The Impact of APIs on Firm Performance

Seth G. Benzell, Guillermo Lagarda, Marshall Van Alstyne *

May 21, 2017

Abstract

Do firms benefit from external as well as internal performance enhancements? Using proprietary information from a significant fraction of the API tool provision industry, we explore the impact of API adoption and complementary investments on firm performance. Data include external as well as internal developers. We use a difference in difference approach centered on the date of first API use to show that API adoption – measured both as a binary treatment and as a function of the number of calls and amount of data processed – is related to increased sales, operating income, and decreased costs. It is especially tightly related to increased market value. In our preferred specification, binary API adoption predicts a 10.3% increase in a firms' market value. Creation of API developer portals is associated with decrease in R+D expenditure inside the firm, supporting the hypothesis that outside developers can substitute for internal spending. Categorizing APIs by their orientation, we find that B2B, B2C, and Internal API calls are heterogeneous in their association with financial outcomes. Finally, the fact that API calls are associated with contemporaneous increases in firm value suggest that data flow at the boundary of the firm can predict financial performance.

1 Introduction

Information and communication technologies (ICTs) have long been studied as a form of business investment. Application programming interfaces, or APIs, are a newly popular type of ICT. Combining proprietary data from an API management firm and firm financial data, we study, for the first time, the impact of APIs on firm profits and growth. We also examine the contributions of internal versus external developers.

A major reason for recent interest in APIs is their role as the foundation of digital ecosystems. APIs make it easy for individuals to write programs that communicate with online services and shared databases.

*We thank participants at the Toulouse School of Economics, CMU Heinz School of Public Policy, Platform Strategy Research Symposium, and Workshop on Information Systems Economics for their advice and encouragement. We thank Michael Schrage for multiple clarifying conversations and Pedro Ferreira for valuable tests of robustness. This work would not have been possible without the diligent aid of a team of research assistants, especially Jon Tabernero, Danial Salman, Jessie Wang, Lura Bonney, and Ardavan Sepehr.

Although seemingly banal infrastructure, APIs are essential for making the power of systems such as Google Maps, eBay, Amazon, and Twitter available to independent developers. They mediate economic transactions. Their value is not fully determined by the actions of their creators or the preferences of their consumers – also critical are the strategic choices of third parties who connect across systems and reuse components in unanticipated innovations.

Making these tools and data available to outsiders can be a win-win. Independent developers gain new opportunities to increase agility and speed of deployment, while the API providing firm gains complementary added value. In the best case scenario for the platform firm, an API becomes the basis of a thriving ecosystem, with exponentially growing revenues and low marginal costs.

In this paper, we investigate the effect API development has on the developing firm itself. In the first section of the paper, we study the characteristics of API using firms, and trends in their use of APIs. We track several measures of API usage, including calls, data flow in bytes, number of developers, and whether an open development portal exists. In the next section we focus on the binary impact of having at least one functioning API. We seek to understand the causal impact of these types of usage on firm outcomes, such as sales, market capitalization, and growth. We conclude by looking at the relationship between different dimensions of API quality and firm outcomes.

Our primary econometric strategy is difference-in-difference centered on the moment in time when a firm first opens its APIs. API adoption is highly endogenous, with, for example, firms in certain industries being much more likely to adopt. However, a difference-in-difference approach allows us to bypass most of these concerns. We use two types of control groups. Our primary control group is firms that adopt APIs at a different point in time. Our secondary control group is a matched sample of firms that are ex-ante similar to API adopters, but which never adopt. By comparing firms' projected outcome before API adoption to their actual outcomes, we extract the impact of API use.

We find that API adopters see large gains in financial outcomes after adoption. Further, the effect of API use is increasing in the amount of data passing through the API. The orientation of an API is important to its effects as well. Internal and B2B oriented APIs are related to different outcomes than B2C ones.

Finally, we estimate a relationship between the number of issued developer keys and API use, as well as a relationship between whether a firm has a developer portal and financial outcomes. We can say less about a causal relationship in these cases. However, we speculate about economic phenomena that could drive the observed relationships. We look forward to further investigating these questions.

In the following section we review the history and role of APIs as well as related literature. In Section 3, we list hypotheses about the impact of API adoption. In section 4, we discuss the data we use in our study. Section 5 presents our analysis. Section 6 concludes.

2 API Overview and Literature Review

An API, or application programming interface, is a set of routines, protocols, and tools that standardizes building software applications compatible with an associated program or database. APIs are code. They are also contracts (Jacobson et al., 2011). They govern the type and format of calls, or communications, that any given application can make of another associated program. The associated program is agnostic about the source of the call, and the app need not know anything about the internal workings of the associated program.

APIs offer the dual virtues of practical modular design and precise metering. Modular architecture, enabled by APIs, allows designers to independently create, subdivide, modify, and remove components without affecting other parts of a larger system. It also facilitates partitioning of decision rights (Tiwana et al., 2010) Modularity combines the advantages of standardization typically associated with high volume processes with the advantages of customization typically associated with bespoke processes (Baldwin and Clark, 2000). APIs also enable precise metering of access permissions to these key resources. Metered access permissions ensure that anyone and anything that consumes system resources adheres to technical and economic policies designed to ensure system health. As architecture, APIs provide infrastructure for building platforms. As regulators, APIs partition decision rights and provide scalable mechanisms for governing behavior. These dual roles – architecture and governance – provide the foundations for building platforms (Parker et al., 2016).

It is not clear when the first API was created, but they clearly predate the internet. Google's ngram tool lists usage of the phrase 'application programming interface' as early as 1961. Usage increased rapidly in the early 90's.

Many web-pioneers featured APIs as core to their business. Salesforce.com included them in their 2000 launch of the world's first 'software-as-a-service' product. Likewise, eBay launched a developer program in 2000 to a select group of partners, encouraging them to create services that drew information from eBay's API. Having created of one of the first popular 'open APIs,' eBay's decision led to a virtuous cycle of better tools, higher visibility, and more customers.

Perhaps the most iconic effort to place APIs at the center of a firm's strategy was Bezos' famous 'Big Mandate' of 2002. Frustrated by the haphazard way Amazon solved its digital challenges, and hoping to turn hard won lessons into new sources of revenue, he demanded, among other things, that:

- All teams will henceforth expose their data and functionality through service interfaces. . .
- There will be no other form of interprocess communication allowed: no direct linking, no direct reads

of another team's data store, no shared-memory model, no back-doors whatsoever. The only communication allowed is via service interface calls over the network.

- All service interfaces, without exception, must be designed from the ground up to be externalizable. That is to say, the team must plan and design to be able to expose the interface to developers in the outside world. No exceptions. (Rowan, 2011, citing Yegge)

In the years since, Amazon has exploded from a marginal seller of bound paper, into a powerhouse that sells all sorts of goods and services. A greatly reduced share of all products on its website are actually owned by Amazon (yet it manages to own many of the high volume items). In 2013, Amazon's marketplace featured more than 2 million third party sellers, accounting for roughly 40% of total sales. Using partner sales data, Amazon has also moved to vertically integrate into 3% of its partners' top selling products (Zhu and Liu, 2016). Amazon's market capitalization has duly expanded. Arguably, Bezos was right: there's more money in managing bytes than managing books.

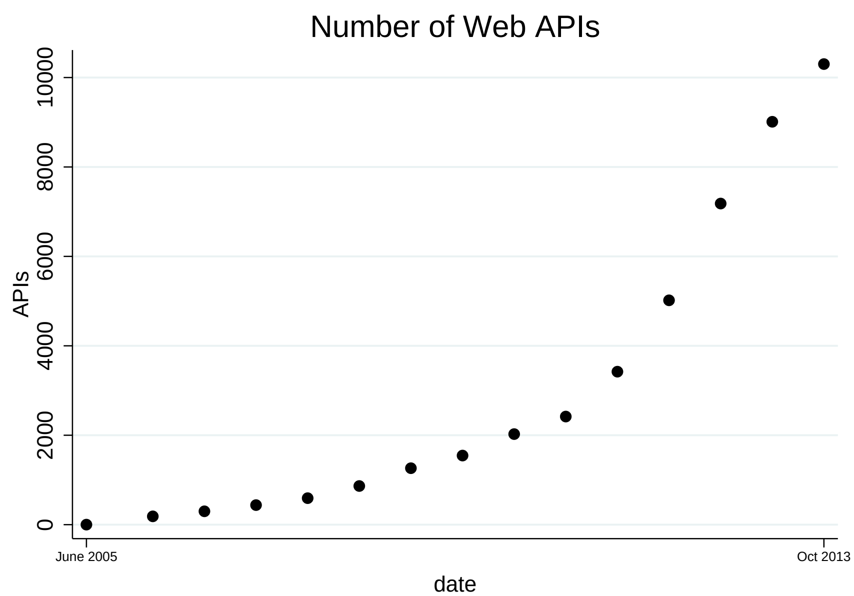


Figure 1: Number of Web APIs over time (source: Programmable Web (2016), Authors' figure)

Driven by Amazon's inspiration, believers have made large investments in APIs. According to Programmable Web, the number of Web APIs has increased from a few hundred in 2005 to over ten thousand today (See Figure 1).

Large, technologically savvy, tech firms create and maintain their own APIs using a variety of programming languages. Smaller firms, and ones for whom APIs are not a core competency, are more likely to seek the help of API management firms. These firms help companies design and implement APIs to manage

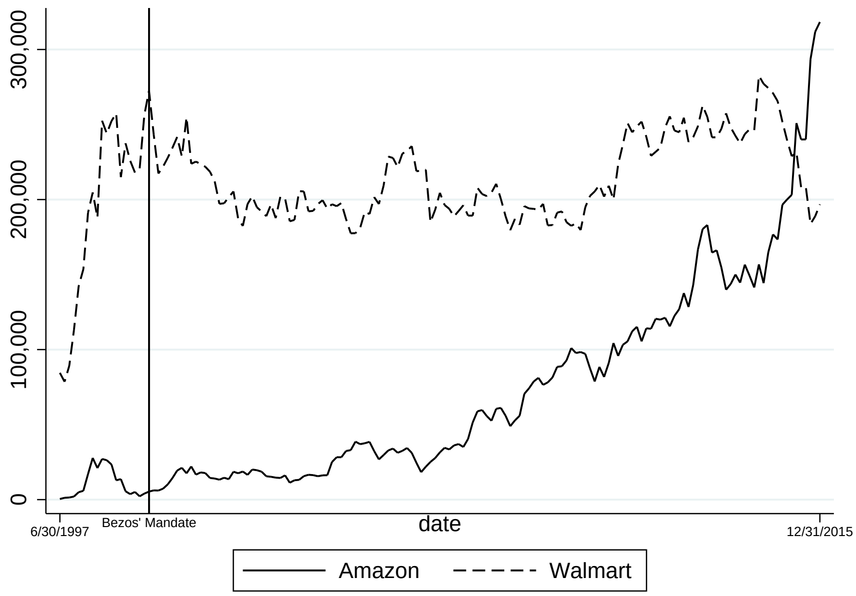


Figure 2: End of quarter market capitalization in millions. Bezos' Mandate is approximately dated to Q1 2002.

internal and external data flows and communications. Some implement API proxies or cloud services. An API proxy is an intermediary, controlled by the API management company, through which calls to the customer's API pass. These API proxies can provide additional functionality, including security (against hackers or DDOS attacks) or analytic tools. Some API management firms have designed computer languages and developer tools for creating APIs. One example is RAML (RESTful API Modeling Language), which Mulesoft has released under an open source license to encourage adoption.

API boosters make several claims about their technical usefulness. They claim that APIs lower barriers to entry for programmers, and make designing complementary programs easier, faster, and less disruptive to existing business logic. Other claimed benefits include better internal information diffusion and management; higher function and lower cost data management and security; and the ability to serve as the foundation of a digital platform, by making it easier for outsiders to participate in and add value to the ecosystem. The key to all of these is that APIs don't directly lead to the consequences. Rather, they make these impactful investments cheaper. Any firm writing their own code will likely find it useful, if not immediately necessary, to create associated APIs.

APIs come in two main flavors: open and closed. This distinction depends on whether access is outward facing (open) or inward facing (closed). This decision should impact the firm on several levels. A recent survey argues that open innovation occurs across inter-firm, firm, and intra-firm levels Bogers et al. (2016). Our research examines openness at a fine grain level, access to individual byte transfers, facilitating analysis across these levels. We can classify API access as B2B (upstream), internal, and B2C (downstream). Thus performance differences can be analyzed for internal or "company-specific" platforms versus "industry-wide" platforms Gawer and Cusumano (2014) or even the upstream and downstream ends of the value chain Adner and Kapoor (2010).

A closed API is only accessible to individuals working for the firm, or close associates. The bulk of APIs are closed (Jacobson et al., 2011). This proprietary usage is designed to enhance internal agility. A firm that includes well designed APIs in its internal software will likely find it easier to modify or create new programs using the same information . For example, Amazon Web Services began as an internal project to reduce duplication of effort (Huckman et al., 2012). Verizon has employees use an app in the course of on-site service calls and to activate new cable lines. If the firm wants to enable employees to perform the same tasks on a new device, it will only need to create a new app to talk to the API. Similarly, the company could easily add additional functionality to a current app by having it communicate with a new API. An app that communicates to multiple APIs is called a mashup.

While there is significant research on the economics of digital platforms, there is much less on APIs themselves. To our knowledge, this is the first paper statistically investigating the impact of API usage on firms. However, APIs can be thought of as a type of ICT. There is a very large literature on the impact of ICTs on firms. A good summary is Brynjolfsson and Hitt (2000). Brnjolfsson and Hitt report productivity impacts from ICT no matter how measured – but mostly at the firm level rather than in aggregate. They emphasize the importance of taking into account complementary investments – both organizational and in software.

