CONTENTS

1

Introduction

Walk down Broad Street toward the southern tip of Manhattan, and you pass the imposing neoclassical façade of the New York Stock Exchange, police barriers, and—in normal times—tourists taking photographs. Throughout the twentieth century, that famous building, crammed with human traders, epitomized what "finance" meant. A couple of minutes' walk farther south, you would most likely pass 50 Broad Street without a second glance. It has a handsome frontage, and has been renovated internally, but is otherwise an ordinary Manhattan office building (see figure 1.1). In 1993, that stretch of Broad Street, then scruffy and neglected, struck a *New York Times* journalist as exemplifying downtown's decline.[1] More than in any other single place, though, what happened at 50 Broad Street in the 1990s and early 2000s transformed the world's financial markets. Now, just one trace of that role remains: inscribed in panels attached to the stonework above a storefront (which, despite the area's revival, has been empty for years) is the word "island."[2]

Island, launched in 1996, was an electronic venue for the trading of US shares. It was not the first such venue, but none of its predecessors had changed the financial system radically. Some had gone out of business; some had been assimilated into existing ways of doing things; some had succeeded modestly but had not come to occupy central roles. Island was different. Its computer system, packed into the basement of 50 Broad Street, consisted almost entirely of cheap machines of the kind you could have bought in a computer store, but it was blazingly fast by the standards of the 1990s. The

FIGURE 1.1. 50 Broad Street. Author's photograph.

interviewee I am calling AF told me that if Island's system received both a bid to buy shares and an offer to sell the same shares at the same price, it could execute a trade in a couple of milliseconds (thousandths of a second), a thousand times faster than the more mainstream electronic system to which it was most comparable, Instinet. To human eyes, trading on Island appeared instantaneous.

Just as consequential as Island's speed was that machines started to trade on it. There had been previous efforts to automate trading, but often they had not gone smoothly. It could be difficult for an automated trading system to interact seamlessly with exchanges' systems, which in the 1980s and 1990s were usually designed on the assumption that traders were human beings, not machines. Indeed, those who ran exchanges' early electronic trading systems often protected their human users from "unfair" automated competition by prohibiting the direct connection of computers to them. In the privacy of their offices, traders found ways to circumvent the

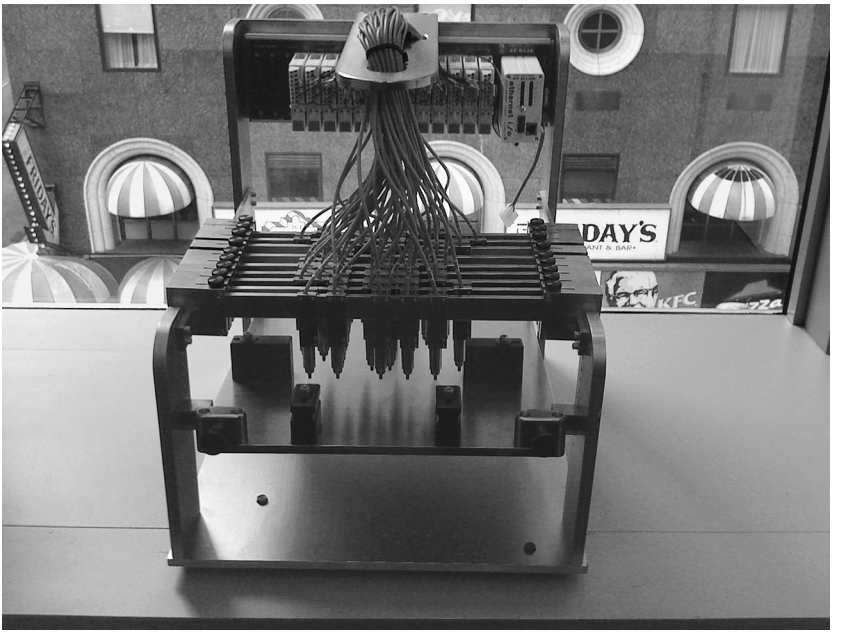

FIGURE 1.2. Lehman Brothers "Clackatron" (ca. 2002), used to strike the keys of an EBS (Electronic Broking Services) foreign-exchange trading keypad. Photograph courtesy of interviewee FL.

prohibition—sometimes even constructing robotic devices to hit the keys of terminals designed for human users (one such device is shown in figure 1.2)—but doing this was cumbersome.[3] Island, in contrast, was machine friendly from the outset. At its core was a set of "order books": electronic files, one for each stock, of the bids to buy the shares in question and of the offers to sell them. Every time Island's computer system executed a trade or received a bid, an offer, or a cancellation of an order, an electronic message was sent out via a continuous datafeed that traders' computers could use to maintain an up-to-date electronic mirror of Island's order books. It was also straightforward for those computers to send Island bids and offers in a fast, succinct, standardized electronic format.

As the machines that traded on Island got faster, the delays that were inevitable if their orders needed to be transmitted to lower Manhattan through hundreds of miles of fiber-optic cable became ever more salient. Dave Cummings, founder of the Kansas City high-frequency trading firm Tradebot ("Trading Robot"), told the *Wall Street Journal* in 2006 that he had come to realize that the 10 milliseconds it took a signal to get from Kansas

City to 50 Broad Street put his firm at a disadvantage: "We were excluded because of the speed of light" (Lucchetti 2006). Starting around 2002, the firms whose machines traded on Island began to move them into 50 Broad Street, at first informally (a web-services firm that had offices in the building hosted their computer servers) and then—in a formal, paid-for arrangement with Island—placing them in Island's computer room in the building's basement, next to Island's heart, the "matching engine": the system that managed its order books and executed trades.

What emerged in and around 50 Broad Street ("emerged" is the right word: no one planned it) is this book's topic: high-frequency trading, or HFT. The practice emerged before the name did; as far as I can tell, the term first came into use at the Chicago hedge fund Citadel in the early 2000s. HFT is "proprietary" automated trading that takes place at speeds far faster than an unaided human can trade and in which trading's profitability is inherently dependent on its speed.[4] (The goal of proprietary trading is direct trading profit, rather than, for example, earning fees by executing trades on behalf of others.) Although the human beings employed by HFT firms to design and supervise trading algorithms often refer to themselves as traders, the trading itself is actually done by those computer algorithms. Humans write the algorithms and (less often now than in HFT's early years) sometimes tweak their parameters during the trading day, but the decisions to place bids to buy and offers to sell are made by the algorithm, not the human being.

HFT algorithms trade both with each other and with other categories of algorithm, such as the "execution algorithms" used by institutional investors—and by banks or other brokers acting on behalf of these investors—to break up a large order to buy or sell shares (or other financial instruments) into much smaller, low-profile "child" orders.[5] HFT firms' algorithms also interact with orders placed manually by human beings, for example by those whom market participants refer to as "retail" (individual investors). Only a minority of retail orders, though, end up being traded on exchanges such as the New York Stock Exchange. Most are executed directly by what are sometimes called wholesalers (which are often branches of HFT firms), who pay the brokers via whom retail investors trade to send them these orders.[6]

HFT firms, in aggregate, trade on a giant scale. For example, as we will see in chapter 4, in just over two months in 2015, eight HFT firms traded Treasurys worth in total about $7 trillion. (Treasurys are the sovereign debt securities of the United States. A trillion is a million million.) The anonymity of most of today's trading makes it difficult in most cases to be certain just

how much of it is HFT, but observers often estimate that HFT accounts for around half of all trading on many of the world's most important markets (see, e.g., Meyer and Bullock 2017; Meyer, Bullock, and Rennison 2018).

