See important backtest disclosure.

Innovation's performance is much smoother after removing its factor exposures. Neutralizing its growth exposure helps avoid getting whipsawed in the dot-com bubble. This allows the strategy to isolate pure innovation without picking up the occasional bet on expensive (or cheap) stocks.

Long-term returns are also higher. This is in large part due to the effect of removing innovation's negative exposure to factors that have historically enjoyed positive long-term returns, such as value and profitability.

Innovation as an asset class is a reasonable idea. Not only is innovation distinct from traditional style exposures (e.g., size and value), but pure innovation has also been rewarded with a long-term premium.

Don't Blame Innovation

The recent drawdown of many popular innovation funds has ignited a swirling debate over innovation stocks. Skeptics proclaim the "death of innovation" while proponents rejoice that "innovation is on sale."

However, we believe that this debate is misplaced. It turns out the recent drawdown was not in innovation, but rather in more mundane factors such as growth and junk. We will show this using ARKK as an example (not to pick on them but because they have the longest track record 🙏).

We start by decomposing ARKK's factor exposures using the regression model from the last section. However, in addition to the Fama-French factors, we include our innovation factor (expressed in long-short form relative to the market).

Exhibit 28
ARKK Betas

Factor	Definition	Beta	Standard Error	t-stat
Market	Market - Risk-Free Rate	1.14	0.03	44.2*
Size	Small - Large Cap	0.47	0.05	9.0*
Value	Value - Growth	-0.61	0.04	-13.6*
Profitability	Quality - Junk	-0.96	0.06	-16.4*
Investment	Conservative - Aggressive	-0.45	0.08	-5.7*
Momentum	High - Low Momentum	-0.01	0.03	-0.1
Innovation	Innovation - Market	0.79	0.06	13.1*
			R-sqr	0.80

Source: ARK Invest, S&P, Ken French, USPTO, Sparkline. Robust standard errors. Daily returns. 11/3/2014 to 2/28/2022. *Significant at the 1% level.

ARKK has significant style exposure. First, its 0.47 beta to size indicates a focus on smaller-cap stocks. Second, its -0.61 beta to value indicates a strong bias toward expensive growth stocks. Third, its -0.96 beta to profitability indicates a very strong willingness to hold unprofitable firms. Fourth, its -0.45 beta to investment indicates a tendency to hold firms with growing balance sheets.

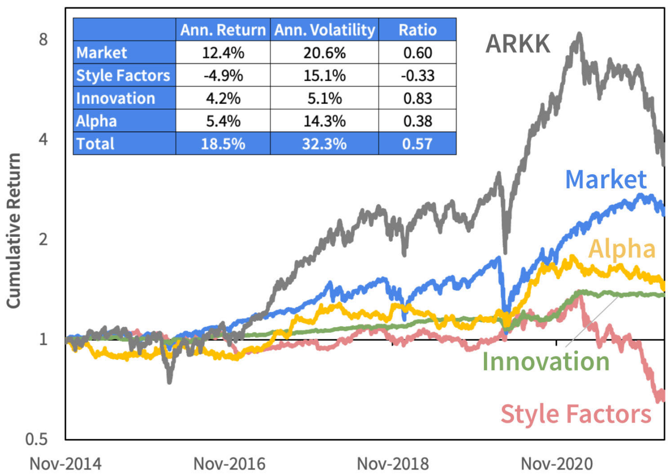
Most importantly, ARKK has a high 0.79 beta to our innovation factor. True to its brand, the fund is investing heavily in disruptive innovation.

Using this model, we can decompose ARKK's returns into four buckets. We combine the style factors into a single bucket for simplicity. Alpha is the residual return that cannot be explained by the factor model.

ARKK = Market + Style Factors + Innovation + Alpha

The next exhibit shows the historical return attribution. We run the model using trailing 1-year rolling periods to remove hindsight bias.