One type of ICT that has received particular attention are electronic data interchanges (EDIs). EDIs are systems by which firms or organizations exchange information within or between each other electronically. They arose simultaneously with the rise of the internet in the 1980s. In papers such as Mukhopadhyay et al. (1995), they are estimated to create large cost savings. APIs are capable of serving in the traditional roles of EDIs in a cost effective manner. APIs used in this role would not necessarily require an open developer portal, as internal developers would create the complementary apps.

By contrast, open APIs allow third party access to internal resources. The idea behind open APIs is to build the foundation for an ecosystem. By making an additional investment in creating a developer portal and making API information available (to either a select set of partners given 'keys' or the general public) a firm can enable others to make complementary products. For example, Walgreens opened an API for photo printing services at its pharmacies. Many developers duly incorporated Walgreens printing into their apps, helping users to print photos from their phones, social networks, and cloud accounts (Iyer and Subramanian, 2015). For a modest investment, Walgreens enabled others to promote its photo services and drive in-store traffic.

Running a product ecosystem brings challenges of its own – including questions around pricing, network effects, and curation.[1] However, it is a tremendous opportunity if third parties do participate. The market maker typically holds a low cost, high margin position. For example, when an additional used book is sold on Amazon, the company faces almost no cost, but captures a share of the sale price.

Basing an ecosystem around an API makes it straightforward for a firm to scale and expand. Therefore, the potential for growth through complementary network effects is enormous. The benefits of strategies around ecosystem management are explored in Parker et al. (2016). Firms successfully adopting platform strategies are either midsize publicly non publicly traded or with unusually large market capitalization to employment ratios.

Another way to think about API investments is as the crystallized output of programmers that will make future coding easier. Measuring the productivity of programmers is a difficult task, but some preliminary evidence exists for the presence of such coding spillovers. Spillovers are when the productivity of programmers at a firm increase with the firm's scale or over time. This hypothesized phenomenon could be driven by code accumulation or by programmers 'learning by doing'.

For example, microeconomic theory (Porter and Stern, 2000; Parker and Van Alstyne, 2016) argues for the importance of code spillovers. A recent study of SourceForge also offers evidence that spillovers exist (Eilhard and Ménière, 2009). Benzell et al. (2016) argues for the macroeconomic importance of these spillovers. However, many empirical questions about digital spillovers remain. There is reason to believe that within-firm code spillovers count towards a firm's intangible assets. Therefore, any effect we measure of API usage on intangible assets likely corresponds in part to this phenomenon.

3 Hypotheses – Should APIs Matter?

In the previous section we described the roles that APIs play in firms. Their purported usefulness in these roles makes it natural to hypothesize that adoption of an API strategy has positive consequences for firms that are able to do so.

Firms adopting an API strategy may see immediate increases in profitability. For example, APIs might make it easier for a firm to sell or market its existing products through complementary apps. Additionally, the tighter control an API gives firms over outside code access (Tiwana et al., 2010) could help firms price discriminate with existing digital products or sell new ones. Productivity (we use the term productivity to mean an increase in output for a certain level of inputs) is measured by a firms' net income and operating

[1] For a discussion of some of the considerations involved, see Boudreau and Hagiu (2009)

income.

While there may be immediate benefits to adopting an API strategy, most API driven boosts to productivity may only become apparent after a lag. As discussed above, this is because a main function of APIs is to increase the speed and productivity of programmers. APIs are more modular than traditional code, potentially increasing data and software access, reuse, and recombination (Yoo et al., 2012) (Baldwin and Clark, 2000). Therefore APIs should help firms get more effective labor from a certain expenditure on programmers. If programmers create investments that give benefits over time, then APIs should have benefits that only manifest over time as well.

APIs may also facilitate the networking of disconnected pockets of expertise (Purvis et al., 2001), integration of new software into legacy software (Joseph et al., 2016) and speed IT deployment (Iyer and Subramanian, 2015). All of these may lead a firm to become more innovative, helping it over time to develop new products and reduce inefficiencies internally. Perhaps most importantly, having an API helps a firm develop as a platform. Outside partners will be more likely to develop complementary products. Opening an ecosystem, through more permissive licensing (which may be enabled by APIs), has been shown to increase complementary device development among handset manufacturers (Boudreau, 2010). Moving the boundaries of the firm based on openness can also affect performance (Adner and Kapoor, 2010).

Measuring the longer term impact of API adoption is complicated. A firm's market capitalization after adoption should capture anticipated long term productivity gains. However, some of this impact may be priced in before adoption due to a firm being likely to adopt. Any estimate of the impact of API adoption on market cap is therefore likely to be underestimated.

Another way API driven productivity gains may manifest in firm level financial data is through intangible assets. Intangible assets are another measure of a firms' 'q' or super normal return. So investments in API development are likely to be recorded as intangible assets. In previous studies, each \$1 invested in ICT hardware has been shown to correlate with \$10 in market capitalization. This is believed to represent capitalized hidden assets in the form of internally developed software, IT training, and organizational process knowledge (Saunders and Brynjolfsson, 2016). In the present case, an increase in intangible assets due to API adoption might represent anticipated future growth or third party investment. However, measures of the impact of API adoption on intangible assets are likely to be underestimated for the same reason that the impact on market cap is. To a certain extent, these problems may be balanced out by firms that are likely to adopt APIs, or have the potential to, not adopting (at least in a certain period).

There is also the issue that announcement of the adoption of an API strategy may lead or lag our measure of an API's first implementation. This could also confound our estimates.

Formally, we seek to test whether

- **H1** API strategy adoption, as proxied by a firm possessing at least one observed API, increases short term firm profitability, as measured by net income and operating revenue

- **H2** API strategy adoption, as proxied by a firm possessing at least one observed API, increases long term profitability, as measured by market value and intangible assets

We are also interested in measuring the correlation between the intensity of API usage and revenue and cost outcomes. While ideally we would like to measure the causal impact of more intense API usage, identifying this effect is more difficult than measuring the binary treatment of API adoption. Therefore, we focus on correlations.

We hypothesize that the relationship between API intensity and productivity outcomes is a composite of a constant direct marginal effect and an indirect effect. The latter would be an effect conditional on the total amount of time that the firm has been using an API and the degree of usage. Time dependency could be associated with a learning process where firms become better at exploiting their APIs.

Intensity of API use might be linked to financial outcomes in a non-linear way. For example, API usage may face decreasing returns, have a logarithmic impact, or have an impact proportional to the intensity of API usage.

We observe API usage intensity along several dimensions. One is the amount of data and calls processed by a firm's APIs. Another is the number of APIs a firm has. Firms that have used APIs for a longer period may also 'learn by doing' and see ongoing increases in productivity. Finally, firms that have invested in a developer portal may see increased gains through better developing their product ecosystem.

One relationship we focus on is intensity of API use and firm financial outcomes as a function of the amount of time since first adoption. Firms may be observed using APIs less intensely over time. This could be due to learning by doing as firms phase out unnecessary data usage. While it is ambiguous whether time-since-first-adoption should be thought of as an intensity measure, such learning-by-doing is one indication it should be treated as such.

An API with an open developer portal allows outsiders to use the API as a resource in their own projects. This can shift the locus of value creating activity from inside to outside the firm (Parker et al., 2017). In the short term, opening up an API to more and more external developers creates additional costs. However, in the long run the return can be large, insofar as it turns the API into a digital platform. Outsiders would then create complements which boosting value.

We acknowledge that these measures have problems: a more sophisticated API will deliver more business value per call or unit of data than a less sophisticated call. While we do not directly observe the quality of a call or a developer portal, we assume that firms can adopt and update APIs easily through the API management provider, which will likely homogenize their effects.

Formally, we seek to test whether

- **H3** API associated productivity improvements are proportional to the intensity of API utilization.
- **H4.a** API usage is increasing in the number of developers and whether a developer portal exists.
- **H4.b** API associated productivity improvements are increasing the number of developers and whether a developer portal exists.

Finally, we are interested in how the type of APIs used influence their relationship to productivity. We split APIs into different functional types.

Some APIs are clearly customer oriented (e.g. they power online sales), some are clearly internally focused (e.g. they facilitate internal communications), while others are clearly supplier oriented or handle other business to business purposes (e.g. they make restocking easier and cheaper). B2B and internal APIs seem similar to traditional EDIs in their role, and so seem likely to reduce costs. Meanwhile B2C APIs seem more directly related to sales. For each type of call we investigate the relationship between the amount of these calls, sales, and costs.

We hypothesize we will find the following relationships

- **H5.a** Consumer facing APIs are more likely to increase productivity through increased revenue
- **H5.b** Internally oriented APIs are more likely to increase productivity through reduced costs.
- **H5.c** B2B APIs are more likely to increase productivity by jointly reducing costs and increasing revenues

To test our causal hypotheses, we use difference-in-difference estimation to compare firms that adopt APIs to those that do not. This approach controls for bias that, due to an overall time trend in productivity following the great recession, might create false positives. It does not control, however, for the fact that API adopting firms are more likely to expand due to a latent variable that governs both the likelihood to adopt an API and future profits.

To account for this possibility, we we use an approach where identification is based on the timing of API adoption. We run difference-in-difference regressions for API adopters, using other API adopting firms (who have yet to adopt) as the control group.

4 Data Description

Our analysis requires matching firm level financial data with API usage. Our API usage data comes from a major API management firm. Companies which partner with this firm use its tools to create APIs. The API management firm also helps partners regulate calls, data flows, and developer keys. The API management firms have provided us with records of these stocks (granted development keys and whether a developer portals exists) and flows (number of calls and amount of data delivered). This data was supplemented research into the purpose of these APIs, and whether the firm had a developer portal.

For each company, we observe every API in use during a month. For each of these APIs, we observe the API's name, the number of calls made to the API that month and the amount of data delivered.[2]. At the firm level we observe the amount of internal and external developer keys granted prior to a given month. Developer keys are judged as internal or external based on the associated email address. All of this data is available at the monthly level. While this API management company operated prior to 2013, 2013 is the first year for which data is available. For our analysis, therefore, we do not consider firms that were first observed in January 2013 using APIs as being treated in that period. The final month of data is September 2016.

It is important to note that our measures of API usage are imperfect on several dimensions. There are cases of companies which had privately developed API systems in place before they collaborated with our API management firm. Private development of APIs, and development with the assistance of other API management companies, also means that our control group of never treated firms may be contaminated by some API users. We have no solution for this problem, except to note that it is unlikely that a firm with an ex-ante satisfactory API situation would contract with the API management firm which is the source of our data.

API Data

Table 1 summarizes our API usage data. In the first year, we observe 58 firms using APIs. The average firm in 2013 processes 84 billion bytes of data per month and less than 10 million calls. It has about 7 APIs in any given month, and has granted about 10 developer keys. Less than one of these was assigned to an email address linked to the firm. In total, firms in 2013 processed 60 million calls and about 15 billion bytes of data.

In 2016 through September the average firm had 91 billion bytes of API data per month and about 20 million calls. It has about 27 APIs in any given month, and has granted about 15 developer keys. About four of these are assigned to an email address linked to the firm. In total, firms in 2016 (through September) processed 800 million calls and 64 trillion bytes (10 terabytes) of data.

The average API using firm in 2016 has increased calls, bytes of data flow, APIs and number of developers

[2] We also observe whether the firm lists the API as being in a testing or live environment. However, this information is unreliable as many API users do not correctly label their test and production environments. In the body of the paper we do not distinguish by APIs on this basis.

Year	Statistic	Tot. Calls	Avg. Calls	Avg. Data	Tot. Data	Avg. APIs	Avg. Tot. Dev.	Avg. Int. Dev.
2013	N	58						
	Mean or Total	0.06	0.00	84.08	1492.25	6.50	10.12	0.65
	Std. Dev.		0.01	322.80		7.12	42.33	1.51
	Min		0.00	0.00		0.08	0.00	0.00
	Max		0.07	2201.39		38.83	311.80	9.57
2014	N	120						
	Mean or Total	1.04	0.04	368.01	8990.47	20.34	10.78	1.93
	Std. Dev		0.14	1566.36		27.80	28.20	4.74
	Min		0.00	0.00		1.00	0.00	0.00
	Max		1.04	14522.34		177.58	181.26	39.00
2015	N	168						
	Mean or Total	1.63	0.06	574.44	18971.13	31.79	10.27	2.52
	Std. Dev		0.26	3792.22		54.96	27.91	6.94
	Min		0.00	0.00		1.00	0.00	0.00
	Max		2.90	44047.49		389.58	211.69	78.08
2016	N	239						
(through	Mean or Total	0.80	0.02	91.00	6406.95	26.65	14.99	4.16
September)	Std. Dev		0.08	503.17		50.83	42.61	11.39
	Min		0.00	0.00		1.00	0.00	0.00
	Max		0.79	5730.18		429.58	348.99	147.92

Table 1: Usage of APIs in different years. Tot. Calls and Data indicate total amounts in our sample. Other variables are per firm-quarter averages. Calls are in billions and data is in billions of bytes (gigabytes). N indicates the number of firms observed using APIs in that year.

versus the average firm in 2013. This increase in the amount of calls, data and APIs per firm could be due to changes at the intensive (i.e. old adopters use their APIs more intensely over time) or extensive (i.e. late adopters tend to be more intense API users) margins.

Table 2 uses the same data but organizes it in terms of how many quarters post adoption the firm is. The average firm processes 456 gigabytes of data in the first quarter of adoption. This increases to 8.6 terabytes of data in the sixth quarter after adoption. After six quarters post-adoption, calls per quarter remain relatively stable while data per quarter declines. However, the number of firm-quarters of API use we observe more than 8 quarters after first adoption is low. So any trends that far out should be treated more cautiously, and may be due to the fact that smaller firms were the earliest adopters. Simultaneously, average calls per quarter increase from 20 million to 300 million. We can safely conclude that the increase in average API usage intensity from 2013 through 2016 is driven by increased use within firms.