The HFT firms that are responsible for these huge volumes of trading are typically recently established and small. Only a small number date from before 2000, and even an HFT firm with no more than a few dozen employees can be a significant player. Consider, for example, Virtu, an HFT firm whose headquarters, as it happens, are just a few blocks away from 50 Broad Street. Virtu's primary activity is "market-making"—continuously posting both bids to buy shares or other financial instruments and slightly higher-priced offers to sell them—and it does this in more than 25,000 different instruments traded in 36 countries. It is responsible, for example, for around a fifth of all US share trading.[7] It rose to its dominant position, my interviewees report, while employing no more than 150 people (its headcount has risen recently because of its acquisition of two firms with more labor-intensive businesses).[8]

In particular niches, even firms with only a handful of employees can be important. In 2019, an interviewee calmly told me that his tiny European HFT firm was responsible for 5 percent of all the share trading in India. Some big banks used to be active in HFT, but their efforts were often less than fully successful; the rapid development of the fast, highly specialized software systems that are needed can be difficult in a large, bureaucratic organization. Banks are still engaged in market-making in some classes of financial instrument (such as those discussed in chapter 4: foreign exchange and governments' sovereign bonds), albeit often using systems that are slow by HFT standards, but large-scale use of other HFT strategies by banks was effectively ended by the curbs on banks' proprietary trading that followed the 2008 banking crisis.

The HFT firms I have visited differ widely. Some had offices in unremarkable or even scruffy buildings; others had spectacular views over Lake Michigan, Manhattan, or Greater London. The décor is generally bland, although as I sat waiting for an interviewee in one HFT firm's new offices, some of the owner's art collection was ready to be hung. The paintings were wrapped and unlabeled, but I'm told they are very fine: the owner has good taste and the firm has been highly successful. More often, though, HFT firms' premises could pass for those of a generic dot-com firm, and they usually have something of the relaxed feel of a software start-up. The employees of HFT firms are mostly young and—at least in the roles closest to trading—mostly male. Office kitchens, for example, often contain multiple boxes of

breakfast cereal, stereotypically young men's food. I am happy to report, though, that the sexist pinups that sometimes used to disfigure trading floors are no longer to be seen. Almost no one in HFT routinely wears a business suit—it is common for me, as the visitor, to be the only person wearing a tie, and I've been told off for being overdressed—and the shouting and swearing that used to be heard on banks' trading floors is less common in HFT firms. That might, of course, be because of my presence, but interviewees tell me that such behavior is indeed less prevalent. As discussed below, I have visited firms only in the US and Europe. There, at least, white faces predominate, though often intermingled with those of South Asian or Chinese extraction, while African Americans, for example, seem rarer.

The internal organization of the HFT firms from which my interviewees come varies. Some operate as unified entities, without even the traditional individual P&L (a trader's profit or loss, the prime determinant of her/his bonus); one firm had a computerized "signal library"—an electronic compendium of data patterns useful to HFT algorithms—that was accessible to all its traders and software developers. Just as Lange (2016) discovered, though, other HFT firms are divided into strictly separate trading teams, with deliberate barriers to communication. One firm, for example, physically separates teams by placing a row of administrative staff between them, and in its main offices even plays white noise between the rows to reduce the chance of members of one team overhearing what is said by members of another. Another firm compartmentalizes its trading by dividing up its long, narrow trading room with white curtains that prevent members of one team from seeing what others are doing. At one compartmentalized firm, said a young trader (interviewee AC) who worked there, "you . . . could get in trouble for being in the next room talking to someone you're not supposed to talk to."[9]

High-frequency trading, however, does not actually happen in these rooms. Instead, it takes place in the computer datacenters of exchanges and other trading venues, which typically contain both the exchange's computer system and the systems of HFT firms and other algorithmic traders, of banks, of communications suppliers, and so on.[10] Exchanges' datacenters aren't generally found in city centers, but in suburban areas in which real estate is cheaper. The datacenters important to HFT are mostly large buildings, and indeed they usually look like suburban warehouses, with, for example, few windows. They are packed with tens of thousands of computer servers, typically on racks in wire-mesh cages (although sometimes the cages have opaque walls, so that a trading firm's competitors cannot see the

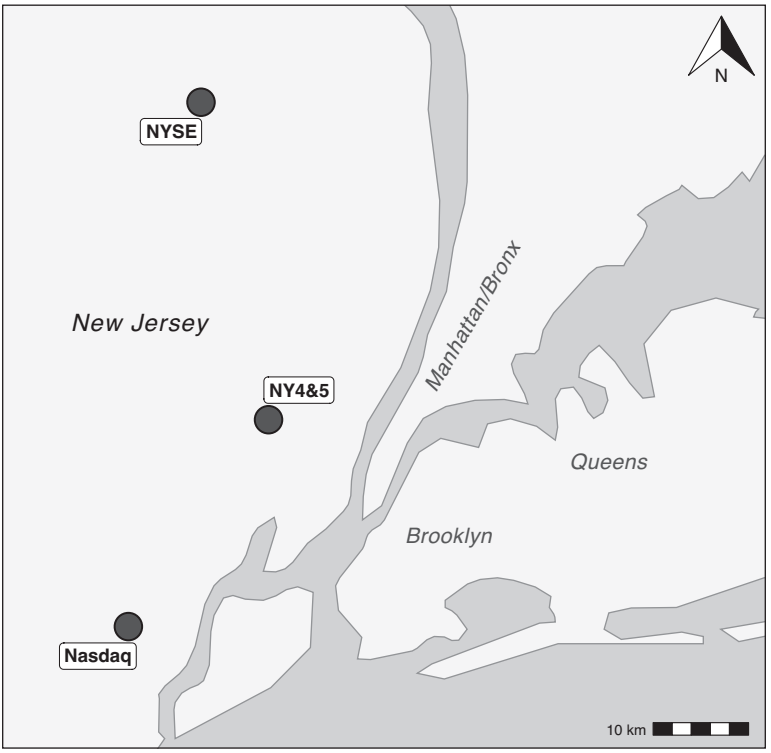

FIGURE 1.3. The "equities triangle" in New Jersey. The Nasdaq and NYSE (New York Stock Exchange) datacenters host the share-trading exchanges run by those groups; NY4 and NY5, which are in effect a single datacenter, host the share-trading exchanges run by the third of the exchange groups, the Chicago Board Options Exchange. (Locations in this and other maps are given only approximately.)

equipment it is using). The servers are interconnected by mile upon mile of cabling, typically running above the racks in what looks to an outsider like an incomprehensibly complex spaghetti of different types of cable. In aggregate, those servers consume very large quantities of electricity and generate large amounts of heat, making a powerful cooling system also a requisite. Normally, few human beings are to be found in these datacenters, just a small number of security and maintenance staff, along with (at least some of the time) engineers from exchanges, trading firms, or communications suppliers who may be visiting to fix problems or install new equipment.

No more than around two dozen datacenters globally host the bulk of the world's financial trading and the vast majority of its HFT. Most US share trading, for example, takes place in the four datacenters in northern New Jersey shown in figure 1.3. One is owned by what is now the New York

Stock Exchange's parent company, the Intercontinental Exchange. Another is leased by Nasdaq, traditionally the main rival to the NYSE as a trading venue for US shares. Two further datacenters (NY4 and NY5) host the systems of multiple trading venues, including the third main group of US stock exchanges, now owned by the Chicago Board Options Exchange. NY4 and NY5 are close together, and in practice are run as a single datacenter. Because of this, market participants often refer to the NYSE datacenter, its Nasdaq equivalent, and NY4/5 as the "equities triangle." (An "equity" is simply another word for a share.)

All of the most important US stocks are traded in all of these datacenters. That makes the automated trading going on in one share-trading datacenter a vitally important source of data for algorithms trading shares in the other datacenters: a vital class of "signal," as market practitioners would call it. A signal is a pattern of data that informs an algorithm's trading, for example by prompting it to bid to buy shares or offer to sell them, or perhaps to cancel an existing bid or offer. A signal of the kind used by HFT algorithms is typically a very short-lived pattern of information: in 2008–9 its duration was usually "less than 3–4 seconds" (Brogaard, Hendershott, and Riordan 2014: 2302). By 2015, a signal may have flickered into life for as little as 10 microseconds—in other words, 10 millionths of a second (Aquilina, Budish, and O'Neill 2020: 55). Another source of signals of this kind, which is of great importance to algorithms trading US shares, is what is going on in the share-index futures market, which is not in New Jersey but in a datacenter in the suburbs of Chicago; see figure 1.4. (A "future" is a standardized, exchange-traded contract that is economically close to the equivalent of one party undertaking to buy, and the other to sell, a set quantity of some underlying asset on a given future date but at a price agreed upon at the inception of the contract.) For reasons to be discussed in chapter 2, the prices of share-index futures in Chicago tend to move a tiny fraction of a second before the corresponding movements in the prices of the underlying shares in the New Jersey datacenters.