Exhibit 29
ARKK Attribution

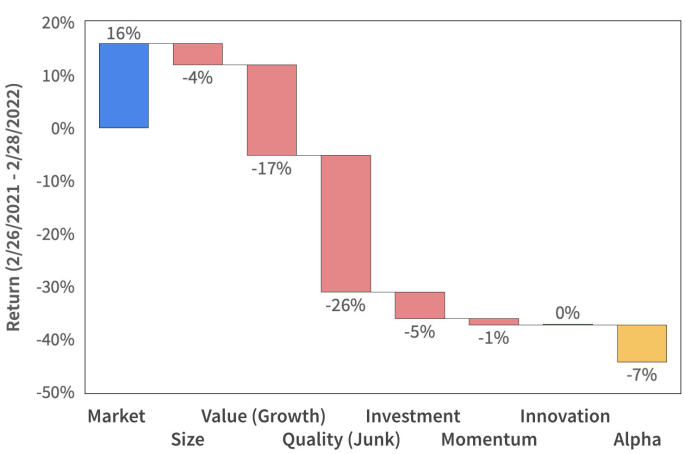


Source: ARK Invest, S&P, Ken French, USPTO, Sparkline. Innovation holds stocks with high shares of innovative patents. We use rolling 1-year regressions on the FF5 model (+momentum). We rebalance daily and exclude transaction and financing costs. From 11/3/2014 to 2/28/2022. See important backtest disclosure.

Since inception, ARKK investors have profited from the fund's market beta (12.4% annualized) and prescient bet on innovation (4.2% annualized). ARKK also justified its active risk and fees with significant alpha (5.4% annualized). Until last year, traditional factors also added value. But style exposure has been disastrous the past year!

The waterfall chart below shows the contribution of each factor to last year's return. ARKK earned +16% from its market beta, but this was more than offset by a whopping -43% loss from its junk and growth exposure!

Exhibit 30
ARKK 1-Year Decomposition



Source: ARK Invest, S&P, Ken French, USPTO, Sparkline. Innovation holds stocks with high shares of innovative patents. We use rolling 1-year regressions on the FF5 model (+momentum). We rebalance daily and exclude transaction and financing costs. From 2/26/2021 to 2/28/2022. See important backtest disclosure.

The popular storyline is that ARKK's recent losses point to the failure of innovation investing. However, this is not true. Innovation has weathered the past year just fine. But ARKK doesn't only invest in innovation. It also has a large bet on expensive, speculative firms, which have had a tough year.

Value and Innovation

Innovation Crashes

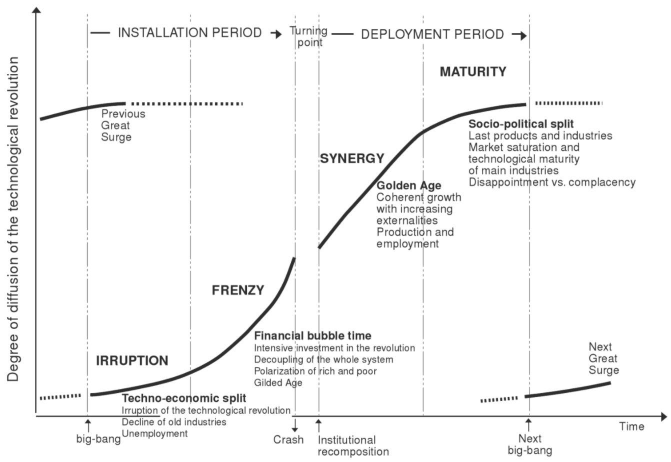
> "At the root of all financial bubbles is a good idea carried to excess."
>
> 🇺🇸 Seth Klarman

Innovation investing has generated a long-term return premium. However, we have seen that it is also prone to buy expensive growth stocks. This was most salient in the dot-com bubble, when it rotated into overpriced internet stocks and suffered massive losses in the subsequent crash.

The link between innovation and speculation is rooted in human nature. Throughout history, the siren of progress has seduced investors. Technological revolutions have almost always been accompanied by a hype cycle consisting of overinflated expectations, bubbles, and crashes.