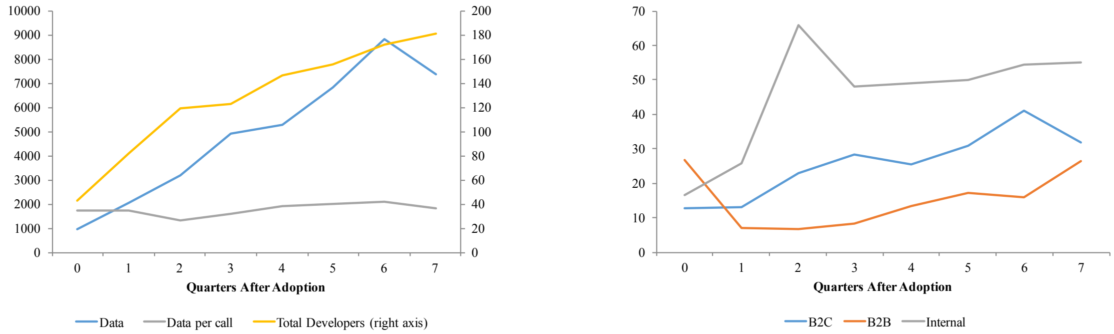
Figures 1.a and 1.b display trends in API usage graphically. These figures are restricted to the subset of firms with 8 full quarters of post-adoption data. Figure 9 in the appendix recreates this figure for all firms. The figure on the right reports the average data use per firm and per call. It shows that in the quarters after adoption data flow per firm and per call increases dramatically. The amount of developer keys issued by the firm increase over time as well, but this is unsurprising as developer keys are rarely rescinded. The amount of data per call is constant over time.

Figure 1.b shows how the total number of calls by type varies as a function of periods since adoption. Firms increase they amount of internal and B2C calls over time. However, the number of B2B calls is

Quarters Post-Adopt	Variable	Mean	Std. Dev.	Min	Max	Quarters Post-Adopt	Variable	Mean	Std. Dev.	Min	Max
Q0	N	244.66				Q7	N	202			
	Data	456.55	3290.38	0.00	41566.51		Data	8284.09	32649.59	0.00	266294.60
	Calls	0.02	0.14	0.00	1.83		Calls	0.32	1.02	0.00	5.44
	Developers	6.48	38.52	0.00	499		Developers	25.37	105.95	0.00	1193
	Internal Developers	1.55	7.41	0.00	105		Internal Developers	6.15	20.50	0.00	213
Q1	N	355.66				Q8	N	172.66			
	Data	1614.11	10093.21	0.00	98843.62		Data	4828.83	12091.20	0.00	63017.62
	Calls	0.04	0.21	0.00	2.11		Calls	0.30	1.17	0.00	8.50
	Developers	10.11	64.87	0.00	1559		Developers	27.65	116.20	0.00	1308
	Internal Developers	2.18	8.16	0.00	117		Internal Developers	6.62	20.59	0.00	212
Q2	N	341.33				Q9	N	117.66			
	Data	4557.99	33064.96	0.00	400426.00		Data	3645.03	8225.85	0.00	42259.04
	Calls	0.11	0.73	0.00	7.67		Calls	0.44	1.85	0.00	12.30
	Developers	15.31	82.49	0.00	1041		Developers	25.90	127.37	0.00	1340
	Internal Developers	2.95	10.60	0.00	152		Internal Developers	5.24	12.81	0.00	127
Q3	N	319				Q10	N	78.66			
	Data	6374.74	35627.66	0.00	359383.70		Data	2280.52	3627.94	0.01	13775.23
	Calls	0.12	0.70	0.00	6.04		Calls	0.55	2.21	0.00	12.87
	Developers	19.09	105.14	0.00	1501		Developers	27.61	157.21	0.00	1442
	Internal Developers	3.33	11.35	0.00	132		Internal Developers	3.83	6.74	0.00	29
Q4	N	289.33				Q11	N	50			
	Data	7254.10	36628.00	0.00	334036.10		Data	1967.99	2576.31	0.32	9799.02
	Calls	0.14	0.89	0.00	9.44		Calls	0.27	0.71	0.00	3.17
	Developers	23.34	109.20	0.00	1396		Developers	29.04	159.97	0.00	1405
	Internal Developers	4.29	13.81	0.00	170		Internal Developers	3.42	6.69	0.00	30
Q5	N	260.66				Q12	N	31			
	Data	7394.80	33375.84	0.00	295625.50		Data	1582.06	2155.99	0.22	7758.82
	Calls	0.21	1.16	0.00	11.56		Calls	0.10	0.17	0.00	0.66
	Developers	24.34	110.50	0.00	1454		Developers	13.73	26.76	0.00	158
	Internal Developers	4.88	15.89	0.00	178		Internal Developers	3.91	7.77	0.00	32
Q6	N	222.66				Q13	N	15.33			
	Data	8693.11	44361.73	0.00	410781.70		Data	1061.02	1504.88	0.00	4631.16
	Calls	0.30	1.28	0.00	10.53		Calls	0.04	0.08	0.00	0.19
	Developers	23.16	97.60	0.00	1185		Developers	13.56	25.19	0.00	106
	Internal Developers	5.06	17.85	0.00	189		Internal Developers	4.17	8.61	0.00	37

Table 2: Quarterly API Usage as a Function of Quarters Since Adoption. N is the number of firm-quarters observed that number of quarters after adoption. The number can be a fraction, because data is at the month level and accumulated up to the quarter shown. Calls and bytes are in billions.

roughly constant. Perhaps this trend is due to firms debugging their inter-firm communication technology eliminating unnecessary B2B calls.



(a) Data and Developers

(b) Calls per Type

Figure 3: Average amount of data flow through APIs and calls by quarters since adoption. In (a) data per firm is in billions of bytes (gigabytes), while data per call is in thousands of bytes (megabytes). A small subset of calls are not assigned to any of these categories. B2B is on the left axis, others on the right axis. This charts for the complete sample are shown in the Appendix. Restricted to the subset of firms with 8 full quarters of API use data.

We also determine when firms first opened a developer portal and categorize APIs by function and orientation. Determining whether a firm had a developer portal was relatively straightforward. The Wayback Machine at the Internet Archive was used to determine as precisely as possible the earliest quarter this portal was available.

Determining the orientation and functions of APIs was more difficult. To do so we primarily relied on the name of the API. We categorize all APIs in our data with more than 100,000 lifetime calls.

The orientation of the API indicates whether the API was meant to be interacted with by consumers, internal employees and systems, or other businesses. For example, APIs with names such as ‘checkout’ were categorized as B2C APIs, APIs with names indicating logistical or technical functions were listed as internal APIs, and APIs with names that indicated partnerships with other companies were judged to be B2B. Internet searches on the name of the API and firm using the API were often helpful. The function of an API is its role. For example, a ‘checkout’ API would be listed as having a sales function. For a full list of the categories employed and their definitions, see appendix A.

B2C APIs generally correspond to online sales, customer loyalty programs, and other systems that make firm information available to consumer apps. Internal APIs relate to logistics, management, internal communications, or serve some internal tech support role. B2B APIs facilitate sales or logistical arrangements with other businesses. A small subset of APIs calls are peer to peer, mixed, or unclassified due to limited use.

API usage varies according to the industry. Figures 4 and 5 show API usage by purpose and industry.

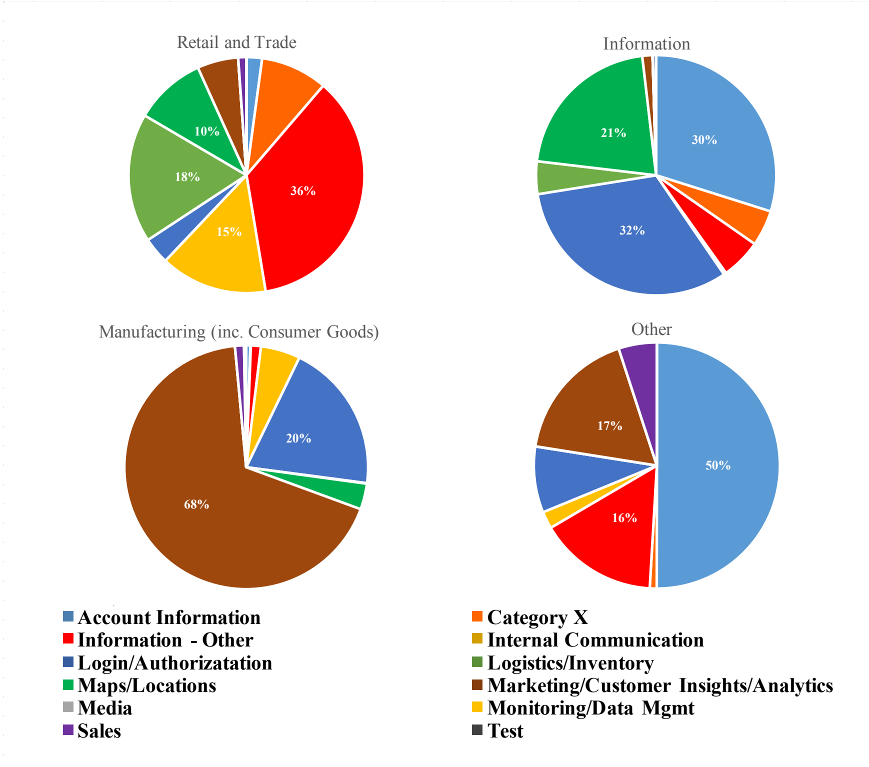


Figure 4: Share of API Calls by Function and Industry

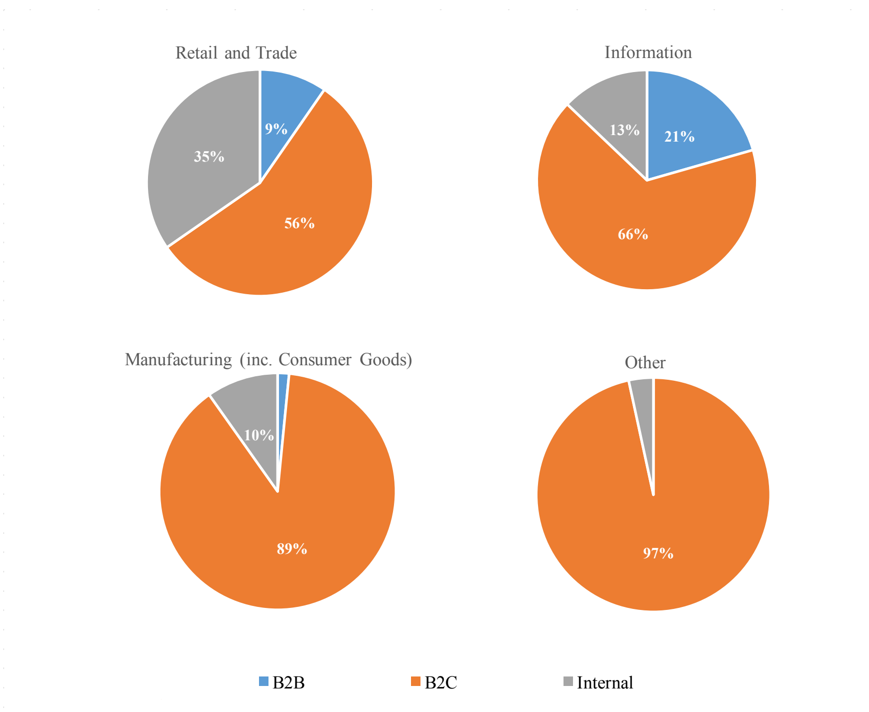


Figure 5: Share of API Calls by Orientation and Industry

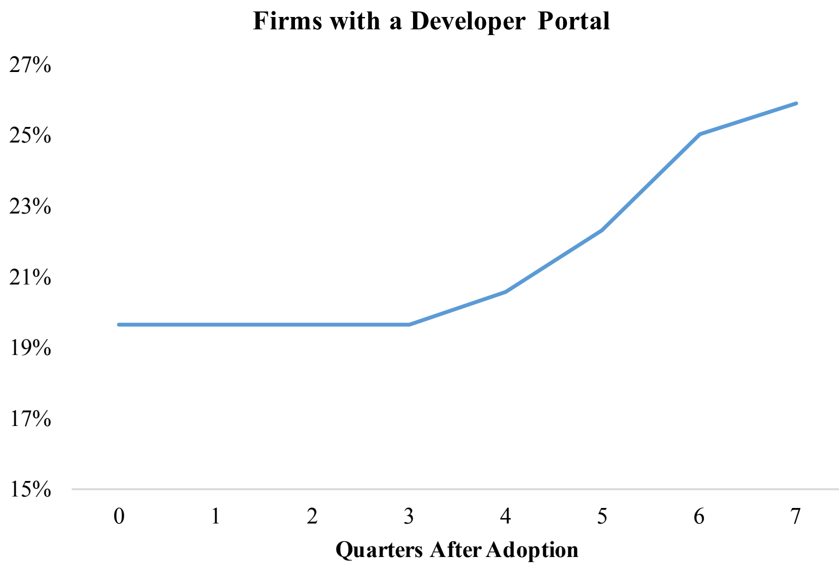


Figure 6: Share of firms with a developer portal by quarters since adoption. Sample restricted to the firms with at least 8 quarters of post-adoption data.

APIs in the retail and trade industry are primarily used for various sorts logistical and data management purposes. In the information industry, they are primarily used for account management and logins. In other industries, a large share is devoted to account information as well, with a sizable amount dedicated directly to sales. Perhaps the most surprising result is how common customer analytics and engagement APIs are in the manufacturing (including consumer goods) industry. This result is driven by a number of brand name clothing and accessory companies in our data which have many calls of this type.