When I began my research, I imagined that the data patterns that informed trading by HFT algorithms would be quite complicated, and that those involved would have had to use sophisticated machine learning to discover those patterns. Although machine learning does play a role in the activity (there are examples of this in chapter 6), it is less central than I had assumed. In many ways the most crucial signals for HFT are the kind of relatively simple data patterns just discussed, patterns that often arise (as the following chapters will show) from the way trading is organized and regulated. Those

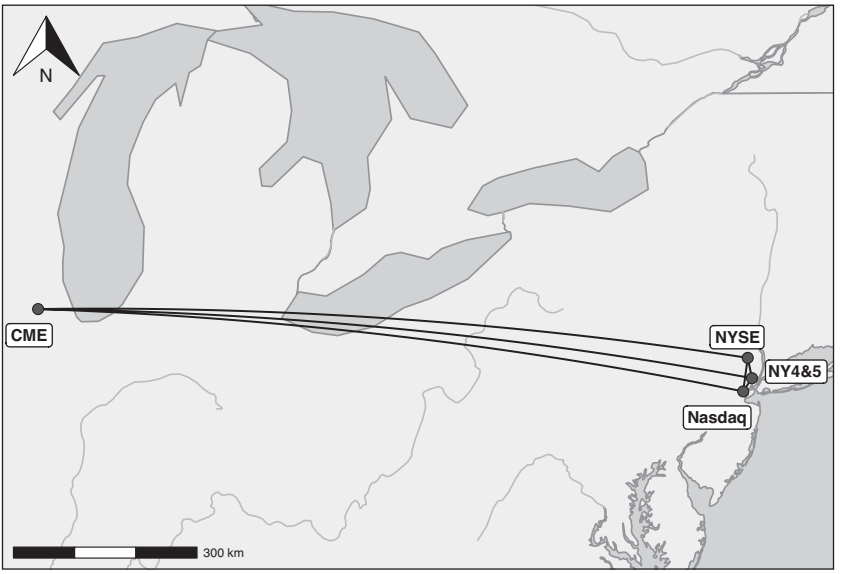

FIGURE 1.4. Geodesics from Chicago to the New Jersey share-trading datacenters. "CME" is the main Chicago Mercantile Exchange datacenter.

patterns are common knowledge in the sector, which means that how fast an algorithm can respond to a signal such as a price movement in share-index futures is vital to whether an algorithm's trading is profitable or loss-making.

Because HFT is, and has to be, very fast (we will see just how fast in the next section), the speed of transmission of signals among datacenters is crucial. That makes the geodesics among them the site of intense activity by communications suppliers, originally mainly using fiber-optic cables, but now using wireless links as well. (A geodesic, or great circle, is the shortest path on the surface of the earth between two given points.) Indeed, US share trading now takes place in what—were it not for the fact that no one planned it and it has no fully coherent overall design—could be called a large technical system, made up of the tens of thousands of machines in the datacenters whose locations are shown in figures 1.3 and 1.4, and of the communications links along the geodesics among these datacenters. Huge volumes of electronic messages (above all, reporting changes in exchanges' order books) flow through this system. The market-data-processing firm Exegy continuously measures the numbers of messages flowing through its equipment in NY4; at the time of writing, the peak recorded on its system was a burst equivalent to 105.3 million messages per second, at 2:39 p.m. on July 19, 2018.[11]

The core systems of automated trading can keep working with little direct human intervention. That became evident in March 2020, as lockdowns belatedly began in Western countries and it finally became clear to their financial markets just how serious the coronavirus epidemic was. Huge amounts of turbulent trading took place, and crucial markets were badly disrupted, including the market for the traditionally safest of safe assets, Treasurys, which as already noted are the sovereign debt securities of the United States. In April, the prices of oil futures even briefly became negative, as a result of the combination of reduced demand for oil and difficulties in storing it. Nevertheless, "the market's plumbing held up" (Osipovich 2020). The turmoil was not exacerbated by major failures of the infrastructures of automated markets. While there are certainly risks involved in automated trading (as discussed in chapter 7), this quiet achievement should also be recognized.

Material Political Economy

This book belongs in the "social studies of finance," the collective name for research on finance not by economists but in wider social-science disciplines such as anthropology, sociology, politics, and science and technology studies. That research has grown rapidly in the last twenty years, and includes, for example, research on HFT. Although my research builds on the work of many colleagues, this book's readers will not all be specialists, so I've put discussion of the existing literature on HFT (including the work of economists) into an appendix at the end of the book. I do, though, need to explain the approach this book takes to the analysis of HFT, which I call "material political economy." It is a single idea, not three ideas, but let me explain it by taking each of the words in turn: all three—"material," "political," and "economy"—are significant.

"Material" indicates a fundamental feature of this book. The previous section has already begun to sketch the material arrangements of today's US share trading. The chapters that follow (especially chapter 5, but not that alone) focus, in as great a depth as is relevant to the book's themes and as my research data allow, on HFT's materiality. Human beings' bodies are part of the material world—if you have any doubts as to whether a human body is material, wait until you have an aging one—and, as Borch, Hansen, and Lange (2015) discuss, the mundane materiality of human bodies is important to HFT. Consider what human eyes and brains can process and what they can't, because it's too fast; what one trader, interviewee OG, calls the "toilet

test" (do you trust an algorithm sufficiently to leave it running unsupervised while you attend to bodily needs?); and what you may need to do to stay focused and awake in the long hours, especially overnight, in which there is often little activity in financial markets.

Nonhuman forms of materiality are, however, much more salient than human bodies in the chapters that follow. HFT is trading *by* machines (trading firms' computer servers and other equipment) *on* machines: all modern exchanges are, at their heart, computer systems. The characteristics of machines, and how those characteristics have changed through time, are hugely important to HFT. Materiality, though, does not refer only to solid objects. Light and other forms of electromagnetic radiation are just as material, and just as salient to HFT, as cables and silicon chips are. The reference to the speed of light in this book's title refers to the need in HFT for the fastest possible transmission of data and orders to buy or to sell.

I think of the materiality of HFT as "Einsteinian." By introducing the name of the celebrated physicist, I don't mean to imply that it's necessary to apply his theory of relativity to understand the aspects of HFT covered in this book, because I don't think that's so, except in limited respects.[12] Rather, the Einstein I invoke is the one portrayed by the historian of physics Peter Galison: an Einstein who was not just a theoretical physicist, but also an inspector in the patent office in Bern, Switzerland, familiar with the technologies of measurement and the practical problem of ensuring the synchronicity of clocks in different spatial locations—Einstein as what Galison (2003: 255) calls a "patent officer-scientist." (Clock synchronization, it is worth noting, is just as prominent a problem in HFT as it was in the railway networks of the late nineteenth and early twentieth centuries. One of my HFT interviewees, CQ, told me how his firm's trading had been badly disrupted by a failure of synchronization.) Einstein's thinking about practical, technological issues such as synchronization, Galison suggests, lay in the background of his development of the theory of special relativity, with its famous postulate that the fastest any signal can travel is the speed of light in a vacuum.

That limit is the fundamental material constraint on HFT. In the early years of HFT, transmission between datacenters was generally via laser-generated pulses of light in fiber-optic cables, but (as described in chapter 5) that gets you only around two-thirds of the way to Einstein's maximum signal speed, because light pulses in these cables are slowed by the materials from which their strands are made, which are specialized forms of glass. In contrast, a wireless signal sent through Earth's atmosphere travels at very nearly

the speed of light in a vacuum. Because, however, wireless transmission for HFT requires radio frequencies that are in high demand, tailor-made radios, and antennas in specific locations (see chapter 5), it is much more expensive than the routine use of fiber-optic cable usually is. One interviewee, indeed, spoke of trying to avoid what he called "radio-frequency markets": those in which an HFT firm has no alternative but to use signals transmitted through the atmosphere.