Carlota Perez describes this Schumpeterian capital cycle in great detail. In her model, the irruption of a new technology attracts a frenzy of speculative capital, inflating a bubble and social unrest. These imbalances ultimately unwind in an epic crash. This forces institutions to be reconfigured, which leads to a golden age. Eventually, innovation stagnates until a new disruptive wave begins and the cycle repeats.

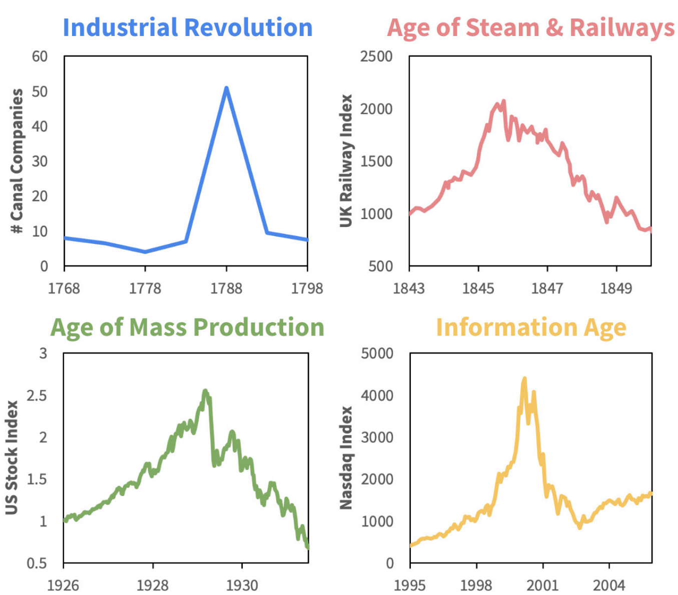
Exhibit 31
Cycle of Technological Revolutions



Source: Perez (2002).

While financial bubbles can occur without a technological narrative (e.g., tulips, silver, yen), almost all technological revolutions have been accompanied by bubbles. One might even argue that speculative capital provides the essential kindling to light the spark of moonshot innovation.

Exhibit 32
Bubble Bloodbath



Source: Perez (2009), Campbell and Turner (2010), Ken French, Nasdaq, Sparkline.

Unfortunately, it is nearly impossible to know in real time if one is in the midst of a bubble. Even if one possessed this knowledge, it alone is insufficient to produce profits due to the challenges of path dependency and timing.

Investors who avoid innovation due to bubble anxiety leave a lot of money on the table. But aping into overpriced story stocks is also not right. We'll show how a value lens can help investors participate in innovation while avoiding many of its speculative pitfalls.

Disruption at a Reasonable Price

> "It's far better to buy a wonderful company at a fair price than a fair company at a wonderful price."
>
> 🇺🇸 Warren Buffett

In Value Investor's Guide to Web3 (Jan 2022), we discussed the challenges of the hype cycle in cryptocurrencies. We formed fundamental value metrics from blockchain, GitHub and social data to help investors navigate the extreme hype and volatility in crypto markets.

Let's see if this approach works more generally across all emerging technologies, not just blockchain. We will remain fully invested in innovation stocks but will seek to avoid the most expensive stocks at each point in time. Let's call this strategy DARP – "disruption at a reasonable price."

We start by building a metric for the price of innovation – an "innovation yield." Value investors often compare price to earnings, book value, or dividends. We do the same but use "innovative patents." This produces a measure for the amount of innovation obtained per billion dollars invested.

Exhibit 33
Innovation Yield Example

	# Innovative Patents	Market Cap	Innovative Patent Yield
Alibaba Group	106	$292.4B	36%
Ford	12	$70.0B	17%
Meta	86	$605.3B	14%
VISA	20	$465.3B	4%

Source: S&P, USPTO, Sparkline. # innovative patents calculated over a two-year window. As of 3/31/2022.

Scaling innovation by price imbues our strategy with a value tilt. This manifests in more attractive value characteristics. Innovative patent yield (i.e., # innovative patents divided by market cap) surges from 13% to 49%. Traditional valuation ratios also improve. Finally, its historical average growth beta declines from 0.18 to 0.08.