Figure 6 lists the share of firms who have a developer portal available by periods since adoption. Most firms that eventually set up a developer portal already have one available during their first adoption quarter. However, the share of firms that have a developer portal is still low overall. On average in our post-adoption data, firms have developer portals in 16.2 percent of quarters.

Financial Data

We match the above information on API usage to firm financial records. Our main sources of financial data are Orbis and Compustat. To supplement this we source additional financial information from Hoover's and other sources.

Orbis, a product of Bureau Van Dijk, is a well-regarded data set on the financial information of public and private firms in every country. Public firms regularly report - usually quarterly- financial statements in order to be listed on a stock market. Non-publicly traded firms are not required to do so, but voluntary reports from some of these firms are recorded as well. While this data has many gaps, it is the most wide ranging data on firm financial information available.

Occasionally, we supplemented gaps in the Orbis data set with data from Hoover's, another source of firm financial data. In some other cases we searched websites and online public financial statements to capture as much data as possible. Ultimately, we were able to match financial information for 132 firms, about half of the firms or which we have API usage data. It is due to these gaps that we have more data on firm API use than can be used in our regression analysis.

Table 3 gives financial information for firms which at some point work with our API management firm.[3] We are able to match 132 API using firms to financial data. Firms adopting APIs are large on average, with an average market value of about 2.3 billion dollars in 2013. The variance of firm values is large as well. The smallest API adopting firm had a value of less than 10 million while the largest has a value of over 180 billion.

In 2013 the average API adopting firm had on average 890 million dollars in sales 115 million in net income, and 689 million dollars in production costs. In 2015, the final year for which complete data is available, they had on average 848 million dollars in sales, 656 million dollars in cost of production, and 78 million in net income. In 2016, on average, the firms' valuations were 2.5 billion dollars - approximately 10 percent higher than in 2013.

This divergence between strong growth in API using firms' market value and lackluster to stagnant growth in sales indicates two possibilities. The first is that the majority of the benefits of API use come through decreased costs rather than increased sales. The second is that API use is associated with a debt-driven increase to firm size that increases firm valuation (perhaps through increased anticipated revenues) but not short term profits. We explore these possibilities in the analysis.

Table 4 gives additional information on the types of firms which have adopted APIs by year. In 2015 adopting firms are concentrated in the Information (35%), Retail (22%), and Manufacturing (19%) industries. Overall, they are mature firms with a modest growth in market cap (.3%). They have on average about 13.6 thousand employees in 2015. Spatially they are concentrated in North America.

R&D expenses correspond to all costs that relate to the development of new products or services. Market share is calculated by dividing the sales of a firm by the sales of all firms in that industry using Compustat NA data.

[3] Table 14 in the appendix reports financial variables for firms that are not part of the API-adopting subset.

Table 3: Financial info for firms which adopt APIs. Table shows the quarterly means at every year. For an annual approximation the reader can multiply times 4 each value. Sales, cost of production, net income, operating revenues, R&D expenses and gross profits in millions. Market value and intangible assets are averages, in millions. Market share is a percentage. R&D is quarterly.

Year	Variable	Firms	Mean	Std.Dev	Min	Max
2013	Sales	132	890.87	7070.07	0.00	59276.63
	Cost of Production	132	689.34	5458.81	0.00	46771.11
	Gross Profit	132	427.60	3334.87	-0.02	27507.28
	Operating Revenues	132	129.43	1083.06	-0.52	10163.59
	Net Income	132	115.26	920.64	-1.23	8049.54
	Market Value	132	2333.02	17891.05	0.00	150016.00
	Intangible Assets	132	68.86	471.49	0.00	4076.95
	R&D Expenses	34	210.02	527.84	0.00	2190.75
	Market Share	39	0.06	0.08	0.00	0.30
2014	Sales	132	795.05	6339.12	0.00	53675.33
	Cost of Production	132	643.73	5080.12	0.00	42764.86
	Gross Profit	132	378.85	2967.31	0.00	25619.48
	Operating Revenues	132	53.59	594.50	-2.14	8488.80
	Net Income	132	89.63	729.05	-2.23	7484.68
	Market Value	132	2650.48	20450.94	0.00	168088.20
	Intangible Assets	132	73.19	507.71	0.00	4684.75
	R&D Expenses	34	178.11	448.70	0.00	2168.00
	Market Share	39	0.06	0.09	0.00	0.33
2015	Sales	132	848.34	6455.33	0.00	53315.45
	Cost of Production	132	656.66	4976.42	0.00	41789.48
	Gross Profit	132	419.14	3149.28	-0.82	25257.63
	Operating Revenues	132	59.23	622.79	-0.15	7393.37
	Net Income	132	78.74	612.45	-1.22	5626.73
	Market Value	132	2932.80	22406.76	0.00	180728.60
	Intangible Assets	132	89.67	628.08	0.00	5560.30
	R&D Expenses	34	186.91	476.41	0.00	2317.50
	Market Share	39	0.06	0.09	0.00	0.39
2016	Sales	132	680.80	5794.83	0.00	50937.12
(through	Cost of Production	132	510.85	4344.40	0.00	37638.12
September)	Gross Profit	132	343.35	2928.34	0.00	26247.74
	Operating Revenues	132	43.09	564.48	-1.45	8143.95
	Net Income	132	73.98	636.78	-0.77	5826.18
	Market Value	132	2535.83	21145.97	0.01	181842.50
	Intangible Assets	132	84.33	614.36	0.00	5375.03
	R&D Expenses	34	181.01	453.32	0.00	2153.00
	Market Share	39	0.05	0.09	0.00	0.42

Table 4: Characteristics of firms that have adopted APIs by year.

Distribution of Firms Using APIs % of Total Firms Using APIs

	2013	2014	2015
Main Sector			
Information (NAICS 51)	34.2	33.3	35.0
Retail and Trade (NAICS 44-45)	21.1	21.3	23.8
Manufacturing inc. Consumer Goods (NAICS 31-33)	19.7	20.0	16.3
Financial (NAICS 52)	7.9	6.7	6.3
Other	17.1	18.7	18.8
Region (Publicly Traded)			
USA	58.8	58.8	61.3
Europe	23.5	23.5	25.0
Asia	5.9	5.9	6.3
Oceania	4.7	4.7	2.5
Canada	3.5	3.5	2.5
Americas	3.5	3.5	2.5
Size			
Employees (Thousands)	13.49	13.42	13.64
Number of Firms	47	108	123

5 Analysis

Our analysis has two main objectives. First, we seek to understand the causal impact of API adoption on firm outcomes. Second, we want to explore empirical regularities in the the relationship between measures of API usage intensity, internal and external developer keys, and firm financial outcomes.

Our analyses proceed in the following order. First, we run a simple difference in sample means analysis. Alongside, we run a placebo test. Then we graph firm financial performance before and after adoption. Next we test to see whether the common path assumption holds, and examine evidence showing that the impact of API use is not constant over time. We then confirm that after including time fixed effects the common path assumption holds. Having established its validity, we run difference in difference regressions with firm and time fixed effects to generate our main estimates of the impact of API adoption.

To achieve the second objective we run a set of regressions relating the intensity of different forms of API usage and financial outcomes.

Event Study

To achieve the first objective we perform an event study. To do so we conceptualize a firm adopting APIs as being subject to a policy treatment. Firms must exhibit parallel trends in their financial outcomes before treatment for this approach to be valid. In other words, firm outcomes need to be uncorrelated with their likelihood of being treated, except through variables we explicitly control for. We test to see whether outcomes are different between the subset of firms which did and did not adopt. This difference, or the difference in difference (DD), is what we estimate.

Conducting a DD study requires careful set up. The quality of the comparison group determines the quality of the evaluation. Thus, our first task is to define treatment and control groups.

As noted above, in our main analysis we proxy treatment as beginning when a firm has an API call for the first time. We run our analysis using two different control groups. The first control group we use is a nearest-neighbor matching of firms that never adopt APIs. The second control group we use are firms that do adopt APIs but do so in a different period.

Our main hypotheses are that the adoption of an API strategy increase short and long term profits. Our measures of short term impact are net income and operating profits, while our measure of long term impacts are market value and intangible assets.

Our first approach to calculating the impact of API adoption is a comparison of sample means for adopting and non-adopting firms before and after the adopting firms' adoption. The left column of table 5 displays mean outcomes for firms that do and do not adopt APIs in a given quarter before and after that period.[4] In other words, the control group used is the subset of firms that will adopt APIs, but have not yet. The third row is the first difference in outcomes for the treated and control group before and after treatments. The difference between these two differences is the sample mean DD estimate of the impact of API adoption. This is reported in the lower right hand corner. Alongside and to the right, the p-value is reported.

We find that API adoption is related to a statistically significant increase in net income. The point estimate is large. API adopters earn 3 million more per year than non-adopters. Because the typical firm is treated for longer than a single period, these sample mean estimates represent the average accumulated impact over time. Comparing the point estimate with the average net income of firms in the sample, this 3 million dollar increase corresponds to an approximately 3 percent increase in net profits.

Our point estimate of the increase in market value, 206 million, is also significant at the 10 percent level. It is similar to the size of the point estimate of the increase in intangible assets (although this estimate is not significant). Both these estimates are also very large. Can they be justified by the estimated increase in net income? The present discounted value of an indefinite stream of 3 million in profits is 206 million with a constant annual discount rate of 1.4 percent. This is a very low implied discount rate, but not outrageously so. The Wall Street Journal (2016) reported that in late 2016 that rates on 7-to-10-year bonds of high-quality U.S. companies were 3.14 percent.

Still, that implies that the increase in market value and intangible assets is at least double what would be inferred from the estimated increase in profits. One possible explanation is that investors expect still larger benefits to manifest from API adoption in the medium and long runs (i.e. more than 2 or 3 years post adoption). This fits our hypothesis that firms might receive the largest benefits from API adoption only in the longer term.

In other words, API adoption sends a signal to investors. Firms that adopt APIs might be seen by investors as being more adaptable and able to take advantage of future opportunities. These opportunities might explicitly entail API use, such as opportunities to found online marketplaces, monetize firm data, or sell digital services. Another possibility is that the ability to exploit new non-API opportunities be strongly correlated with API adoption. Investors see a business implementing APIs and they assume that it has the sort of bold, dynamic leadership willing to implement disruptive innovations and best practices. Even if the firm does not gain directly from the APIs it implements, its stock price may gain from this reputation. Of course, the signal of API adoption may be interpreted incorrectly or irrationally. Investors may be hyped up by the implementation of a new API initiative and over-correct the price of companies involved.

[4] The exception is the period Q1 and Q2 2013. January firms are discarded, and February and March adopters are consolidated with Q2 2013 adopters

A possible concern about this analysis is that some non-API related trend is driving the result. To test for this possibility we perform a placebo test. From the general Compustat and Orbis sample, we selected two groups of firms. They were selected to be a similar as possible to the sample of API adopting firms before adoption in both financial attributes and number.[5] These samples should have experienced comparable aggregate and firm level shocks as the real sample. We assigned to the 'adopter' group a distribution of placebo adoption periods matching the actual treatment distribution. The right hand side of table 5 reports the result of this test.

In the placebo test none of these estimates are close to being significantly different from zero. We conclude that the significant outcomes we estimated for adopters were not due to some aggregate trend or random chance alone.[6]

Table 5: Difference in sample means for API adopters and non-adopters. Control group is firms who have not yet adopted APIs in a given quarter, but will in the future. Outcomes are in billions. Difference-in-difference treatment estimate is bold.

Sample Means Difference in Difference					Placebo Sample Means Difference in Difference				
	Adopters	Non-Adopters	Difference	P-value		Placebo	No Placebo	Difference	P-value
	Mean	Mean				Mean	Mean		
Net Income					Net Income				
Before Treatment	0.050	0.012	0.038		Before Treatment	0.016	0.031	-0.015	
After Treatment	0.056	0.016	0.041		After Treatment	0.027	0.031	-0.004	
Change in Mean	0.007	0.004	**0.003**	*0.019*	Change in Mean	0.011	0.000	**0.011**	*0.876*
Operating Income					Operating Income				
Before Treatment	0.006	0.016	-0.002		Before Treatment	0.036	0.036	0.000	
After Treatment	0.011	0.008	-0.005		After Treatment	0.038	0.036	0.002	
Change in Mean	0.005	-0.008	**0.013**	*0.084*	Change in Mean	0.002	0.000	**0.002**	*0.977*
Sales					Sales				
Before Treatment	0.109	0.120	-0.010		Before Treatment	0.242	0.386	-0.144	
After Treatment	0.161	0.134	0.028		After Treatment	0.246	0.335	-0.089	
Change in Mean	0.052	0.014	**0.038**	*0.143*	Change in Mean	0.004	-0.052	**0.056**	*0.797*
Cost of Production					Cost of Production				
Before Treatment	0.053	0.092	-0.039		Before Treatment	0.120	0.133	-0.013	
After Treatment	0.032	0.093	-0.061		After Treatment	0.105	0.128	-0.023	
Change in Mean	-0.022	0.001	**-0.022**	*0.159*	Change in Mean	-0.015	-0.004	**-0.010**	*0.563*
Market Value					Market Value				
Before Treatment	0.937	1.019	0.082		Before Treatment	1.076	1.051	0.025	
After Treatment	1.101	0.977	-0.039		After Treatment	0.961	0.918	0.043	
Change in Mean	0.164	-0.043	**0.206**	*0.077*	Change in Mean	-0.116	-0.133	**0.017**	*0.980*
Intangibles					Intangibles				
Before Treatment	0.115	1.219	-1.104		Before Treatment	0.143	0.129	0.014	
After Treatment	0.100	1.030	-0.930		After Treatment	0.135	0.088	0.047	
Change in Mean	-0.015	-0.189	**0.174**	*0.34*	Change in Mean	-0.008	-0.041	**0.033**	*0.709*

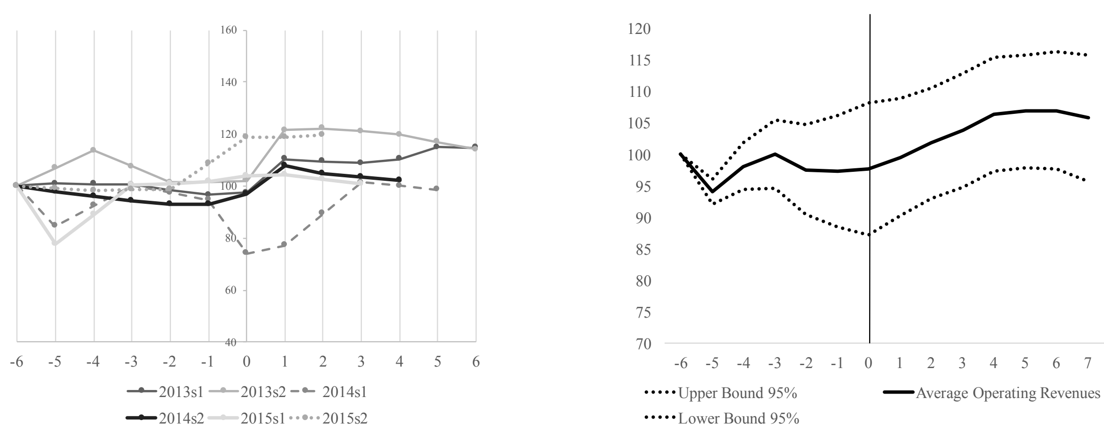
Another approach to understanding the impact of API adoption is to plot outcomes as a function of the treatment and perform a visual inspection. Figure 7 shows how operating profits vary for firms before

[5]Specifically, we matched as best as possible non yet adopting firms with never adopting firms in year 2011. We ensure that the number of firms in each group was similar.