One way of gauging the speed of HFT is the response time of an HFT firm's system: the delay between the arrival of a "signal" (a pattern of data that informs an algorithm's trading) and an action—the dispatch of an order or a cancellation of an order—in response to that signal. In March 2019, an interviewee told me that, although his own systems were slower than this, he had learned of the achievement of response times as low as 42 nanoseconds.[13] A nanosecond is a billionth of a second, and in a nanosecond, even light in a vacuum, or a wireless signal in the atmosphere, travels no more than around thirty centimeters, or roughly a foot.

That nanoseconds are important in HFT makes its world Einsteinian: for HFT, that no signal can travel faster than the speed of light in a vacuum is a practical constraint, not just a theoretical limit. For a signal to travel even as short a distance as a meter takes what is potentially an economically consequential amount of time, and that makes HFT exquisitely sensitive to the precise location of technical equipment and to how closely the path of a fiber-optic cable or wireless link hugs the geodesic between datacenters.[14] The materiality of HFT is, therefore, above all a *spatial* materiality. It's easy to think of what is sometimes called today's postmodernity as involving the shrinking of both time and space.[15] In an Einsteinian world, though, as time shrinks, space becomes ever more salient.

The computer specialists who work for HFT firms have to be materialists in their thinking and their practices. One such specialist with whom I chatted during a coffee break in a traders' conference in Amsterdam told me that he had had to unlearn the attitude that he had unwittingly picked up during his time as a computer-science student. He could not, as he had implicitly been taught, safely abstract away from the physicality of the hardware on which his algorithms run. A computer, from the viewpoint of HFT, is not an abstract information processor, but a material assemblage of plastic, metal, and silicon through which electrical signals flow, and making them flow as quickly as possible is a vital practical concern. When I use the word *algorithm* in this book, I don't mean the word in the dominant sense in which it is used in computer science: a "recipe" that achieves a goal or solves a problem

in a finite number of precise, unambiguous steps, and which is abstract in the sense that it can be implemented in different programming languages running on different machines. Rather—and I am following my interviewees' predominant usage here—an algorithm is a recipe of this kind written in a particular programming language, running on particular physical hardware, and having material effects on other systems.[16]

I didn't begin my research on HFT with the concept of material political economy in mind. The notion evolved as I conducted my fieldwork, and it seems to me a useful way of framing the research and of capturing its findings. I don't want, however, to try to draw an ontological divide between "material" and "nonmaterial" phenomena, or to suggest that we should focus on the materiality of economic life and exclude everything else. Nor do I see "material political economy" as making redundant other ways of studying economic phenomena, such as "cultural economy" (du Gay and Pryke 2002), "cultural political economy" (Jessop 2009), or, for example, the various forms of "international political economy" pursued by scholars in politics. Even a quintessentially material business such as HFT is influenced by factors that we wouldn't ordinarily think of as "material": beliefs, metaphors, epistemic authority, legitimacy, and so on. (Ultimately, all of these factors come down to material phenomena: words or images on paper or in other media, and sometimes other physical objects; the soundwaves that encode speech, and so on, including material patterns of neural activity in human brains. However, while the materiality of "culture" in this sense is indeed sometimes important, it would be facile to argue on these a priori grounds that it should always be focused on.)

Consider legitimacy, for example. As I will shortly discuss, the history of HFT has been marked by systematic conflicts with trading's incumbents. Scandals that have undermined those incumbents' legitimacy—such as the Nasdaq scandal in the 1990s touched on in chapter 3—have been important in creating opportunities for the rise of HFT. Similarly, as will be described in chapter 6, a crucial internal divide in HFT is between "market-making" strategies (which, as already noted, involve continuously posting both bids and offers in order books that others can execute against) and "liquidity-taking" strategies, which involve executing against orders that are already present in order books. Market-making inherits the legitimacy of a traditional human role in markets, and some—although by no means all—of my interviewees regard it as a preferable, even a more moral, economic activity than liquidity-taking. It is true that the extent to which this preference shapes particular forms of trading, rather than simply being invoked to justify them,

is questionable. After one of the leaders of an HFT firm emphatically presented its activity to me as market-making, another interviewee (who had recently left the firm) described that image as a "legitimatory" move rather than fully factual. However, that the role played by legitimacy is often ambiguous in this way certainly does not imply that it should be ignored. The materiality of human activity is inseparably bound up with questions of belief, the achievement (or loss) of authority, and so on, and to close one's eyes to this would be to pursue an impoverished form of research.

Material *Political* Economy

HFT's materiality is nevertheless crucial, and the approach I'm taking owes a great deal to the perspective—called actor-network theory—that has in recent years done most to place materiality at the heart of social sciences.[17] (Marxism was also originally a thoroughly materialist intellectual enterprise—that is certainly the case, for example, in the memorable passages on machinery in volume 1 of Marx's *Capital* [Marx 1976]—but twentieth-century developments, such as the rise of more philosophically oriented Western Marxism, in practice rather weakened this emphasis.)[18] In a lovely contribution to actor-network theory, John Law and Annemarie Mol (2008) discuss what they call "material politics." The idea is in essence simple, and it has antecedents—we will encounter one of them at the start of chapter 7, in research conducted in the 1930s by the historian Marc Bloch—but is elegantly laid out by Mol and Law. As they argue, it is possible to arrange the material world in different ways, and at least sometimes the issue of which of these ways becomes real has a political dimension. There are, for example, "roads not taken" in the development of technology, as the historian Ruth Schwartz Cowan (1983) points out, drawing the phrase from the poet Robert Frost: technologies that could have been developed but weren't, and not necessarily because they were simply less efficient than the successful alternative, but sometimes for reasons that have more to do with class, gender, and ethnic divides, state power, and so on.[19]

"Material politics" is a pivotal aspect of HFT. As already noted, the activity's history is characterized by incumbent-challenger conflicts, in which HFT firms have traditionally been in the role of challenger; some of those conflicts continue today. Conflicts of this kind are emphasized by the sociological perspective known as field theory, and elsewhere I've argued that the materiality of actor-network theory needs to be complemented by the field-theory emphasis on such issues, despite the occasional bitter clashes

between the two perspectives.[20] Incumbent-challenger conflicts in finance are played out in multiple dimensions, but the material arrangements of trading are certainly one of those dimensions. In chapter 6, for example, we will discuss a material procedure ("last look") that has protected the incumbents in foreign-exchange trading from HFT firms' faster systems.

The divide, already mentioned, between "market-making" and "liquidity-taking" algorithms fuels much of today's material politics of HFT. The existing bids and offers that liquidity-taking algorithms execute against are, in most of the world's leading electronic markets, often those that have been submitted by HFT market-making algorithms, so that a good proportion of the profits made by liquidity-takers is most likely at the expense of the latter. As discussed in chapter 6, there is at least a degree of differentiation among trading groups and sometimes even entire firms in the extent to which they specialize in either market-making or liquidity-taking, and the material arrangements of trading can tilt the playing field in favor of one or the other. These arrangements are, therefore, a form of material politics.

More mundanely, HFT firms (even though small and often closely steered by their founders) are not immune from internal disharmony and office politics. For example, even in a compartmentalized HFT firm with strictly separate trading groups, there is usually a common technical and communications infrastructure, and how access to that infrastructure is shared among the groups can reflect what interviewee DC calls "a . . . political dynamic." When, for example, the firm for which he worked started to lease wireless bandwidth, "it was the politics of who gets to use that line. . . . It was such limited capacity, then people started really fighting over it." Since the trading groups in a compartmentalized firm have their own—possibly very different—approaches to HFT, the overall trading activity of such a firm may sometimes be shaped substantially by the outcomes of struggles of this kind.

"Politics" and "political" are, of course, elastic words. In this book, I use them mainly in their broadest sense, which refers to the full gamut of phenomena that shape and are shaped by actors' power and position, the status and respect they enjoy or fail to receive, their economic resources, and other factors. However, politics in the narrower sense of political parties, members of Congress, congressional committees, and so on, has also played a part in the development of HFT, especially via the interaction between the political system and the regulation of finance. (When writing about politics in the narrower sense, I will sometimes for clarity use the ordinary term "political system," although once again "system" suggests a coherence that

is often absent.) The interaction between the political system and financial regulation takes a variety of forms, including, for example, differences among political parties in their typical attitudes to regulation. In the US, for instance, Democrats have generally preferred stricter forms of regulation, and Republicans often have deregulatory impulses. Politicians, furthermore, are often lobbied by financial-sector interests, and money from the financial sector can form an important part of campaign contributions to politicians (there is an example of this, and of its consequences, in chapter 2).

