Center for Technology Innovation at BROOKINGS

September 2013

Cyber Theft of Competitive Data:
Asking the Right Questions

Allan A. Friedman

Allan A. Friedman
is a fellow in Governance Studies and research director of the Center for Technology Innovation at the Brookings Institution.

INTRODUCTION

Cybersecurity has captured the attention of policymakers around a range of technical and economic questions, but this convenient term can muddy different threats and challenges. Perhaps the most important and least understood is the question of what some have labeled economic espionage, and we call Cyber Theft of Competitive Data. Properly identifying and clearly understanding the scope of this issue is critical for several reasons. First, it has become central to the US cyber agenda, particularly as the Administration works with the Chinese government. Both Vice President Biden and Treasury Secretary Lew specifically addressed the importance of addressing this issue in the recent Strategic and Economic Dialogues with China.

Second, from a policy perspective, a better understanding of the problem will help promote the urgency of a response from Congress and the White House. At the same time, it will help tamp down some of the aggressive rhetoric that has surrounded this issue. A member of the independent Commission on the Theft of American Intellectual Property has testified that the scope theft of American intellectual property was "comparable to the current annual level of U.S. exports to Asia," and pointed to China as the chief culprit.[1] For their part, the Chinese call the accusations of economic espionage 'unfounded,' and caution that they are 'hyped into something that would overshadow and obstruct' cooperation between the US and China. This issue has dominated headlines for the past year, with bold statements and confusion on all sides.[2] Even the official in charge of defending American cyberspace, Cyber

1. Slade Gorton. Testimony before the House Energy & Commerce Committee; Subcommittee on Oversight and Investigations. July 9, 2013. http://docs.house.gov/meetings/IF/IF02/20130709/101104/HHRG-113-IF02-Wstate-GortonS-20130709-U1.pdf

2. Yang Quingchaun. Commentary: Don't let cyber security overshadow key China-U.S. dialogue. Xinhua News Service. July 9, 2013. http://news.xinhuanet.com/english/indepth/2013-07/09/c_132525189.htm

Command's General Alexander, does not seem above inflammatory rhetoric, calling it "the greatest transfer of wealth in history."[3]

Perhaps most importantly, however, a complete understanding of the mechanics of CTCD will help policy makers understand *how* to respond. If we do not understand the nature and mechanics of the threat, then sizable technical investments and elaborate diplomatic negotiations may fail to actually address problem. Different sectors depend on different types of data to generate value, and face different threats from theft of that data. Only by appreciating the underlying processes that harm the American economy can we explore the efficacy of mitigating solutions.

WHY 'CYBER THEFT OF COMPETITIVE DATA'?

We use the term 'competitive data theft' for several reasons. First, the use of 'espionage' in economic espionage is a politically charged concept. State-sponsored espionage is itself legal, and there are few countries in the world that don't practice it. Therefore, discussions of 'cyber attacks' frequently get mired in definitional dead ends in international dialogue. A commentary from China's official press agency objected to American claims of Chinese cyber attacks out of fairness. Since the US also engages in cyber exploitation, the reasoning went, American claims that its cyber espionage isn't economically motivated just "sounds like an attempt to cover one's old mistake with a new excuse, and a bad one."[4] Indeed, during the recent Strategic & Economic Dialogues between the two countries, American officials seemed careful to avoid the term economic espionage.

To avoid the ambiguity of espionage, we first argue that the research framing should focus on 'competitive data' to distinguish it from strategic data that might serve as the target for more legitimate intelligence activities. Defining competitive data involves looking at its potential use, both from those generating it and those who wish to exploit it. Competitive data is used to support the behavior of commercial entities in a market, and its value depends—at least in part—on its secrecy. That is, not only must the data be assumed to be treated as secret, as per US trade secret laws, but that the owner's successful use of this data relies on other competitive actors not sharing it. Similarly, to qualify this data as competitive data, any other party that might exploit them must be in a market context, although not necessarily in the same context. By focusing on the use of competitive data, we explicitly ignore other types of intellectual property, such as patent licensing or trademark counterfeits that make up large portions of other discussions of IP theft.