[6]The placebo exercise can also be estimated through a re-sampling procedure on the adopting group. In other words, within that sub-sample we can randomly reassign the date of treatment. Using that procedure we also find no significant effect from the placebo treatment.

and after API adoption. In the left panel de-trending reveals an increase in operating revenues one quarter after adoption. There seems to be no pattern indicating that early or late adopters gain a larger benefit on average. Figure 7.b aggregates the observations by centering all firms' adoption at period zero. This figure suggests that the peak benefits on operating revenues arise four periods after first adoption. While we do not find any evidence of benefits increasing further after this period, this could be due to few observations of firms with more than 6 to 8 quarters of treatment.

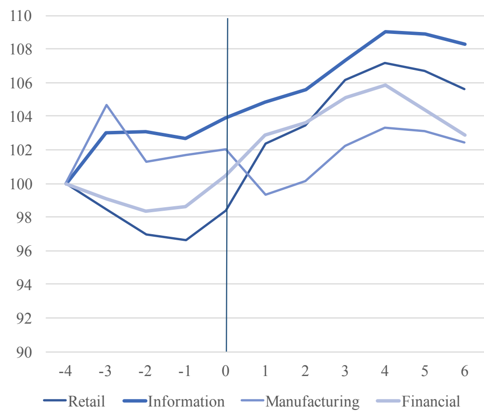
Another question is whether the effect of API adoption varies by industry. Figure 8 graphs operating income as a function of quarters since adoption as well. However, it breaks down the trends by industry. Retail companies who adopt see the largest increase in operating income after adoption. Their operating incomes decrease by roughly 3 percent in the 3 quarters before API adoption. However, over the three quarters over and post adoption, it rises to 3.5 percent above their average 4 quarter prior to adoption level. In the following year, operating incomes for retail companies grew a further 2 percentage points. Financial and information firms saw similar, though slightly smaller gains. The only industry that seems on average to have not done better after adoption is manufacturing firms. Their operating incomes increased by about 1.5 percent in the 3 quarters prior to adoption. In the 7 subsequent periods operating income for manufacturing grew very slowly, by only one additional percentage point. This result is unsurprising as Gartner and other API analysts argue that retail firms, information sector firms, and financial firms are particularly good candidates for an API strategy.



(a) By Adoption Period, De-trended

(b) Consolidated, With Confidence Interval

Figure 7: Indexed Operating Income by Quarter Since Adoption. In both figures, API adoption dates realigned so all occur in period zero. The left figure reports operating income before and after API adoption. It is first differenced, to take out average growth over the period. Half-year periods. Operating revenue is indexed to 100 for all adopters in by their average operating incomes in the final two quarters of 2009. Different line types indicate different half-year adoption cohorts. The right figure reports average indexed operating for all groups consolidated. Operating income is indexed at 100 for all groups 6 periods before adoption. The inner line is average operating income as a function of period before/after adoption. The outer pair of lines correspond to a 95 percent confidence interval.



(a) Indexed Operating Income by Quarter Since Adoption and Industry

Figure 8: Operating revenue subdivided by NAICS Industry for the four most common NAICS industries of API users.

From figure 8 we also learn that firms in different industries had importantly different characteristics. This suggests that we should control for firm level fixed effects or trends in our analysis. It could also be the fact that firm specific factors are creating the concavity in outcomes in figures 7 and 8.

Therefore we would like to create a new estimate of the impact of APIs controlling for firm level fixed effects. However, before we proceed to that estimate, there are several important factors to consider.

5.1 Obstacles to Identification

In the subsequent section, we estimate the effect of API adoption using a difference in difference approach. Difference in difference estimation can be contaminated by many factors. We pay particular attention to three issues documented by the literature. The first potential issue arises when treatment probability is based on a difference in outcomes prior to the intervention. In other words, difference in difference estimation can be biased towards zero when selection into treatment is influenced by transitory shocks to past outcomes (Ashenfelter, 1978; Chay et al., 2005).

This issue is not simple to address. We restrict our analysis to data beginning in 2011, one year before the earliest treatments to minimize this possibility. This helps to prevent any potential contamination from heterogeneous shocks due to the financial crisis. We also include yearly fixed effects. This controls for transitory shocks in our sample period which are common to all firms.

The second, and perhaps most important, concern is that the assumption of a common path be satisfied. This assumption is also known as parallel trends. Identical counterfactual trends in treatment and control states is an important identifying assumption. Failing to have a common path might mean that firms who were likely to adopt and those are unlikely to adopt were trending towards different outcomes regardless of treatment. We selected our two control groups to make this unlikely. However, we test to see whether it remains a concern by testing for non-parallel trends directly. While our previous visual inspection presented no obvious evidence of non-parallel trends, in a model with multiple treatment groups and multiple periods, it becomes more difficult to provide a simple visual evidence for the common trends, as discussed in Card and Krueger (2000) and Hastings and Gilbert (2005).

We test our assumption of parallel trends by estimating a regression of the outcomes of interest on leads and lags of the treatment. If none of the estimates of the effect of leads of treatment are significant, then we cannot reject the hypothesis that those who are about to adopt and those who are not are exhibiting a common path. This procedure is widely used in the literature.[7] We estimate

$$Y_{i,t} = \alpha_i + \sum_{l=-n}^{q} \beta_l T_{i,t}(t{=}k{+}l) + \epsilon_{i,t}, \tag{1}$$

where k is the adoption date and α_i is a firm fixed effect. Instead of a single treatment effect, we have include n leads and q lags of the treatment effect. β_l is the coefficient on the *l-th* lead or lag.[8] A test of the differences in differences assumption is $\beta_l = 0$, i.e. the coefficients on all leads of the treatment should be zero. Moreover, the β_l for $l > 0$ may not be identical. For example, the effect of the treatment could accumulate over time, so that β_l increases in l.

Table 6 provides estimates of (1). The right row of the figure augments (1) with a linear and quadratic year control. [9] Specifically, we include indicator variables for years 1 to 3 before adoption, years 1–2 after adoption and year 3 forward. In addition to telling us about whether common paths exist, this specification allows us to analyze how the impact of the treatment varies with the amount of periods post adoption.

In table 6 the first 7 columns show that firms close to adoption do have significantly different outcomes than firms that have also not adopted but are further from adoption. This is evidence against the existence of a common path. However, the addition of a quadratic trend as a function of time period resolves the issue. The right eight columns include a quadratic trend as a function of time in years. This time trend is not a function of time until adoption, but is rather absolute in the year. The inclusion of the trend makes insignificant the coefficient of the treatment on pre-adoption periods. We can proceed with the knowledge that after controlling for this time trend, a common path exists.

Beyond the validity of the common path assumption, table 6 gives us information about when and how API benefits arise. A key result is that, for all financial outcomes, the post-treatment effects differ from year to year. This suggests that the effect of API adoption is not a one time boost, but modifies the firms' growth

[7]For instance David (2003) or Mora and Reggio (2012) who document several more examples and a generalized procedure.

[8]Following the impact evaluation literature, if k is the treatment date, a lead corresponds to the period t=k-l while a lag is t=k+l.

[9]Standard errors for these regressions and all other estimates are clustered at the firm level.

Table 6: Testing the Common Paths Assumption

Estimation Lags and Leads, Testing for Common Path												
	(1) Net Income	(2) Oper.-Income	(3) Market Cap	(5) Sales	(6) Costs	(7) Intangibles	(8) Net Income	(9) Oper.-Income	(10) Market Cap	(11) Sales	(12) Costs	(13) Intangibles
Pre-Treatment (t-3)	-0.00332**	-0.00214+	-0.172**	-0.0736**	-0.0629**	-0.187**	-0.00901	-0.000955	0.157	0.0306+	0.032	0.117+
	(-3.12)	(-1.46)	(-129.08)	(-15.80)	(-15.25)	(-55.17)	(-0.58)	(-0.20)	0.81	1.23	0.4	1.11
Pre-Treatment (t-2)	-0.00195*	-0.00219+	-0.177**	-0.0519**	-0.0429**	-0.227**	-0.00515	-0.00047	0.0443	0.0209+	0.0213	0.0467+
	(-1.89)	(-1.55)	(-136.22)	(-11.49)	(-10.72)	(-69.00)	1.23	(-0.15)	1.18	1.09	0.41	1.16
Pre-Treatment (t-1)	-0.00171*	0.00	-0.145**	-0.0284**	-0.0232**	-0.229**	-0.00318+	0.00053	-0.016	0.0157	0.0145	-0.0258
	(-1.71)	(-0.77)	(-115.30)	(-6.53)	(-6.01)	(-72.72)	(-1.06)	0.26	(-0.54)	0.35	0.44	(0.51)
Adopt	0.0904**	0.0798**	0.0678*	0.0245**	-0.242**	0.07	0.0906**	0.0792**	0.0148**	0.0264**	-0.157**	0.0858*
	3.46	3.52	1.92	2.37	(-2.57)	1.04	3.47	3.49	2.42	2.56	(-2.74)	1.79
Pot-Treatment (t+1)	0.173**	0.01	-0.01	0.12	-0.208**	0.00	0.0173+	0.0783	0.0121**	0.0413	-0.115**	0.0253
	7.01	0.41	(-0.29)	1.26	-2.35	(-0.03)	1.71	0.41	2.37	1.34	(-2.42)	0.38
Pot-Treatment (t>1)	0.03	0.0701**	0.0952**	0.339**	-0.214**	0.108*	0.0129	0.0699**	0.0640**	0.0352**	-0.104**	0.0553+
	1.36	3.01	2.94	3.58	(-2.48)	1.65	1.37	2.99	1.99	3.72	(-2.60)	1.34
Abs. Time Trend							-0.000856*	-0.00034	0.0331**	0.00880**	0.00903**	-0.00402**
							(-1.73)	(-0.52)	53.25	4.15	4.79	(-2.56)
Abs. Time Trend Sq.							0.0000235+	0.0000211	-0.000503**	-0.0000856	-0.000128**	0.00131**
							1.51	1.01	(-25.48)	(-1.27)	(-2.14)	26.36
Ho: Adoption (t0-t4)=0	0.981	0.694	0.155	0.115	0.464	0.272	0.981	0.691	0.666	0.125	0.556	0.673
Adj. R-sq	0.59	0.68	0.16	0.6	0.59	0.42	0.59	0.68	0.04	0.60	0.59	0.40

t statistics in parentheses
+ p<0.10, * p<0.05, ** p<0.01

path. Moreover, the point estimates seem to weakly track the concave shape we observe in figures 7-8. All of the financial outcomes but market value seem to confirm an initial increase that slows down by the end of the third year after adoption. This finding is rather interesting. Our original hypothesis was that firms should see increasing benefits from API adoption over time. This hypothesis was consistent with our finding that increases in market value were larger than would be implied by the observed increases in net income. This result suggests that the benefits of APIs on net income decrease over time. However, this downward trend may just be due to how our sample becomes more limited more quarters post adoption. In appendix table table 13 we show that the common path assumption holds if we substitute the time trend with year fixed effects. It estimates equation (2) alternatively with both time fixed effects and a year trend. It shows that the two lead to equivalent estimates of the coefficients.

5.2 Main Results

Having established that there is a common path after controlling for the year, we proceed to our main regression. These regressions are all run on the only adopters sample, that is, those firms who eventually adopt an API.

To estimate the impact of API adoption, we run the following regression with firm and time fixed effects.[10] Our general specification is

$$Y_{i,t} = \lambda_t + \alpha_i + \beta \mathbf{T_{i,t}} + \gamma_{i,t} \mathbf{Z_{i,t}} + \epsilon_{i,t}, \tag{2}$$

where i denotes the firm and t denotes the time period in quarters. λ_t is a time fixed effect and α_i are firm

[10] As we have many time periods and groups, the econometric logic follows (Bertrand et al., 2004).

fixed effects. $T_{i,t}$ indicates the whether firm i has been treated in period t. In this section we maintain our focus on a binary adoption treatment. We also include covariates $Z_{i,t}$. The covariates we focus on are firm leverage and installed capital. $\epsilon_{i,t}$ is an error term.