The tightness of the connections between financial regulation and the political system should not, however, be overstated. Again, sociological field theory is relevant. "Politics" and "regulation" (and, indeed, extending the point, "exchanges" and "trading") are all what sociologists such as Pierre Bourdieu and Neil Fligstein call "fields," or what the Chicago sociologist Andrew Abbott calls "ecologies." In other words, they are specific domains of social and economic life, characterized by differently positioned actors competing and collaborating to achieve rewards that are often specific to that field or ecology. (The key reward for politicians, for example, is usually votes.) Fields are often characterized by an implicit sense of what forms of behavior are legitimate, and there are sometimes explicit rules about what actors should and should not do. We should not, however, expect to find consensus. Rules often privilege some actors over others, and one form that competition can take is for challengers to seek to change a field's norms and rules.[21]

Although different fields thus differ in their specific dynamics, in their norms and rules, and in the rewards at stake, developments in one field can affect adjacent fields in important ways. One way is via what Abbott (2005) calls a "hinge": a process that generates rewards in more than one field or, in his terminology, more than one ecology.[22] As Abbott points out, though, the political system differs from many other fields or ecologies in that issues that are continuously important in those other fields (for example, licensing in nineteenth-century medicine; or the organization and regulation of trading) are salient only sporadically in the political system, because most of the time there are few directly political rewards for pursuing them. The hinges that link the field of politics (in the narrow sense) to the fields of finance—or, in Abbott's chief example, to the ecology of the professions—are thus often transient and contingent, even idiosyncratic. That does not mean, however, that they are unimportant. In chapter 2, for example, we will see how the lasting effects of one such idiosyncratic 1970s-era hinge have profoundly shaped automated trading in the US.

Material Political *Economy*

Why, though, do we need the third word, "economy"? Surely, everything to do with finance is self-evidently economic? The reason for emphasizing the economic (indeed, the monetary) aspects of finance is that they have often been given insufficient attention in the specialist area of the social sciences to which this book belongs. That area, to repeat, involves the application to the study of finance not of economics, nor of individualistic "behavioral finance," but of wider social-science disciplines such as anthropology and sociology. This specialism first crystallized in the late 1990s, and it's understandable that in its early years it usually focused not directly on money-making—perhaps implicitly seen as covered by existing scholarship, especially in economics—but on other aspects of finance of a kind more familiar in those wider disciplines. (I intend no criticism of my colleagues here—that was just as true of my own work in that period as it was of that of others—and some of them, especially Olivier Godechot, did pay attention to money-making; see, especially, Godechot 2007.)

What in particular has been neglected is what I think of as the "mundane" political economy of finance: the undramatic, everyday ways in which money is made, often individually fairly small amounts of money, but time and time again.[23] It is easy when one is new to the study of finance—as I was, twenty years ago—to focus on its dramatic aspects (finance's giant crises; the making or losing of huge sums; and so on), and ignore the undramatic. What was for me, therefore, a surprising side benefit of researching HFT is how often its development throws light on preexisting mundane ways of money-making. Money-making was and is often made possible by the occupancy of favorable positions in what practitioners call "market structure," by which they mean the way in which a market is organized, especially the formal and informal rules of the game that dictate matters such as who or what can trade with whom, and on what terms; how information flows; where its flows are blocked, and so on.[24] Precisely because the rise of HFT has often involved challenges to aspects of market structure such as these, it thereby renders them visible. Although sociological field theory has had no influence that I can detect on practitioners' use of the notion of market structure, it is perfectly reasonable to see market structures as the very core of the various fields that make up the financial system.

Let me give an example of a mundane but consequential aspect of market structure. As I've already mentioned, trading on Island was, and trading on the majority of the most crucial of today's electronic markets is, organized

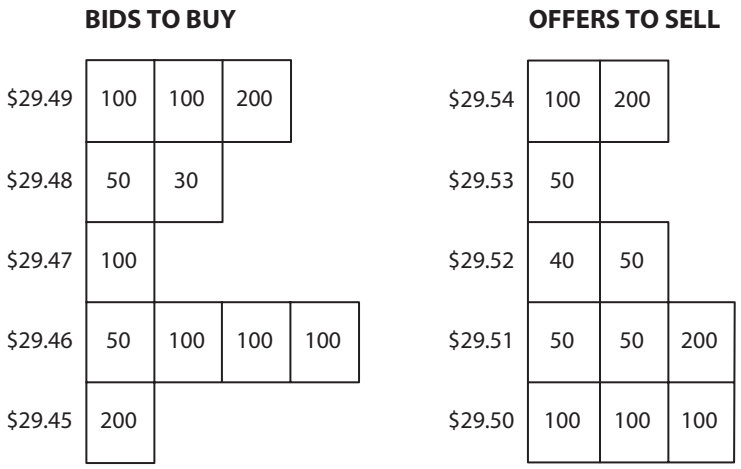

FIGURE 1.5. An order book. *Source:* author's interviews and observations of trading.

around "order books"—lists of the bids to buy and offers to sell the stock or other financial instrument being traded that have not yet been executed. (For a visual representation, designed for human eyes, of an order book, see figure 1.5.) On Island, and in most of the other electronic markets discussed in these chapters, order books are visible—in electronically mediated ways—to all the humans and machines trading on those markets. That, however, has not always been the case. As will be described in chapter 3, until the early 2000s a stock's order book on the New York Stock Exchange was largely private to the designated "specialist," the trader who coordinated trading in that stock. Initially, indeed, an NYSE order book was almost literally a book, made up of preprinted forms on which orders were handwritten by specialists or their clerks. An NYSE rule introduced in June 1991 required specialists to "share general information about their books with other floor traders on an informal basis when asked" (Harris and Panchapagesan 2005: 26). Prior to that, order books were—at least in principle—accessible only to specialists, their clerks, and NYSE officials.

An order book, whether handwritten or electronic, often contains information that is extremely helpful to trading. If, for example, the book contains many more bids to buy than offers to sell, it is reasonable to anticipate that prices are about to rise. (The balance between bids and offers and how that balance is changing—along with the sequences of executed trades—form a crucial class of "signals" for HFT.) The mundane issue of who or what has access to the order book can therefore be economically consequential.

Access, furthermore, is an inherently material process, an issue of which data flow to which computer systems and when, or—in the case of human traders—what they can and cannot see. As touched on in chapter 3, when the NYSE's order books were handwritten on paper, a trading-floor broker could sometimes catch a useful glimpse of a book's contents, because a specialist or a clerk had to open it to write down the broker's order. (A broker in a financial market is an intermediary who either executes trades on behalf of a customer—that is the main sense in which the term is used in the NYSE—or arranges trades between others.)

As the sociologist Mitchel Abolafia showed in his pioneering early 1990s ethnographic research on the NYSE, its specialists' exploitation of their central role in trading was constrained by phenomena of a kind familiar to academics in the social studies of finance like me: the NYSE's formal rules and its monitoring of specialists' trading, its informal trading-floor culture, and floor brokers' countervailing power (Abolafia 1996; see also Mattli 2019). Yet that conclusion needs to be balanced against the reality that the NYSE's market structure *did* give its specialists an informational advantage. Using three months of unusually detailed NYSE data (November 1990–January 1991, which is within the period of Abolafia's observations), Harris and Panchapagesan showed that the contents of order books were predictive of price movements, and they concluded that specialists "use this information in ways that favor them" (2005: 25). However, their data did not enable them to determine "[w]hether this advantage produces significant trading profits" (2005: 27) or whether such profits outweighed the risks inherent in the specialists' obligation to continuously "make a market," bidding even when there were few other buyers or offering to sell when there were few sellers.