3. NSA Chief

4. Yang Qingchuan. "Commentary: Don't let cyber security overshadow key China-U.S. dialogue." *Xinhua News*. http://news.xinhuanet.com/english/indepth/2013-07/09/c_132525189.htm

Second, given the context of cyber theft, the data must be in digital form. What is this data? In public discussions of this problem, many assume it revolves around advanced technology, which would allow the exploiter to replicate the owner's products and compete in global markets in high value industries. While this is undoubtedly a major concern, competitive data theft spans beyond what we might think of as the drivers of high-end innovation. We divide competitive data into two categories: proprietary technology data, and tactical data. Proprietary technology data, under a broad definition of technology can be anything that supports the creation of a good or service. Tactical data, on the other hand, supports company decisions, from pricing information to long-term strategies. Between these two categories, we capture all of the means that market actors can create value from data, while not including other forms of intellectual property that are the products themselves, such as copyrighted works.

Finally, we should address the notion of theft. It is important to acknowledge that theft of information is not the same thing as theft of physical goods, since the legitimate owner usually retains the data even after it is copied.[5] However, absent a better term,[6] we feel that theft puts this in the context of an illegitimate appropriation of something of value. This notion is also constrained to *cyber* theft from cyber intrusions of independent systems. We do not consider competitive data gained from joint ventures, even when the technology transfer violates law or contract.

WHAT ARE THE QUESTIONS?

A natural starting point in examining an emerging policy problem is to assess the magnitude. Estimates to the size of the problem range from the billions all the way up to a trillion dollars. Most estimates are based on very broad assumptions, if there is any concrete data behind them at all. In 2011, the UK Cabinet Office estimated that industrial espionage cost that country alone $12 billion. That year, however, one of the most comprehensive studies of cybercrime led by Cambridge University's Ross Anderson examined the same question and "there is no reliable evidence of the extent or cost of industrial cyber-espionage."[7] In the US, the most systematic efforts have been by the Center for Security and International Studies, who attempted to estimate the harms to the US economy of all cyber crime by comparing it with other major social problems. They were able to bound the estimate between .05% and 1% of national GDP, or between $70 billion and $140 billion annually.[8]

5. Jessica Littman. "Sharing and Stealing" *Hastings Communications and Entertainment Law Journal*, 27 (2004) http://papers.ssrn.com/sol3/papers.cfm?abstract_id=621261

6. http://xkcd.com/1228/

7. Ross Anderson et al. "Measuring the Cost of Cybercrime" *Workshop on the Economics of Information Security* (2012) http://weis2012.econinfosec.org/papers/Anderson_WEIS2012.pdf

8. http://csis.org/files/publication/60396rpt_cybercrime-cost_0713_ph4_0.pdf

Understanding the rough size of the problem may help us understand the importance of the problem, but there are two drawbacks to a focus on size – lack of both data and differentiation. First, we simply don't have the data. Firms have been reluctant to discuss security breaches even after an SEC guidance recommending disclosure, and may not understand how an attack might hurt them. And that's assuming the firm is aware. As cybersecurity researcher and entrepreneur Dmitri Alperovitch says, "I divide the entire set of Fortune Global 2000 firms into two categories: those that know they've been compromised and those that don't yet know."[9]

Indeed, there are many policy issues where one can make estimates with incomplete data. Yet we face a second and more prominent drawback that existing discussions of competitive data theft tend to gloss over: different types of theft can lead to radically different types of harms.

Consider an oil company. An adversary might steal just four bytes of data—thirty-two 1s and 0s—and know how much they are planning to bid on an oil field. Exploitation of this information might cost the firm a huge percentage of their projected revenue for ten years if they were outbid. At the other end of the spectrum, one might think of a leading advertising company. Even if they were to lose every bit of data on every computer, the long-term value of that firm would not suffer significantly. Yes, some clients might be annoyed in the short run, but the value of that firm is tacit, embedded in the creativity, personal skills, and relationships of its employees. Their competitiveness does not depend on their data.

In the middle of this spectrum, we might place a consumer electronics firm. Knowing the details of an upcoming product or some technical advantage would enable a competitor to bring a similar product to market more quickly without the costs of development. The amount of time that firm would have a market lead would shrink. Firms in this sector, however, have probably faced a shrinking market lead window for the past 30 years, as global competition has increased. They have had to compete on other factors—quality, customer service, complementary goods—so competitive data, though still important, does not drive the entire market. These three examples highlight the need to explore competitive data theft in a more nuanced fashion. The focus must be on how firms use competitive data, and the harms that come from its theft.

WHAT ARE THE HARMS?