Difference in Difference Based on Fixed Effects Regressions

	Net Income		Operating Income		Sales		Costs		Market	Intangible	Research &	Market
	Levels	Δ%	Levels	Δ%	Levels	Δ%	Levels	Δ%	Cap	Assets	Development	Share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
i. Adopters Sample												
Dif in Dif Regression (DD)	0.009	0.0152	0.0362**	0.004	0.025	0.0173	-0.108+	-0.0168	0.127**	0.157**	0.0336	0.00514**
	(0.47)	(0.68)	(2.00)	(1.69)	(0.34)	(0.08)	(-1.58)	(-1.63)	(4.82)	(3.02)	(0.21)	(2.51)
Firm and Quarter FE (F&Q)	0.001	0.001	0.0146	0.0012+	0.038	0.0091	-0.0114+	-0.0101+	0.0223*	0.0180+	0.0715	0.00781
	(1.35)	(1.55)	(0.71)	(1.71)	(1.11)	(0.44)	(-1.76)	(-1.81)	(1.9)	(1.69)	(0.88)	(0.39)
Firm and Year FE (F&Y)	0.007	0.0053	0.0341*	0.002+	0.022	0.0138+	-0.147**	-0.0181	0.0453*	0.0880*	0.0294+	0.00622
	(0.35)	(0.95)	(1.88)	(1.77)	(0.30)	(1.78)	(-2.16)	(-1.11)	(1.77)	(1.71)	(1.72)	(1.34)
Controls and F&Y FE	0.002	0.0049	0.024	0.002	0.343**	0.0137	-0.420**	-0.0106	0.103**	0.141**	0.0194	0.0061
	(0.11)	(1.43)	(1.09)	(1.42)	(5.63)	(1.66)	(-7.13)	(-1.55)	(3.99)	(2.71)	(1.25)	(1.42)
Observations	1363	1231	1378	1246	1353	1221	1341	1209	1308	1214	366	165
i. Adopters Sample (Randomized)												
Dif in Dif Regression (DD)	0.006*	0.0121	0.029*	0.003	0.044+	0.0166	-0.0107+	-0.0104+	0.071**	0.061	0.0116	0.0011+
	(1.89)	(0.94)	(1.88)	(1.16)	(1.75)	(0.66)	(-1.84)	(-1.77)	(3.10)	(1.01)	(1.11)	(1.75)
Firm and Quarter FE (F&Q)	0.0017	0.001	0.0014	0.0007	0.036+	0.0088	-0.0008	-0.0072+	0.013*	0.020*	0.007	0.0087
	(1.20)	(0.84)	(1.31)	(1.66)	(1.78)	(1.33)	(-1.64)	(-1.84)	(1.99)	(1.77)	(0.68)	(0.21)
Firm and Year FE (F&Y)	0.0062*	0.0044+	0.026*	0.002+	0.029+	0.0121	-0.022+	-0.01	0.0306**	0.077**	0.0109+	0.00069
	(1.86)	(1.71)	(1.84)	(1.81)	(1.770)	(1.43)	(-1.81)	(-1.56)	(2.73)	(2.97)	(1.83)	(0.93)
Controls and F&Y FE	0.0017	0.0033+	0.0241*	0.002	0.0309**	0.012	-0.025**	-0.0102	0.017**	0.002	0.011+	0.00061
	(1.57)	(1.73)	(1.830)	(1.67)	(2.11)	(1.60)	(-2.91)	(-1.31)	(2.88)	(0.10)	(1.77)	(0.84)

t statistics in parentheses
+p<0.1, *p<0.05, **p<0.01

Table 7: First four rows are estimates of variations on equation (2). Bottom four rows randomly reassign treatment dates.

The first four rows of 7 performs difference in difference estimates following equation (2). It reports the estimated average treatment effect of API adoption. It is the β estimated in (2). Net income, operating income, sales, and costs are annualized and measured in billions of dollars or log billions. Market value and intangible assets are percentage deviations from the firms' long term averages. R&D and Market Share outcomes are measured as percentage deviations from the firms' last 5 year pre-treatment average.

The first four rows differ in terms of which fixed effects and covariates are included. In the first row, no fixed effects are included. In the second, year and quarter fixed effects are included. In the third, firm and year fixed effects are included. In the final row controls for contemporaneous installed capital and leverage are included in addition to firm and year fixed effects.[11]

While the precise estimates of the impact of APIs are sensitive to the specification, a positive relationship between API adoption and increases in market value is our most robust result. When fixed effects are not included, we estimate that API adoption leads to a 12.7 percent increase in market capitalization. In our most conservative specification, including both quarter and firm fixed effects, we find that API adoption increased firms' market cap by 2.23 percent. However, this is still a very large effect in level terms (taking into account the average size of companies in our sample), and it is significant at the five percent level.

Another effect which is significant at the five percent level in almost all specifications (except after

[11]We calculate installed capital as total assets less intangible assets and cash.

including quarter fixed effects) is the impact of API adoption on intangivle assets. The magnitude of the estimates is similar to, or slightly larger than the percentage increase in market capitalization.

Given that the average API adopting firm has a market value of 2.3 billion dollars in 2013, a 10 percent increase in market value indicates an increase in value on the scale of 200 million dollars. The scale of the estimated increase in intangible assets is similar though slightly larger. We would anticipate that any technological advantage conferred by API adoption would mainly do so through increasing the productivity of a firm. API adoption inuitively would have no strong direct effect on the amount of capital owned by a firm (and is controlled for explicitly in the fourth specification). Therefore, the bulk of the impact of API adoption should be realized as increased intangible assets. Given that intangible assets are only about 3 percent of market value in 2016, we would expect the percentage change in intangible assets would be much higher than the change in market value. The fact that it is not may be due to how intangible assets are accounted for in the compustat.

The effect of API adoption on other financial outcomes follows the anticipated signs, but is not very significant, except for in level of costs. Costs go down by 420 million dollars a year after API adoption (after controlling for increases in leverage and capital stock) and by 11.4 million (although this is only significant at the 10 percent level) with only year and quarter fixed effects.

Regarding income and sales, the effect of API adoption is only significant in some specifications. But the point estimates are all positive and reasonable. They are a bit small however, compared to the large estimated increases in market cap. In order to justify an increase in valuation of 200 million, and assuming a 5 percent discount rate, a firm should see the equivalent of an indefinite yearly increase in profits of 10 million. This is slightly higher than the point estimates of the impact of API adoption on net income seen.

There are at least two riddles. The first is why the increase in market cap is so large if the increase in net income is so small. The other is why the increase in net income is so small if the decreases in costs are so large.

To answer the second question first, a large part of the explanation seems to be that firms which adopt APIs are also financing growth using debt. Table 17 in the appendix shows that adopters are have increasingly high leverage ratios over time, from a mean of 2.32 in 2011 to 4.07 in 2016. Interest payments seem to be soaking up any gains from a decrease in costs. The fact that firms which see increases in APIs have larger increases in market caps than are predicted by their increases in their net incomes we interpret as being due to the fact that profitability benefits of API adoption are only manifested in the long run.

Another way an increase in potential future profitability might show up in the data is as an increase in market share. Under the platform business model of 'ubiquity first', firms first seek to maximize their revenues and user engagement, and only later worry about profitability. Indeed, under some specifications we see a signficant increase in market share. No robust result is found on R+D expenditure.

Finally, across specificatins, findings tend to be more significant when using levels as the outcome rather than logs. This suggests that API adoption tends to have a level effect on outcomes rather than a percentage change effect.

In the bottom half of the table, for robustness we perform a robustness analysis. We randomly reassign the date of treatment. This removes any potential endogeneity in the exact date of adoption, and provides a lower bound on the true effect.

5.3 Developer Portals

We discussed the evidence for the beneficial effects of adopting on short and long run outcomes. We refine our estimations by performing a second round of regressions interacting the adoption event with an indicator variable that assigns 1 if a particular API has opened a developer portal. Table 8 shows the results of these estimations.

Difference in Difference Based on Fixed Effects Regressions

	Net Income		Operating Income		Sales		Costs		Market	Intangible	Research &	Market
	Levels	Δ%	Levels	Δ%	Levels	Δ%	Levels	Δ%	Cap	Assets	Development	Share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
i. Adopters Sample												
Difference in Difference												
Adopt	0.00692	0.0165	0.014	0.004	0.0392**	0.0173	-0.0266	-0.0133	0.0815*	0.00563	0.021+	0.00689+
	(1.18)	(0.88)	(0.92)	(1.09)	(3.08)	(0.08)	(-0.34)	(-1.13)	(2.25)	(0.28)	(1.77)	(1.81)
Adopt x Open Portal	0.0101*	0.0303+	0.0228**	0.002	0.0107	0.007	0.00134*	0.0003	0.0885	0.0283+	-0.031*	0.01968+
	(2.06)	(1.86)	(2.72)	(0.54)	(1.22)	(0.22)	(2.03)	(1.11)	(1.34)	(1.68)	(-1.92)	(1.92)
Observations	201	201	194	194	162	162	200	200	166	201	175	161
Fixed- Effects Regression												
Adopt	0.0060+	0.0055	0.0108	0.002	0.0387**	0.0138+	-0.0218	-0.0118	0.061*	0.00732	0.0185*	0.00847**
	(1.89)	(1.05)	(1.06)	(1.16)	(3.18)	(1.78)	(-0.39)	(-1.31)	(2.22)	(0.36)	(1.95)	(-3.02)
Adopt x Open Portal	0.00803+	0.0101+	0.0213*	0.013+	0.0216*	0.0021*	0.00126+	0.007+	0.0948+	0.0295+	-0.0414*	0.0161+
	(1.81)	(1.72)	(2.54)	(1.74)	(2.00)	(1.90)	(1.88)	(1.74)	(1.77)	(1.73)	(-1.88)	(1.84)
Observations	197	197	194	194	162	162	197	197	166	197	161	158
Controlled Fixed Effects												
Adopt	0.00879	0.0051	0.01894	0.002	0.0437*	0.0137	-0.0282	-0.0112	0.066+	0.00303	0.0078**	0.00815*
	(1.47)	(1.04)	(1.41)	(1.56)	(2.33)	(1.66)	(-0.89)	(-1.32)	(1.81)	(0.61)	(2.18)	(2.08)
Adopt x Open Portal	0.00802	0.0233	0.0245	0.014	0.0147	0.0017	-0.00011	0.001	0.076+	0.0188	-0.035*	0.01889
	(0.94)	(0.42)	(1.07)	(1.43)	(0.13)	(1.41)	(-0.43)	(1.46)	(1.76)	(0.35)	(-1.95)	(1.50)
Observations	197	197	194	194	162	162	197	197	166	197	161	158
ii. Reshuffled Treatment Dates (1000 Draws)												
Difference in Difference												
Adopt	0.0031		0.022		0.02**		-0.00419		0.115+	0.00171	0.0193+	0.00511**
	(1.11)		(0.51)		(2.81)		(-0.45)		(2.75)	(0.06)	(1.81)	(2.88)
Adopt x Open Portal	0.005*		0.043+		0.0107		-0.060*		0.0982	0.0416+	-0.0101*	0.0211+
	(2.22)		(1.79)		(1.22)		(-2.05)		(1.34)	(1.81)	(-1.87)	(1.89)
Observations	201		194		162		200		166	196	175	161
Fixed- Effects Regression												
Adopt	0.001+		0.018		0.0472**		-0.0285+		0.091*	0.00188	0.00921**	0.00684**
	(2.79)		(1.46)		(3.10)		(-1.71)		(2.13)	(0.07)	(2.09)	(3.66)
Adopt x Open Portal	0.005+		0.0234		0.0201+		-0.0150+		0.0881+	0.0424+	-0.037*	0.0111+
	(3.31)		(1.54)		(1.70)		(-1.89)		(1.85)	(1.91)	(-1.94)	(1.97)
Observations	197		194		162		197		166	194	161	158
Controlled Fixed Effects												
Adopt	0.001+		0.0149		0.0313**		-0.0158+		0.0805+	0.00376	0.00588**	0.00722*
	(1.74)		(1.41)		(3.33)		(-1.86)		(1.71)	(1.35)	(5.22)	(2.22)
Adopt x Open Portal	0.002		0.0315+		0.0107		0.0106		0.0799	0.0341	-0.0304**	0.00889
	(1.09)		(1.77)		(0.56)		(0.48)		(1.86)+	(0.68)	(-2.11)	(1.63)
Observations	197		194		162		197		166	194	161	158

t statistics in parentheses
+p<0.1, *p<0.05, **p<0.01

Table 8: Developer Portal Estimations: Second and fifth row include year and firm fixed effects. Third and sixth row include these as well as leverage and installed capital controls.

The most striking finding in this table is the result that, across specifications, opening a developer portal is significantly negatively related with R+D expenditure. This fits with the hypothesis that developer portals allows for outside app makers to substitute for R+D expenditure within the firm.

5.4 Intensity and Predictive Results

It seems likely that the intensity of API use is related to outcomes. Our third hypothesis is that firms that use APIs more intensely are likely to have more positive financial outcomes than those which do not. Table 9 reports an estimate based on specification (1), but with log calls and log data as explanatory variables.