The rise of HFT has eroded or eliminated a variety of older market structures, such as that of the NYSE's trading rooms, and created in their place a different set of mundane political economies. To put it in capsule form, the shift has been from structures whose pivot was *who* could see the order book (or its loose equivalents in other preexisting market structures) to structures in which money-making is affected crucially by *when* the order book can be "seen." The book is electronically visible to all, but how long it takes their systems to receive, process, and respond to order-book updates is a critical determinant of whether a trading firm makes or loses money. (There are exceptions to the visibility of order books even today, which include the trading venues called "dark pools."[25] If a dark pool has an order book—and not all do—it is not visible to participants. I have written about

dark pools in an open-access article [MacKenzie2019d], but have decided not to discuss them in this book to avoid overcomplicating it.)

In HFT, the margin between making and losing money is slender. The profits of HFT are mundane in the sense that what can typically be earned on individual HFT trades is very small. Drawing on documents released at the time of the initial public offering of the shares of Virtu, the HFT firm discussed earlier in this chapter, Laughlin (2014) estimates that Virtu was on average earning trading revenues of 0.27 cents per share traded. He calculates that if it were typical of HFT at large, it would imply that HFT firms were earning revenues, in aggregate, of around $2.5 billion annually trading US shares; by comparison, the investment bank Goldman Sachs's total revenues in 2013 were $34.2 billion (Goldman Sachs 2014: 1). A quarter of a cent per share traded is mundane enough, but that quarter of a cent is considerably reduced when HFT firms' heavy expenses are taken into account.

Against, for example, Virtu's 2013 trading revenues of $624 million must be set operating expenses of $477 million, which included $195 million on brokerage, exchange, and clearing fees and $65 million spent on technology and communications links (Virtu Financial 2014: 73). I have found it difficult to get HFT interviewees to talk freely about the activity's profitability after fees and other expenses. It is a sensitive topic, because HFT firms quite often fail financially, and my impression is that the most common way in which they do so is not by losses in trading but when revenues from trading are swamped by expenses. Nevertheless, there seems to be some consensus among interviewees that 0.05–0.1 cents per share traded (or its rough equivalent in other asset classes) is a healthy rate of profits net of expenses. Even if that is an underestimate (and it may not be), it does indicate the narrowness of the economic difference in HFT between success—regular tiny profits on huge volumes of trades do add up[26]—and failure: the inability to earn revenues that exceed a firm's expenses.

Chapter 5 will discuss the technologies and communication links on which HFT firms spend so much money, and chapter 6 why they often have no alternative but to do so, but the question of fees is worth briefly discussing now because it throws light on another aspect of the mundane economics of HFT. One of this book's themes is that it is impossible to entirely separate out discussion of HFT, and its place in the "fields," or "ecologies," of trading, from analysis of the exchanges and other trading venues on which it takes place. Those venues form a distinct ecology. At least to a first approximation, exchanges compete with each other, not with HFT firms.[27] As discussed in chapter 3, however, Abbott's "hinges" are at work here. The mundane

economics of exchanges (which are nowadays almost all profit-making commercial firms or subsidiaries of such firms) has become interwoven with the economics of HFT.[28] Many exchanges earn much of their income from HFT firms, not just in trading fees but also in fees for receiving the fastest versions of the streams of electronic-update messages necessary to mirror exchanges' order books, and for the "cross-connect" cables within datacenters through which those streams are transmitted. Incumbent exchanges therefore profit from HFT's speed race, which (as Budish, Lee, and Shim 2019 point out) may make them reluctant to materially reorganize trading to mitigate this speed race.

Once again, I must emphasize that I see attention to finance's mundane money-making as complementary to research on its more dramatic aspects such as financial crises, not as displacing it. (Indeed, crises often have their roots *in* mundane money-making.) Mundane money-making, though, should be an important topic for researchers in the social studies of finance because it is a vital determinant of what goes on: for example, which technological or organizational changes are readily adopted and which are resisted fiercely. My favorite, because sublimely mundane, example of resistance is what I think of as "the battle of the asterisk," but I save that example for chapter 7.

There is a second, quite different, argument for more research on finance's mundane money-making, implicitly revealed in a remarkable article by the economist Thomas Philippon (2015).[29] I've written about his article elsewhere—initially in the *London Review of Books* (MacKenzie 2016)—but the point is worth repeating: it sounds esoteric, but its implications are large. What Philippon has done is to measure the "efficiency"—the unit cost of financial intermediation—of the US financial system through time.[30] Strikingly, and quite surprisingly, his data, shown in figure 1.6, do not reveal any clear tendency for finance's efficiency to increase between the 1880s (the era of clerks writing in ledgers by pen, perhaps by gaslight) and 2012, in the epoch of HFT and the iPhone.

If Philippon's data are correct in suggesting that the unit cost of financial intermediation has changed little over this long period, why should that be so, given that the information and communication technologies that underpin finance have improved so radically? One possible answer is that most of those efficiency gains have been captured by mundane money-making within the financial system, and take the form of extremely high remuneration for elite employees of banks and other financial firms, along with (less reliably high) dividends and capital gains for the shareholders in those firms. If that is so, we have here a crucial component of societal inequality in income and wealth,

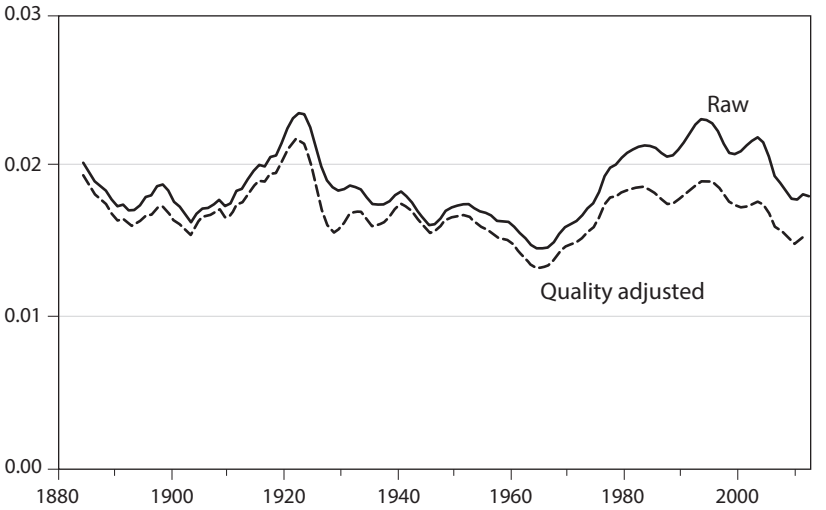

FIGURE 1.6. The unit cost of financial intermediation in the United States, 1884–2012. Data courtesy of Thomas Philippon. For details, see endnote 30 of chapter 1 and Philippon (2015).

because high rewards in the financial system have contributed palpably to that inequality in recent decades.[31] For example, between the 1940s and 1980s, workers in the US financial sector earned salaries similar to those in other sectors with comparable educational qualifications. Thereafter, their relative salaries rose markedly, so that eventually senior executives in finance were earning two and a half times more than their equivalents elsewhere (Philippon and Reshef 2012). Although there is no definitive proof, what may be involved in this dramatic change in the fortunes of the finance sector is an increase in what an economist would call "rent," defined by the commentator Martin Wolf (2019) as "rewards over and above those required to induce the desired supply of goods, services, land or labour."[32] In effect, the financial system may be exacting rents from the rest of the economy, and, of course, among the expected consequences of this would be slower growth of the wider economy.

Let me not raise false expectations about the chapters that follow: on no plausible calculation are the aggregate profits, salaries, and bonuses earned via HFT large enough to contribute in anything other than a minor way to overall inequality of income and wealth. Indeed, HFT's slender profit margins suggest that the automation of trading (and the associated changes in market structure) may actually have reduced, quite considerably, the substantial rents that trading's well-placed insiders could traditionally exact. Rent, though, may still be relevant to the mundane economics of HFT; that,

at least, is the conclusion reached by the Chicago economist Eric Budish and colleagues (Budish, Cramton, and Shim 2015; Budish, Lee, and Shim 2019). Consider what happens when an unequivocal "signal" appears in the datafeed from an exchange: if, for example, the price of a share-index future traded on the Chicago Mercantile Exchange rises, or there is selling of Apple's shares on the NYSE, or the bids to buy Amazon's shares in a Nasdaq order book suddenly evaporate. As already noted, my interviews make clear that simple signals such as these are familiar to all high-frequency traders, and their appearance triggers a race to be first to respond to them. Traces of races of this kind can be found in London Stock Exchange electronic-message data from 2015 at an average incidence of around one race per minute for each of the 100 leading stocks on that exchange (Aquilina et al. 2020: 3–4). The algorithm that is fastest to detect and react to a widely employed signal of this kind can, at least on average, make money by "picking off"—that is, executing against—quotes that in the light of the signal are "stale," or out-of-date (or, if the algorithm is following a market-making strategy, it can avoid losing money by canceling its stale bids or offers). As Budish, Cramton, and Shim (2015) put it, "mechanical arbitrage opportunities are built into" the current material arrangements of trading, creating what they call "arbitrage rents" (Budish et al. 2015: 1548).