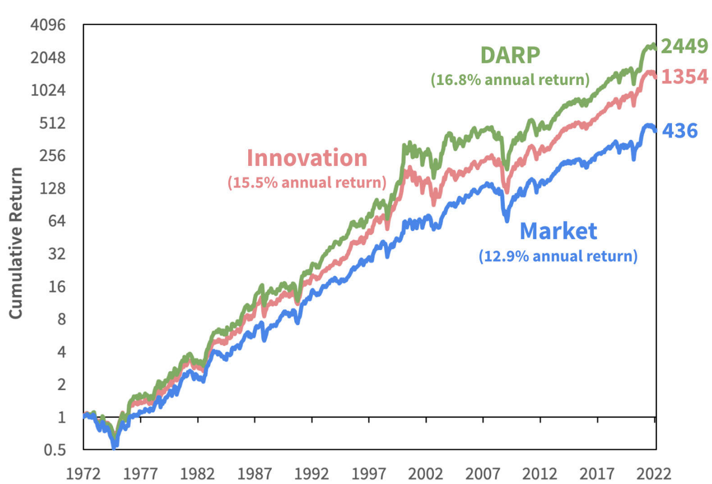
Exhibit 34
Value Characteristics

Factor	DARP	Basic Innovation	Market
Innovative Patent Yield	49.0%	13.0%	6.6%
Patent Yield	7.7%	2.1%	1.1%
Earnings Yield	5.0%	4.0%	4.1%
Free Cash Flow Yield	6.0%	3.1%	3.3%
Sales Yield	47.6%	28.6%	38.5%
Book Yield	28.6%	21.7%	27.8%

Source: S&P, USPTO, Sparkline. Calculations are position-weighted averages. Market is an equal-weighted index of U.S. equities. Innovation holds stocks with high shares of innovative patents. DARP removes the most expensive stocks on innovation yield. Characteristics refer to the underlying stocks and do not predict future performance. *Innovative patents and patents are scaled by billions and millions, respectively. As of 3/31/2022.

These improved value characteristics translate into better returns. DARP outperforms the basic innovation strategy by +1.3% per year with lower volatility. Importantly, DARP greatly mitigates the pain of the dot-com crash, returning to its highs two years sooner. It also fully eliminates the underperformance of the past year.

Exhibit 35
DARP Backtest

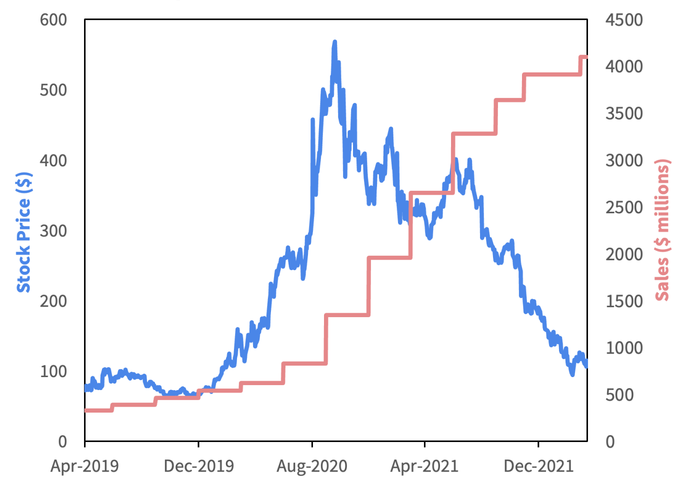


Source: S&P, USPTO, Sparkline. Market is an equal-weighted index of U.S. equities. Innovation holds stocks with high shares of innovative patents. DARP removes the most expensive stocks on innovation yield. We rebalance monthly and exclude transaction and financing costs. From 12/31/1971 to 3/31/2022. See important backtest disclosure.

Why was risk reduced? For very expensive companies, the biggest risk factor is often not fundamental decline but a repricing of overinflated expectations. DARP seeks to avoid stocks that are heavily exposed to this "valuation risk."