A positive relationship between log calls or data and net income, operating income, and market value is robust to several specifications. To get a sense of the size of the effects estimated, recall that the typical firm processes about 7.2 terabytes of API data and 210 million calls per quarter four quarters after first adoption. This is a very sizable amount of data. It is the same amount of data contained in 1440 five gigabyte high resolution movies. It is an incredible amount of text. In uncompressed plain 8-bit text, it is 7.2 trillion characters. The average amount of data per call in that interval is therefore about 34 megabytes, enough for several very high resolution pictures. Table 9 indicates that an order of magnitude increase, say from 720 gigabytes of data to 7.2 terabytes, is related to an approximately 5 percent increase in market value. It is related to an approximately 12 million dollar increase in operating income in our preferred specification, row three.

An order of magnitude increase in calls has similar effects on market value and operating income, but larger impacts on net income. While an increase of 13.4 million a year in net income is large, the amount of data involved is large as well. Neither form of API intensity has much effect on intangible assets or costs. As above, it is unclear why the impact on market value would be lower than the impact on intangibles in percentage terms. However, one should be cautious in interpreting these results. Any concern about the endogeneity of the binary treatment is doubly concerning here, as one can imagine, for example, firm specific shocks to demand increasing both calls and net income.

Table 10 shows the result of a linear regression of a firms log calls on log amount of internal and external developers and other variables.[12] It shows that calls are related strongly to the contemporaneous number of programmers but are uncorrelated with past numbers of developers. This result suggests that API programming is better modeled as a contemporaneous input rather than as an investment. However, we know that API developer keys are rarely retracted, so developers and lagged developers are highly co-linear.

The results in our intensity section indicate that a one unit increase in log calls is relatedd to an increase in net income of eight million dollars. The results in table 10 suggest that a roughly one unit increase in

[12] As before + $p<0.10$, * $p<0.05$, ** $p<0.01$. Columns (2)-(4) use Arellano-Bond panel estimation.

Difference in Difference Based on Fixed Effects Regressions								
	Net Income (1)	Operating Income (3)	Sales (5)	Costs (7)	Market Cap (9)	Intangible Assets (10)	Research & Development (11)	Market Share (12)
i. Regression Using Data								
Only Firm FE	0.00606+	0.0165+	0.013	0.000	0.0474*	0.014	-0.00077	0.00944
	(1.83)	*(1.96)*	*(1.57)*	*(-0.83)*	*(2.31)*	*(1.74)*	*(-1.33)*	(0.54)
Firm and Year FE	0.00812*	0.0120*	0.0163*	0.000	0.0508*	0.003	-0.000689**	0.000108
	(1.77)	*(1.78)*	*(1.75)*	*(-1.38)*	*(2.30)*	*(0.31)*	(-3.75)	(1.44)
Additional Controls FE	0.000	0.0127+	0.003	0.000	0.0564+	0.003	-0.000557	0.0000634
	(0.03)	*(1.90)*	*(0.88)*	*(-1.09)*	*(1.94)*	*(0.30)*	(-0.55)	(0.37)
ii. Regression Using Calls								
Only Firm FE	0.0104*	0.0155*	0.0212*	0.000	0.09253+	0.009	-0.0477	0.0218+
	(2.22)	*(2.26)*	*(2.22)*	*(-0.84)*	*(1.95)*	*(1.04)*	*(-1.01)*	(1.77)
Firm and Year FE	0.0126**	0.0176**	0.0250**	0.000	0.054	0.000	-0.0413+	0.046*
	(2.43)	*(2.40)*	*(2.38)*	*(-1.27)*	*(1.32)*	*(0.02)*	(-2.15)	(2.19)
Additional Controls FE	0.0134*	0.0143+	0.000	0.000	0.0588+	0.000	-0.0316	0.0303
	(2.14)	*(1.80)*	*(0.57)*	*(-0.61)*	*(1.93)*	*(0.03)*	(-0.50)	(1.15)

t statistics in parentheses
+p<0.1, *p<0.05, **p<0.01

Table 9: Financial Outcomes As a Function of API Usage Intensity. Only firms which adopt APIs. Net income through cost columns are in billions. Market cap through market share are defined as above.

log developers leads to a one unit increase in log calls. Therefore, assuming no inter-temporal effects and that the average outcome is equal to the marginal outcome, these results suggest that firms are maximizing profits if the cost of maintaining an order of magnitude higher number of developers is eight million dollars.

Table 10: Fixed Effects Regression of Log Developers (with lags) on Log Calls

	(1)	(2)	(3)	(4)	(5)	(6)
Internal	0.742**	1.098**	1.157**	1.118**	0.902**	0.958**
t - statistic	*3.25*	*2.15*	*2.17*	*2.02*	*2.72*	*2.96*
External	0.771**	0.319	0.371	0.32	0.777**	0.747**
	4.55	*1.32*	*1.36*	*1.15*	*3.77*	*3.63*
Calls (t-1)		0.223*	0.0901	0.157		
		1.88	*0.57*	*0.9*		
Internal (t-1)		-0.00998		0.138	0.132	0.223
		(-0.04)		*0.53*	*0.66*	*1.05*
External (t-1)			0.251	0.079		
			1.21	*0.24*		
d2013					(.)	
					(.)	
d2014					0.138+	
					1.58	
d2015					0.0428	
					1.82	
d2016					-0.482	
					(-0.26)	
trend						-0.132**
						(-2.21)
Constant	-10.44**	-8.586**	-10.43**	-9.684**	-11.05**	-9.737**
	(-2.15)	*(-4.04)*	*(-3.76)*	*(-3.36)*	*(-5.75)*	*(-2.96)*

We can also estimate the direct effect of having more developers on firm outcomes. Table 11 does so.[13] An interesting result we observe in this table is the significant relation between internal and external developers and net income but market value is significantly influenced only by external developers. While the

[13] As before + p<0.10, * p<0.05, ** p<0.01.

information to further test robustness checks is limited, we speculate that is the external usage of APIs what might be sending signs to the market of better future performance. Lastly, we find two more correlations that are not surprising: costs reduce with internal developers and intangibles increase as internal developers increase.

Table 11: Contribution of Developers to Financial Firms' Outcomes

Developers and Productivity Estimates based on Fixed Effects Regression												
	Net Income		Oper.-Income		Market Value		Sales		Costs		Intangibles	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Internal	0.000312+	0.000529+	0.000558**	0.0000731	0.0178	0.0123	0.00137*	0.00251*	-0.000832	-0.00291**	0.0392	0.104+
t - statistic	*1.73*	*1.63*	*2.01*	*0.17*	*0.73*	*0.39*	*1.78*	*1.68*	*(-0.82)*	*-2.01*	*0.47*	*1.51*
External	-0.000303	0.000761+	0.000349+	0.00011	0.0474**	0.00572	-0.000457	-0.000959	0.0000405	0.000152	0.022	0.0143
t - statistic	*(-0.93)*	*1.52*	*1.59*	*0.31*	*2.43*	*0.23*	*(-0.56)*	*(-0.82)*	*0.05*	*0.13*	*0.35*	*0.18*
Internal (t-1)		0.000102		-0.000437		0.00627		0.000334		0.000416		0.0128
t - statistic		*0.18*		*(-1.07)*		*0.22*		*0.25*		*0.32*		*1.14*
External (t-1)		0.000111		0.000112		0.0138		-0.0000184		-0.000528		-0.00826
t - statistic		*0.24*		*0.35*		*0.6*		*(-0.02)*		*(-0.51)*		*(-0.11)*
Constant	0.00257**	0.00281*	0.00336**	0.00379**	0.821**	0.905**	0.0394**	0.0359**	0.0233**	0.0156**	0.842**	0.600**
t - statistic	*2.82*	*1.79*	*6.09*	*3.73*	*1.95*	*2.36*	*2.29*	*2.75*	*10.66*	*4.35*	*4.75*	*2.35*
Overall R-sq	0.25	0.32	0.42	0.49	0.45	0.34	0.41	0.49	0.48	0.59	0.19	0.27

Predictive API Results by Orientation and Purpose

A final question is what is the effect of different types of APIs are. We classified APIs into three orientations based on their names, the firms using them, and information publicly available online. Table 12 reports logit regressions of the likelihood of positive changes in financial outcomes as a function of both contemporaneous log calls of different types and contemporaneous increases in calls.

We find that contemporaneous net income has the strongest relationship to B2B calls. And cost decreases have the strongest relationship to internal calls. Increases in API calls of various sorts are also associated with a higher likelihood of positive financial changes. All of these regressions suggest that API flow data could have a very useful role in predicting financial outcomes.

A straightforward analysis would suggest that use of B2C APIs would be more related to sales growth, and use of Internal APIs would be more strongly associated with cost reductions. An interesting addition to these findings are those of Table 12. In this estimation we add lagged variables as controls. We find that usually that API have a higher likelihood of improving outcomes in the second period after adoption. Signs and significance broadly remain the same as the previous estimation, except for market capitalization which reduces to levels of 10%.[14]

One reason this might be the case is if online or digital sales substitute for in store sales. Then B2C

[14] The logit estimation relies on a firm having three type of APIs at all times. This drops observations where at least one of these are missing, thus, reducing the observations

Probability of a Larger Outcome Logit Estimation														
	Net Income		Operating Income		Costs		Sales		Market Cap		R & D		Market Share	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(7)	(8)	(9)	(10)
Type of API														
B2C Calls	0.0086**		0.0562+		-0.054		0.04**		0.142*		0.002		0.002	
	(7.17)		*(1.66)*		*(-0.61)*		*(2.79)*		*1.81*		*(0.19)*		*(0.11)*	
B2B Calls	0.38**		0.072**		-0.0737		0.662**		0.718**		0.022		0.0071	
	(6.11)		*(2.75)*		*(-0.92)*		*(4.39)*		*4.13*		*(0.49)*		*(0.25)*	
Internal Calls	0.04+		0		-0.22**		0.165+		-0.279		-0.052+		-0.0211	
	(1.82)		*(.)*		*(-2.35)*		*(1.83)*		*(-1.57)*		*(-1.71)*		*(-0.71)*	
Δ B2C Calls		0.106+		0.0891*		-0.114		0.091+		0.129*		0.003		0.001+
		(1.89)		*(1.96)*		*(-0.34)*		*(1.73)*		*1.81*		*(0.13)*		*(1.81)*
Δ B2B Calls		0.194**		0.22**		-0.088		0.503+		0.807+		0.0101		0.008
		(2.11)		*(2.71)*		*(-1.02)*		*(1.69)*		*1.64*		*(1.41)*		*(1.22)*
Δ Internal Calls		0.04		0		-0.24+		0.266*		-0.288		0.006		0.00
		(1.11)		*(.)*		*(-1.77)*		*(1.86)*		*(-1.33)*		*(0.16)*		*(1.02)*
Controls														
Leverage and Size	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm and Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	140	75	71	21	278	96	321	188	282	103	78	49	78	49

t statistics in parentheses
+p<0.1, *p<0.05, **p<0.01

Table 12: Likelihood of a larger financial outcome versus the previous quarter as a function of log calls and increase in previous quarter's log of calls.

calls would relate not to additional revenue sources, but rather, lower cost ways to provide previous services. Meanwhile, use of internal APIs might support scaling of internal operations in ways that facilitate growth and support external participation. Both Amazon's Web Services (Huckman et al., 2012) and the NY Times customer services began this way (Jacobson et al., 2011).

6 Discussion and Conclusion

We gathered a unique data set involving more than 120 firms that opened APIs and paired this with measures of firm performance. Using a difference-in-difference approach, we examined whether opening an API predicts performance.

Our results show that firms adopting APIs see increases in sales, net income, market capitalization, and intangible assets. API use also predicts decreases in operating costs in some specifications. The extent to which API adoption is linked to this outcome is sensitive to the econometric specification. Taking a difference in difference of sample means, we found that API adoption was significantly related to increases in net income and operating income. In our regression analysis, we find the most significant relationships to be between API adoption and market value. This finding is robust to the inclusion of time and firm fixed effects, the inclusion of leverage and installed capital as covariates. It is also robust to changes in the selection of control group. Restricting our attention to the regressions using only eventual adopters, we also find significant effects of API adoption on intangible assets and reduced costs. Curiously, net income is not found to be significantly influenced by API adoption in the regression analysis.

We also find strong positive relationships between the intensity of API usage and financial outcomes.

Results linking API usage intensity to outcomes should not be taken as causal relationships however. Endogeneity concerns loom as one can imagine, for example, firm specific shocks to demand increasing both calls and net income. However, firms which have given out more developer keys, either internal or external, have more contemporaneous calls. Firms have direct control over the number of API keys they distribute. This suggests that the number of calls are not purely a function of financial outcomes. These results support the main theses of this research. One surprise finding showed that internal APIs predict increases in sales more strongly than B2C APIs.

The size of the impact of API usage we estimated is large. In our preferred specification, binary API adoption leads to a 12.7 percent increase in market capitalization. While such a large outcome is perhaps not out of the question, it does suggest that an important latent factor drives both the timing of API adoption and market capitalization.

API adoption is a particularly attractive investment for firms in the information technology sector. This is unsurprising, as data management constitutes most of this industry's costs. Other industries, however, see gains to operating revenue after API adoption as well. Similarly, having an open developer portal seems to increase the returns to having an API, as does having APIs with several different orientations (e.g. B2B, internal, B2C) but these are not necessary for witnessing gains.

References

Adner, R. and R. Kapoor (2010). Value creation in innovation ecosystems: How the structure of technological interdependence affects firm performance in new technology generations. *Strategic management journal 31*(3), 306–333.

Ashenfelter, O. (1978). Union relative wage effects: new evidence and a survey of their implications for wage inflation. In *Econometric contributions to public policy*, pp. 31–63. Springer.

Baldwin, C. and K. Clark (2000). *Design rules: The power of modularity*, Volume 1. The MIT Press.

Benzell, S., L. Kotlikoff, G. Lagarda, and J. Sachs (2016). Robots are us: Some economics of human automation. *NBER Working Paper*.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in. *China's mortality decline under Mao 53*.