That usage of the notion of rent may appear abstract, but it translates into rents of a more familiar kind. The race to execute against or cancel stale quotes makes it possible for those who control the means of transmission (such as the "cross-connect" cables that link trading firms' servers to an exchange's computer system), or the locations that are crucial to speed, to charge heavily for access to them. These locations include, for example, racks within the datacenters in which exchanges' matching engines are located, and microwave towers or particular places on datacenter roofs that are especially favorable as sites for wireless antennas. HFT's Einsteinian materiality creates what interviewee DE calls "pinch points," and those who control pinch points can exact rent. (See chapter 6 for another race that takes place among market-making algorithms trying to reach the head of what is in effect a queue for electronic execution.)

Data Sources

I began the research for this book with exploratory interviews in 2010. High-frequency trading (since its inception, almost always a low-profile activity) had been in the news in the previous year, especially following a front-page

story about HFT in the *New York Times* in July 2009, which began, "It is the hot new thing on Wall Street, a way for a handful of traders to master the stock market, peek at investors' orders and, critics say, even subtly manipulate share prices" (Duhigg 2009). I was intrigued. My initial interviewing involved little more than trying to find out more about HFT from people who I thought (sometimes rightly, sometimes wrongly) might be able to tell me more about it.

After those exploratory interviews, I focused on finding—first from published sources, and then by introducing myself to people at traders' conferences or obtaining referrals from earlier interviewees—current employees, former employees, and founders of HFT firms. There are no HFT firms in Edinburgh, Scotland, where I teach, and so to keep my travel from becoming wholly excessive, I concentrated on the four main centers of HFT worldwide: Chicago, New York, London, and Amsterdam. Although there is also HFT in a variety of other markets—Brazil, Canada, India, Singapore, Hong Kong, South Korea, Japan, and Australia, to name a few—I decided for this practical reason to focus my research on the US and Europe. I ended up interviewing 86 practitioners of HFT (see table 1.1), 22 of them more than once, including, for example, one whom I interviewed six times, one seven times, and one nine times. These repeat interviews were particularly useful, because in a first interview an interviewee would often present a somewhat idealized account of events or practices, and—embarrassingly often—I did not at first understand enough technically to be able to focus my questions properly. Multiple interviews with the same person helped me to develop my technical understanding and helped them to begin to trust that I would not misuse the information they gave me, making it possible for me to extend the conversation beyond topics that I had already learned were part of what interviewee AG called "High-Frequency Trading 101."

These interviews with employees or former employees of HFT firms did not follow any set format. They were more like conversations in which I tried to get interviewees to talk about the activity, not always successfully. (One of the most frustrating was a lunch in Chicago with two employees of an HFT firm who had both been traders in Chicago's famous "open-outcry" trading pits. They were more than happy to talk at any length about the pits, but steering the conversation toward HFT was much harder.) Gradually, largely by trial and error, and especially by having more than one meeting with the same person, I began to identify topics that were common knowledge among high-frequency traders. At the same time, I began to get a sense of the kind of thing that would count as an example of "secret sauce," and

TABLE 1.1. Interviewees

Founders, employees, or former employees of HFT firms (AA-DI)	86
Members or staff of exchanges, clearinghouses, and other trading venues (EA-HI)	87
Traders for investment-management firms (IA-IJ)	10
Manual traders (MA-ML)	12
Practitioners of other forms of algorithmic trading (OA-OY)	25
Regulators, lawyers, lobbyists, and politicians (RA-SE)	31
Suppliers of technology and communications links (TA-UF)	32
Researchers and market analysts (VA-VU)	21
Dealers, brokers, and broker-dealers (XA-YG)	33
Total	337

Note: Two-letter codes are used to preserve anonymity.

which interviewees would be reluctant to talk about, perhaps fearing losing their jobs if it became known that they had done so.

There is a fair amount of movement of staff among HFT firms, which makes common knowledge of the techniques of HFT quite extensive. "It's really hard to keep those special nuggets of information special for very long," says interviewee BD. Even without hiring a candidate, interviewing him or her can be a useful way of finding out how things are done at other firms, as well as an occasion on which the senior staff who conduct them can discover that what they have taken to be their "secret sauce" is actually common knowledge.[33] Examples of common knowledge include the three main classes of signal already touched on in this chapter: price movements in relevant futures contracts; the contents of order books; and (in the case of shares or other financial instruments traded on more than one exchange) trading and price movements on those other exchanges. These are all unequivocally part of AG's "High-Frequency Trading 101." In contrast, precisely how a firm's algorithms analyze the changing contents of order books has an element of secret sauce to it. I deliberately avoided asking overly detailed questions about this, and the only time an interviewee volunteered real detail was the last time I met him in his trading group's offices; the group had not been able consistently to earn revenues in excess of its expenses and was closing down.

Some of my interviews took place in cafes, restaurants, bars, or interviewees' homes; one memorable interview involved my taking off my socks, shoes, and trousers and helping an interviewee launch his jet ski into Lake Michigan.

However, many of my interviews with employees or founders of HFT firms took place in those firms' offices, and I was often taken around those offices before or after the interview. Such visits helped give me a sense of the "feel" of those firms, and sometimes conveyed other information as well. (For instance, on my second visit to one HFT firm I noticed that its large, impressively well equipped trading room, fully occupied on my first visit, had emptied out markedly. It was my first direct encounter with the economic precariousness of HFT. A year or so later the firm went out of business.)

Gradually, too, I came to realize that to understand HFT I had to interview more widely than in HFT firms. The opportunities for HFT and important aspects of how it is practiced depend on the rules and material configurations of exchanges and other trading venues, and those venues in turn are influenced, sometimes deeply, by the actions taken (and not taken) by government regulatory agencies. Exchange staff, regulators, lawyers, and others therefore became important categories of interviewee. Because those interviews mostly dealt with topics that were less sensitive, they were typically more straightforward than the interviews with high-frequency traders, and many fewer repeat interviews were required. (Of the 87 employees of exchanges, other trading venues, and clearinghouses that I spoke to, I interviewed only 6 more than once.) In addition, those most intimately familiar with the materiality of HFT's "signals" are often not employees of HFT firms but the specialized communications suppliers to those firms; thus I interviewed them as well. There was a great deal I needed to learn in this sphere; here too, another interviewee allowed me to interview him nine times. Overall, across all categories of interviewee (high-frequency traders, exchange staff, etc.), I conducted 358 interviews with 337 people; see table 1.1 and an endnote.[34]

There are no comprehensive, accessible lists of the relevant populations of high-frequency traders, exchange staff, and the like, so interviewees cannot be guaranteed to be representative, either statistically or in other ways. Despite my best efforts, I am almost certainly guilty of a "seniority bias" in those to whom I spoke. This is partly because relevant senior people are easier to identify, and partly because younger employees, especially in HFT firms, may feel they need to be especially wary of speaking to outsiders about their work. (At least one leading HFT firm explicitly bans its employees from doing so.) I tried wherever possible to interview female occupants of relevant roles, but despite this only 17 of my 337 interviewees are women. High-frequency trading and related fields tend to be male-dominated, but I suspect that seniority bias may also help account for this small proportion of

women. About halfway through the interviewing, I became aware that my sample of high-frequency trading interviewees was biased toward specialists in market-making. My efforts to persuade larger numbers of specialists in liquidity taking to be interviewed were, however, only partially successful.[35]