Zoom provides an illustrative example. A pandemic winner, it was hit as the macro regime shifted toward reopening. Zoom's stock price retreated to pre-pandemic levels, even though sales had climbed 7.5X over this period. Its -80% drop was due not to fundamental decline but to multiple compression (i.e., price-to-sales ratio falling from 110 to 8.5).

Exhibit 36
Zoom Roundtrip



Source: S&P, SEC, Sparkline. As of 4/14/2022.

DARP did not hold Zoom. Despite the disruptive potential of remote work, DARP could not get comfortable with Zoom's astronomical valuation. It was a good miss. We want to own disruption at a reasonable price, not disruption at any price.

Investing in innovation does not require YOLO'ing into speculative, hypergrowth stocks. Prudent value investors can also invest in innovation by following a DARP approach.

Beyond Innovation

In Intangible Value (Jun 2021), we introduced a modernized value investing framework. The economy is increasingly intangible. Thus, rather than measure value with traditional metrics like book value, we use intangible assets. We believe there are four main pillars of intangible value: brand equity, human capital, network effects, and intellectual property.

Exhibit 37
Four Intangible Moats

Brand Equity	Strong brand recognition and customer loyalty
Human Capital	Highly skilled, aligned, and motivated workforce
Intellectual Property	Proprietary knowledge, patents, technology, and data
Network Effects	Engaged ecosystem of external producers and consumers

Source: Sparkline.

Our intellectual property pillar consists of proprietary knowledge held in patents, trade secrets, technology and data. In other words, patents are one but not the only form of intellectual property. And intellectual property in turn is only one of four intangible pillars.

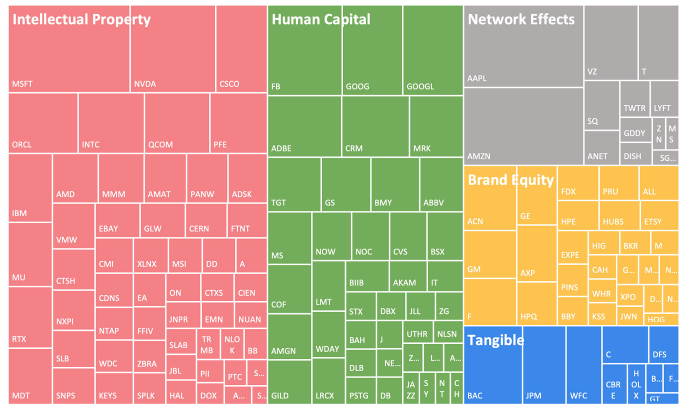
In other words, intellectual property should not be viewed in isolation. Genius alone is rarely enough to drive commercial success. History is littered with brilliant ideas that failed due to an inability to raise capital, attract talent, build brand, or bootstrap network effects.

In addition, we have seen that innovation stocks can get overvalued in speculative frenzies. While avoiding the most overpriced stocks via DARP can help, one can do even better. By viewing innovation as just one of four intangible assets, investors can reallocate capital to other pillars when there is nowhere to hide in innovation.

Stepping back, we believe investors should think about their portfolios in terms of the pillars driving value-creation today. Intellectual property (including non-patented knowledge) is a key moat in the Information Age. But so are brand, human capital, and network effects. Portfolios should strive to have a nice balance of these intangible assets.

The next exhibit shows a sample intangible value portfolio constructed along these lines. We assign each company to its primary intangible pillar (e.g., Harley-Davidson to brand; Goldman Sachs to human capital; Twitter to network effects). Intellectual property is the most crucial pillar today but still constitutes only a minority of the balance sheet.

Exhibit 38
Balance Sheet Composition



Source: Sparkline. As of 12/31/2021.

A portfolio surrounded by multiple moats is more robust than one relying on a single moat. Each intangible pillar comes with unique risks. For example, intellectual property is vulnerable to rapid obsolescence from technological paradigm shifts (e.g., mobile supplanting desktop). Owning multiple pillars helps diversify these varied risks.