When thinking about the impact of data theft, the good news is that there are relatively few ways that a firm can suffer from an adversary compromising the confidentiality of competitive data. Unfortunately, these harms can still have a serious impact. The primary mechanism of loss is loss of sales, but this needs to be broken down further. Rather than just considering the impact of market entrance, it is useful to understand whether the primary effect of new

9. Nicole Perlroth. "Some Victims of Online Hacking Edge Into the Light" *New York Times*. February 21, 2013. http://www.nytimes.com/2013/02/21/technology/hacking-victims-edge-into-light.html?pagewanted=all

knowledge enables a new entrant to improve the quality of the final product, or lower the price of their existing capacity. That is, a firm faces a different competitive environment if the capacity had existed but competition would have been cost-prohibitive, as opposed to enabling competition on a new product domain. By distinguishing between these two, we can better understand the importance of how the adversary might adapt and compromised information. The victimized firm might also suffer a loss of sales through poaching, if the adversary compromises sales data and is able to 'steal' customers based on this tactical data without altering their product.

Another common harm to a firm would be the disruption of its strategic plans, due to the loss of other types of tactical data, such as a merger or acquisition. The victim firm may also be denied key inputs for their product, particularly if they lose their private valuation in a competitive auction.

Finally, a firm's product could lose some intrinsic value, even if not competitive product is introduced. The product might depend on exclusivity of information, and any disclosure could reduce that value, such as the compromise of a cryptographic key in a security product. This also includes the economic harms of what might otherwise be seen as strategic espionage. For instance, while the primary harms of a cyber intrusion on a weapons vendor might be the strategic impact on national security—potentially fair game under the norms of national security—the vendor's product is now worth less to many of its customers.

HOW TO APPROACH THE PROBLEM: THE BROOKINGS MODEL OF CYBER THEFT OF COMPETITIVE DATA

Measuring these problems in an absolute fashion across the entire still requires more data than is currently available, and still would not offer explicit guidance about how to address the threat to the American economy. We propose two key features that will enable a general, widely applicable framework for understanding the long-term impact of cyber theft. First, we focus on the impact of CTCD on specific industrial sectors. By abstracting across individual firms, we can avoid the challenge of intra-sector replacement as well as the hurdles of business-specific exceptions. At the same time, focusing on industrial sectors allows us to capture more specific details about the nature and harms of espionage, and how different types of businesses use data differently. By choosing a middle ground between the macro approach studying the entire economy and the micro approach focusing on the enterprise, we can gain insights into the actual mechanisms of CTCD.

Any policy model should contain within it the ability to understand potential solutions. By opening up the black box driving the relationship between competitive data and the harms of its theft, we can explore how to mitigate these harms. We focus on three classes of interventions that might address data theft. The first is technical: how can an enterprise protect the data using security technology? The relative efficacy and costs of different

technical interventions depends a great deal on the nature of the data to be protected. Small amounts of data with relatively few users are much easier to successfully protect than giant databases with a large sets of diverse users. Second, we look at market interventions by the players themselves. How can a firm adapt in a world with less confidentiality, where similar, cheaper products may be available? Xerox, for example, famously preserved the value of its brand through customer service to generate loyalty even as the market surged with similar products. Finally, we explore how legal and diplomatic protections may have an impact to understand where policies such as greater global protection of property rights may—and may not—improve the situation.

Absent any useful, statistically representative quantitative data about cyber theft, we can still glean useful insights into threats and solutions by mapping the relationships between the usage of data and the harms of data theft for specific industries. These mappings are informed by a collection of over fifty public cases of known economic espionage with specific estimates of harms. The methodology for estimated harms varies, allowing us insight into different legal and economic models, and varies from no loss to the hundreds of millions of dollars. We supplement these cases with interviews in the targeted sectors. By focusing on a key subset of most innovative industries, we limit the initial scope to a feasible project while still offering timely insight.

This framework is designed to be extensible and adaptable. With present data, we can begin to understand the relationship between data use and the harms of CTCD, and consequences of different types of protection failures. As we gain more data from public and confidential sources, we will expand on this work to improve its precision and utility. The goal remains to better understand the problem with an eye to understanding who has the most at stake, and what solutions have the best chance of addressing one of the most talked about and potentially critical threats of the digital age.

Governance Studies
The Brookings Institution
1775 Massachusetts Ave., NW
Washington, DC 20036
Tel: 202.797.6090
Fax: 202.797.6144
brookings.edu/governance.aspx

Editor
Christine Jacobs
Beth Stone

Production & Layout
Beth Stone

EMAIL YOUR COMMENTS TO GSCOMMENTS@BROOKINGS.EDU