Despite difficulties of this kind, there was much that emerged unequivocally from the interviews. As, for example, the interviews began to reveal the main classes of financial instrument in which HFT firms are active, it became clear that there are very large differences between those classes in how trading is organized, and sometimes also substantial differences between the US and Europe. Trying to understand and explain those differences then became one of the goals of the research. In this book, I focus on four classes of instrument—futures, shares, governments' sovereign bonds, and foreign exchange—that are all comparable in that they are simple and highly liquid. (The research also encompassed options and interest-rate swaps, but the greater complexity of these instruments makes them less directly comparable. Apart from chapter 6, in which the options market offers an extreme case of what is being discussed, I have not considered these markets, again to avoid overcomplicating the book.)[36]

Understanding and explaining how trading is organized requires taking a historical perspective, because its organization is "path dependent": how trading was organized in the past affects how it is organized today. Liquidity, for example, is "sticky," in the sense that once traders expect a trading venue to be the most liquid, it tends to remain so because they direct their trades there. (As we shall see, however, there are also "political economy" issues that create stickiness.) There is existing literature on the history of electronic trading to draw upon, notably an exemplary study by Fabian Muniesa of the automation of the Paris Bourse, and a fine book by Juan Pablo Pardo-Guerra on developments in the UK and US (see the appendix on the existing literature). Historically focused interviews were, however, still necessary, especially to cover the history of the trading of financial instruments other than shares, and to throw light on the more recent history of share trading. In unraveling these histories, I found I had to speak to incumbents—for example, to sovereign-bond and foreign-exchange dealers and to the interdealer brokers who arrange trades among those dealers—to understand their conflicts with HFT from their perspectives. (Dealers are intermediaries who trade on their own behalf as well as with clients.)

Most of my interviewees preferred anonymity, and quotations from them are identified only by two-letter labels (see table 1.1). Occasionally, I refer to something someone said that is especially sensitive or that might make

it possible for an insider to HFT to guess to whom a label refers: no label is used in such cases. In the small minority of cases in which interviewees were happy for their names to be used, I do so. As far as possible, I tried to check what one person told me against what others had said. I also supplemented the interviewing by attending six traders' conferences (two in London, two in Chicago, and two in Amsterdam), along with an algorithmic-trading training course in New York, three events focusing on cryptocurrencies (see chapter 7), and a meeting attended by many employees of governments' debt-management offices. At such gatherings, it was often possible to chat informally with people whom I could not interview formally. I was also taken on tours of trading floors, including the two that still had genuinely significant roles: the main trading room of the New York Stock Exchange, which is important during the NYSE's daily opening and closing auctions, and the section of the trading floor of the Chicago Board Options Exchange which trades options on the Standard & Poor's 500 share-price index. (Face-to-face trading had, of course, to be suspended as the coronavirus epidemic peaked, and at the time of writing, in June 2020, it has restarted only partially.)

I'm not an economist, and this book does not try to answer the questions that economists have traditionally asked about HFT, such as whether it increases market liquidity or volatility; see the appendix on the literature on HFT. As described in the appendix, though, I have certainly gained insights from economists' research on HFT. That research is also sometimes helpful as a way of confirming that what interviewees report about the "signals" employed by their algorithms is indeed plausible.[37] Occasionally, too, the wider literature of financial economics (which now stretches back more than fifty years) provides historical evidence. For example, the first unequivocal evidence of the tendency of the share-index futures market to move before the underlying shares do is in Kawaller, Koch, and Koch (1987).

Some particular episodes discussed in this book were covered in the specialist press, and that too usefully complemented the interviews. I visited the archives of the New York Stock Exchange and the records of the Securities and Exchange Commission in the US National Archives, but more productive than either were records of one of the first firms to conduct an early version of what we now would call HFT: Automated Trading Desk, set up in Charleston, South Carolina, in 1989. (I was given access to those records by the firm's cofounder, David Whitcomb, and by another interviewee.) Four individuals important to the episodes discussed in the chapters that follow—Leo Melamed, former chair of the Chicago Mercantile Exchange; Iowa politician Neal Smith; the stockbroker and trader Donald Weeden; and

Dave Cummings, founder of the HFT firm Tradebot—have written autobiographies, and I have occasionally drawn on these too in what follows.[38]

When you are studying a material activity, it helps to get a sense of the physical setting in which it occurs. As noted, HFT is conducted not in trading firms' offices, but in exchanges' computer datacenters. Visits to datacenters are not straightforward to arrange, but I was able to visit two and I walked past others, taking note, for example, of communications infrastructure such as the microwave towers described in chapters 2 and 5. In one sense, I didn't learn much from doing this. Simply walking around inside a datacenter, much less viewing it from outside, doesn't give you much insight into the computations performed there. It was, nevertheless, essential to do it. I was trying to understand a world, so to speak—the material world of HFT, the way it emerged, the conflicts surrounding it, its mundane economics, and so on—and seeing where HFT actually happens was a vitally important part of the process.

Synopses of Chapters

Chapter 2 focuses on the Chicago Mercantile Exchange (CME) and its transformation into one of the world's leading sites of ultrafast automated trading. The chapter describes the deeply conflictual process by which the buying and selling of futures shifted from Chicago's open-outcry trading pits to electronic trading, including, for example, the death threats faced by the latter's most prominent advocate, Leo Melamed. The chapter also tells the story of how the CME achieved its central role in the trading of financial futures, which included crucial interventions in the 1970s in congressional politics and a reworking of the legal boundary between gambling and legitimate futures trading. The chapter ends with a discussion of why price changes in the Chicago share-index futures market tend to lead those in the underlying shares.

Chapter 3 explores the huge HFT-induced transformation of US share trading, viewing that transformation initially through the lens of the early HFT firm Automated Trading Desk. The chapter describes the difficulties faced by the firm in its efforts to trade on pre-HFT share-trading venues. It then shifts to Island, the first "HFT-friendly" share trading venue, and its emergence from what were in the 1990s the disreputable margins of the US financial system. The chapter explores the mutually reinforcing relationship in the US between automated trading and new trading venues such as Island, a relationship that was strengthened, largely inadvertently, by actions taken

by the stock-market regulator, the Securities and Exchange Commission. The chapter ends with brief discussions of the way in which a similar relationship transformed European share trading, and of the "signals" employed by HFT algorithms trading shares in both the US and Europe.

Chapter 4 turns to the trading of sovereign bonds and foreign exchange, where a market structure very different from those that have come to prevail in futures and shares still largely survives. That market structure is organized above all around a distinction in socioeconomic roles, between "dealers" (nowadays, mainly large banks) and "clients," where the latter are not usually private individuals but smaller banks, nonfinancial corporations, and especially institutional investors. Even the world's largest such investors are still usually treated simply as clients in the trading of bonds and currencies. In sovereign bonds, for example, this dealer–client market structure is anchored in the system of government-designated "primary dealers": banks (or sometimes also other securities firms) that commit always to bid in the initial auctions of government bonds and to act continuously as market-makers in their subsequent trading, receiving privileges in return.

Chapter 5 focuses directly on the material technical systems within which HFT takes place. The chapter discusses the cables and wireless links that convey the crucial signals for HFT from one datacenter to another, emphasizing, for example, the huge importance of how closely a link follows the geodesic between datacenters. The chapter then moves to how trading firms' computer systems materially interact, within datacenters, with exchanges' computer systems. Among the phenomena described is how the macro-scale importance of spatial location in fiber-optic and wireless transmission is mirrored in miniature: the designers and programmers of the specialized computer chips (field-programmable gate arrays, or FPGAs) that are now involved in all of the fastest forms of HFT must also pay close attention to where exactly on these chips computations take place.

Chapter 6 focuses on the two main species of HFT algorithm: market-making algorithms that systematically place in exchanges' electronic order books both bids to buy the financial instrument being traded and offers to sell it (at a slightly higher price); and liquidity-taking algorithms that seek to identify opportunities to profit by executing against existing bids or offers. The chapter then turns to the "material politics" efforts by exchanges and other trading venues to alter the interaction of HFT algorithms by deliberately modifying—sometimes overtly, sometimes in lower-profile ways—the material features of the technical systems within which trading takes place. Chapter 7, the conclusion, reviews what the previous chapters have revealed

(continued...)

INDEX