From a statistical standpoint, this can be confirmed by measuring the correlation between pillars. The average correlation is only 10%. We smell a free lunch! 🍜

Exhibit 39
Uncorrelated Moats

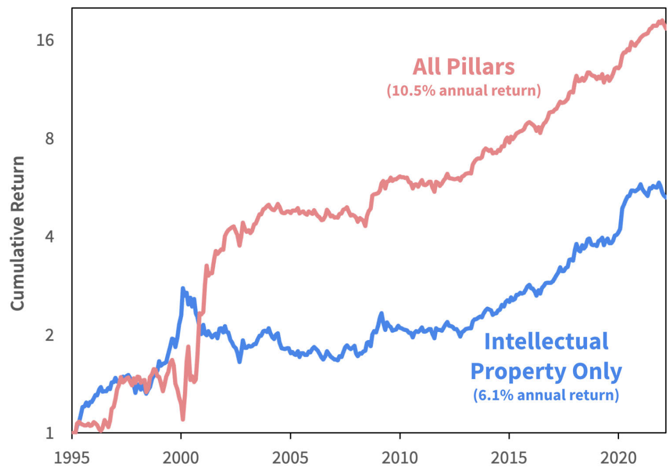
	IP	Brand	Human Capital	Network Effects	Tangible
IP					
Brand	0.21				
Human Capital	0.49	0.28			
Network Effects	0.30	0.12	0.20		
Tangible	-0.11	0.06	-0.12	-0.39	

Source: Sparkline. As of 3/31/2022.

The final exhibit illustrates the value of having multiple intangible moats. First, we build an intellectual property value factor. It is an enhanced version of our patent-based

DARP factor, where we also include metrics to capture non-patented innovation. Next, we compare this factor to an intangible value portfolio consisting of all pillars.

Exhibit 40
Beyond Innovation



Source: Sparkline. Blue line is a long-short portfolio of the top and bottom 150 stocks within the top 1000 U.S. stocks on intellectual property value score. Red line is the same except it uses the full intangible value score. We rebalance monthly and exclude transaction and financing costs. From 12/31/1994 to 3/31/2022. See important backtest disclosure.

The intellectual property factor has positive returns but, as earlier, gets caught up in the dot-com bubble. Fortunately, the other pillars balance out this issue. In fact, the full strategy actually thrives in the dot-com crash. In addition to reducing risk, the full strategy also produces higher returns.

The analysis in this paper supports the idea of allocating some capital to "innovation as an asset class." But investors should not stop there. We believe that investors can do even better in a strategy that finds value not just in innovation but also in the other intangible pillars – brand, human capital, and network effects.

Conclusion

We analyzed two centuries of patent data using machine learning. We unearthed the fossils of past technological revolutions and showed we could predict future tech trends by extrapolating from the past.

Next, we created a rotating list of cutting-edge technologies and invested in the companies innovating in these fields. This innovation strategy outperforms the market and is distinct from growth and other traditional factors. We find the current drawdown is driven not by pure innovation but by a selloff in expensive and unprofitable stocks.

Innovation stocks are prone to become overvalued, as seen in the dot-com bubble. We construct a DARP (i.e., "disruption at a reasonable price") strategy to mitigate this risk. Finally, we show that including other intangible pillars alongside intellectual property both helps to mitigate valuation risk and enhance total returns.

Kai Wu

Founder & CIO, Sparkline Capital LP

Kai Wu is the founder and Chief Investment Officer of Sparkline Capital, an investment management firm applying state-of-the-art machine learning and computing to uncover alpha in large, unstructured data sets.

Prior to Sparkline, Kai co-founded and co-managed Kaleidoscope Capital, a quantitative hedge fund in Boston. With one other partner, he grew Kaleidoscope to $350 million in assets from institutional investors. Kai jointly managed all aspects of the company, including technology, investments, operations, trading, investor relations, and recruiting.

Previously, Kai worked at GMO, where he was a member of Jeremy Grantham's $40 billion asset allocation team. He also worked closely with the firm's equity and macro investment teams in Boston, San Francisco, London, and Sydney.

Kai graduated from Harvard College Magna Cum Laude and Phi Beta Kappa.

Disclaimer

Backtest Disclosure