Bogers, M., A.-K. Zobel, A. Afuah, E. Almirall, S. Brunswicker, L. Dahlander, L. Frederiksen, A. Gawer, M. Gruber, S. Haefliger, J. Hagedoorn, D. Hilgers, K. Laursen, M. Magnusson, A. Majchrzak, I. McCarthy, K. Moeslein, S. Nambisan, F. Piller, A. Radziwon, C. Lamastra, P. di Milano, J. Sims, and A. Ter Wal (2016). The open innovation research landscape: Established perspectives and emerging themes across different levels of analysis.

Boudreau, K. (2010). Open platform strategies and innovation: Granting access versus devolving control. *Management Science 56*(10), 1849–1872.

Boudreau, K. and A. Hagiu (2009). *Platforms, Markets and Innovation*, Chapter Platform rules: Multi-sided platforms as regulators, pp. 163–189. Edward Elgar Publishing Limited, Cheltenham, UK.

Brynjolfsson, E. and L. M. Hitt (2000). Beyond Computation: Information Technology, Organizational Transformation and Business Performance. *Journal of Economic Perspectives—Volume 14*(4), 23–48.

Card, D. and A. B. Krueger (2000). Minimum wages and employment: a case study of the fast-food industry in new jersey and pennsylvania: reply. *The American Economic Review 90*(5), 1397–1420.

Chay, K. Y., P. J. McEwan, and M. Urquiola (2005). The central role of noise in evaluating interventions that use test scores to rank schools. *The American Economic Review 95*(4), 1237–1258.

David, H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of Labor Economics, January*.

Eilhard, J. and Y. Ménière (2009). A look inside the forge: Developer productivity and spillovers in open source projects. *Available at SSRN 1316772*.

Gawer, A. and M. A. Cusumano (2014). Industry platforms and ecosystem innovation. *Journal of Product Innovation Management 31*(3), 417–433.

Hastings, J. S. and R. J. Gilbert (2005). Market power, vertical integration and the wholesale price of gasoline. *The Journal of Industrial Economics 53*(4), 469–492.

Huckman, R., G. Pisano, D. Chen, and L. Kind (2012). Amazon web services. *Harvard Business School Case 9-609-048.*

Iyer, B. and M. Subramanian (2015, January). The strategic value of apis.

Jacobson, D., G. Brail, and D. Woods (2011). *APIs: A strategy guide*. O'Reilly Media, Inc.

Joseph, S., V. Ludford, and B. McAllister (2016, September). Plugging in: Enabling the enterprise for the platform economy. Technical report, Gartner Research Board.

Lahart, J. D. (2016, May). Low-yield blues? corporate bonds are the last ones paying.

Mora, R. and I. Reggio (2012). Treatment effect identification using alternative parallel assumptions.

Mukhopadhyay, T., S. Kekre, and S. Kalathur (1995). Business value of information technology: A study of electronic data interchange. *MIS Quarterly 19*(2), 137–156.

Parker, G., M. Van Alstyne, and X. Jiang (2017). Platform ecosystems: How developers invert the firm. *MIS Quarterly (forthcoming)*.

Parker, G. and M. W. Van Alstyne (2016). Innovation, openness, and platform control. *Management Science Forthcoming*.

Parker, G. G., M. W. Van Alstyne, and S. P. Choudary (2016). *Platform Revolution: How Networked Markets Are Transforming the Economy—And How to Make Them Work for You.* New York: WW Norton & Company.

Porter, M. E. and S. Stern (2000). Measuring the "ideas" production function: Evidence from international patent output. Technical Report Working Paper #7891, National Bureau of Economic Research.

Programmable Web (2016). API Research.

Purvis, R. L., V. Sambamurthy, and R. W. Zmud (2001). The assimilation of knowledge platforms in organizations: An empirical investigation. *Organization science 12*(2), 117–135.

Rowan, R. (2011). Steve Yegge's Platform Rant.

Saunders, A. and E. Brynjolfsson (2016). Valuing it-related intangible assets. *MIS Quarterly, Forthcoming*.

Tiwana, A., B. Konsynski, and A. A. Bush (2010). Research commentary-platform evolution: Coevolution of platform architecture, governance, and environmental dynamics. *Information Systems Research 21*(4), 675–687.

Yoo, Y., R. J. Boland, K. Lyytinen, and A. Majchrzak (2012). Organizing for innovation in the digitized world. *Organization Science 23*(15), 1398–1408.

Zhu, F. and Q. Liu (2016, February 4). Competing with complementors: An empirical look at amazon.com. *Harvard Business School Working Paper 15-044*.

A API Orientations and Functions

Using our proprietary dataset from Apigee, we categorized APIs by their orientation and function. The orientation of an API is determined by whether its primary purpose is for interactions with consumers, internal employees or systems, or other businesses. We divide functions, illustrated in the table below.

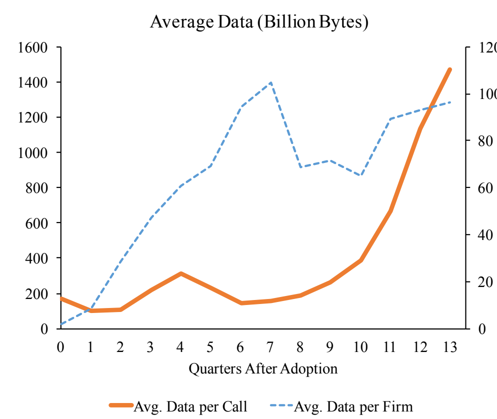
- Account Information: APIs related to storing, retrieving and displaying users' profiles
- Login/Authorization: APIs authenticating users and allowing information to be securely shared with other platforms
- Logistics/Inventory: APIs related to recording, managing and optimizing logistical items and inventory flow such as order delivery
- Maps/Locations: APIs dedicated to maps and GPS platforms, often Google Maps.
- Marketing/Customer Insights/Analytics: APIs related to storing and/or analyzing customer behavior or advertising information
- Media: APIs related to accessing, displaying or linking news or social media content
- Monitoring/Data Traffic Management: APIs related to collecting and managing data traffic

- Sales: APIs related to consumer purchases, especially online shopping
- Test: Any API named a variation on 'test' as well as any other API used for conducting tests of the platform performance
- Internal Communication: APIs for internal communication between employees
- Information – Other: APIs storing and providing information unrelated to the aforementioned categories
- Category X: APIs performing some technical internal function task unrelated to the aforementioned categories

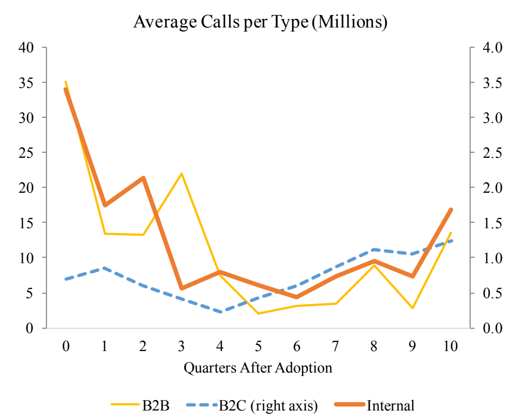
Many APIs have names which directly point to their functions, such as "sales" or "login" APIs. For the APIs with more technical names we searched for their descriptions online, often on developer websites. After classifying hundreds of APIs manually, we noticed patterns between API names and their corresponding functions. Therefore, we were able to identify and use certain keywords to partially automate the categorization. All automatic categorizations were double checked by hand.

At times, it was not so straight-forward to determine which function corresponded to an API. For example, many APIs such as "Pingdom" performed tasks falling in both the Monitoring and Test categories. Similarly, some Marketing and Sales APIs could be classified as either type. We used our best judgment to classify these ambiguous cases.

B Additional Tables



(a) Data and Developers



(b) Calls per Type

Figure 9: Data and Calls After Adoption. Calls of each type made by the average firm. as a function of quarters since adoption. In (a) data per firm is in billions of bytes, while data per call is in megabytes (thousands of bytes). A small subset of calls are not assigned to any of these categories. B2B is on the left axis, others on the right axis. This figure does not restrict to a constant sample of firms.

Table 13: Equivalence Test for Time Fixed Effects and Time Trends

Fixed Effects Regression: Trends and Time Fixed Effects						
	(1) Net Income	(2) Oper.- Income	(3) Market Cap	(5) Sales	(6) Costs	(7) Intangibles
i. Adopters Sample						
Firm FE and Trends	0.00721	0.0346*	0.0395+	0.0288	-0.153**	0.107*
t - statistic	*0.38*	*1.91*	*1.53*	*0.39*	*(-2.24)*	*2.08*
Ho: β_{FE}-β_{Tr}=0	0.099	0.034	0.015	0.043	0.053	0.067
Additional Controls	0.00301	0.0245	0.0981**	0.345**	-0.422**	0.159**
t - statistic	*0.15*	*1.1*	*3.77*	*5.68*	*(-7.18)*	*3.05*
Ho: β_{FE}-β_{Tr}=0	0.017	0.002	0.000	0.016	0.053	0.000
ii. Matched Sample						
Firm FE and Trends	0.0584	0.0069	0.0441	0.504	-0.351	0.0148
t - statistic	*0.03*	*1.02*	*1.18*	*1.4*	*(-1.31)*	*0.27*
Ho: β_{FE}-β_{Tr}=0	0.000	0.000	0.018	0.031	0.071	0.021
Additional Controls	0.222	0.329	0.0594**	0.674	-0.403	0.00382
t - statistic	*1.21*	*1.01*	*2.67*	*0.52*	*(-1.374)*	*1.375*
Ho: β_{FE}-β_{Tr}=0	0.032	0.000	0.000	0.000	0.053	0.000

+ p<0.10, * p<0.05, ** p<0.01

Table 14: Financial info for firms in the sample. All outcomes are in millions except for percentage changes, which are year on year growth rates. Net income and other flow variables are yearly.

Year	Variable	Firms	Mean	Std.Dev	Min	Max	Year	Variable	Firms	Mean	Std.Dev	Min	Max
2010	Sales	93136	18.30	313.55	0.00	34676.00	2014	Sales	110126	36.84	539.37	0.00	54804.47
	Cost of Production	95998	14.96	264.03	0.00	29191.00		Cost of Production	113776	31.52	472.04	0.00	53520.98
	Gross Profit	103487	4.65	88.88	-447.37	9572.00		Gross Profit	122413	7.88	130.03	-1011.09	13769.00
	Operating Revenues	79038	2.45	48.34	-747.88	4306.00		Operating Revenues	101854	3.22	62.74	-1934.65	5937.60
	Net Income	113293	2.28	53.54	-7074.93	5561.90		Net Income	133656	3.07	75.63	-2346.33	7704.97
	Market Value	107445	57.44	769.38	0.00	58689.00		Market Value	137218	106.35	1497.00	0.00	120324.00
	Intangible Assets	86687	5.27	117.96	-0.40	10155.77		Intangible Assets	106320	11.72	231.22	-0.01	26160.43
2011	Sales	96989	29.77	445.06	0.00	43266.00	2015	Sales	107705	38.14	550.28	0.00	63717.45
	Cost of Production	100099	25.07	383.03	0.00	37082.00		Cost of Production	111356	32.37	484.23	0.00	62506.90
	Gross Profit	108564	6.48	107.78	-790.84	11176.00		Gross Profit	120094	8.79	144.44	-2867.91	15358.04
	Operating Revenues	82401	3.40	57.74	-953.17	4901.00		Operating Revenues	99460	3.38	78.84	-3174.28	15536.17
	Net Income	118073	2.84	58.58	-1638.98	4853.00		Net Income	131357	2.90	74.93	-2296.25	6980.96
	Market Value	120958	77.75	997.17	0.00	75838.00		Market Value	137308	119.15	1682.20	0.00	126610.00
	Intangible Assets	91652	7.07	129.49	-0.47	13724.76		Intangible Assets	106904	14.23	268.34	0.00	28863.64
2012	Sales	100925	31.58	486.81	0.00	49566.00	2016	Sales	53076	38.81	536.71	0.00	46321.00
	Cost of Production	104186	26.78	419.59	0.00	42563.00		Cost of Production	54439	32.87	463.16	0.00	40322.00
	Gross Profit	113329	6.65	120.65	-8132.35	12480.00		Gross Profit	62158	8.84	139.18	-934.09	12598.00
	Operating Revenues	86251	3.32	63.06	-825.18	5381.00		Operating Revenues	47595	3.75	62.71	-1019.37	4118.00
	Net Income	123180	2.89	65.50	-2150.68	5512.39		Net Income	63911	3.43	72.75	-1123.67	6245.49
	Market Value	125180	86.93	1172.63	0.00	89814.00		Market Value	67089	131.83	1902.94	0.00	147089.20
	Intangible Assets	96163	8.52	159.80	-0.63	18605.59		Intangible Assets	53419	15.68	301.27	0.00	31415.92

Table 15: Simple Leverage Regression, Arellano-Bond Estimation

Leverage	Coefficient	Robust Std. Err.	Z	P-val	[95% Conf.	Interval]
Leverage (t-1)	0.202	0.111	1.810	0.070	-0.017	0.420
Adopt	4.172	2.411	1.730	0.083	-10.319	1.976
Constant	56.819	54.358	1.050	0.296	-49.722	163.359

Table 16: Change in Leverage Ratio, FE Regression

Leverage	Coefficient	Robust Std. Err.	t	P-val	[95% Conf. Interval]	
Adopt	0.555	0.347	1.600	0.110	-0.125	1.235
Constant	-148.635	0.000	-510.000	0.000	-148.635	-148.634

Table 17: Leverage Ratio Statistics for Adopters

Year	Mean	Std. Dev.
2011	2.32	4.04
2012	2.43	2.55
2013	2.49	2.71
2014	3.08	7.69
2015	2.78	3.07
2016	4.07	10.12