Can incumbent firms launch successful platforms? The interplay of data and power in adding platforms to existing product businesses

Abstract:
Platforms are increasingly becoming a dominant organizational form across sectors. Yet research into how an incumbent firm can develop a platform alongside its existing product or service businesses remains limited. To address this gap, we conducted an in-depth single case study of a large, product-oriented US firm in financial services as it introduced a platform to connect its customers with third party lenders. Our findings show that platform development in established firms is shaped by the interests of a diverse range of actors, including the platform sponsor (the organizational unit that leads platform development), other organizational units dedicated to existing product models, distinct coalitions of complementors, and both new and existing customers. These interests shape, and are shaped by, the data strategy of these distinct groups of actors, resulting in a dynamic process model. Our findings contribute to research on platforms, data strategy, and digital transformation, and carry practical implications for how firms with established products and customers can adopt platform business models.

Keywords: platforms, incumbents, data, power, single case study
INTRODUCTION

Platforms have emerged as a dominant organizational form in the 21st century. Platform-based companies, such as Amazon, Alibaba, Uber, Deliveroo, and Airbnb, have generated significant economic value by leveraging technology platforms for business models that connect actors in new and efficient ways, with a profound impact on the global economy and society. Through the provision of infrastructure and rules (Parker et al. 2016; Van Alstyne et al. 2016; Gawer, 2021) platforms bring market participants together, allowing for transactions and innovation. This way, platforms can generate significant value for complementors (Boudreau, 2010), customers (Parker et al., 2016), and the organizations that manage the governance of the platforms, commonly known as ‘platform sponsors’ (Boudreau and Hagiu, 2009; Adner and Kapoor 2010, Gawer, 2021). For many incumbent organizations, developing platform capabilities has become a central digital transformation challenge as they seek to add platform sponsor positions to their business portfolios (Hess et al., 2016; Majchrzak et al., 2016; ).

Extant studies on platforms establish a central role for data in mediating platform interactions, enabling innovation, and incentivizing platform participation (Boudreau, 2010). They emphasize a digital platforms’ role as a “conduit of data” (Gawer, 2021; 3), facilitating interactions between actors on both sides of a market through information exchange. From the viewpoint of platforms as a conduit of data, a critical challenge for researchers is to understand how platform sponsors aggregate, share, and use data to foster platform growth. Customers’ attitudes towards their own data are shaped by a variety of concerns about trust and privacy (Barrett et al., 2016; Acquisiti et al., 2016; Lin, 2022) and the usage of their data (Athey et al., 2016). Similarly, complementors’ attitudes towards sharing data with emerging platforms may

---

1 For the remainder of this paper, we use the term 'platform sponsor' to refer to organizations that create value by designing architecture and governance to enable interactions between different platform participants, such as complementors and end users (Boudreau & Hagiu, 2009). In contrast, participants like complementors and end users are focused on transactions enabled by the platform infrastructure. As ecosystem architects, platform sponsors have the opportunity to shape the scope and types of interactions between other participants through their governance choices (Rietveld et al., 2019).
be shaped by concerns about trust (Bhargava et al., 2020) and competition, and especially whether their data is being used in ways that aid competitors on the platform (Bhargava et al., 2020). Some scholars have theorized that customers and complementors would be more likely to share data on established platforms (versus emerging ones), since the benefits of engaging with an established platform are higher (Krämer et al., 2019), and concerns about trust are lower (Barrett et al., 2016). This raises a potential chicken-and-egg scenario for emerging platforms: data sharing by customers and complementors is both a prerequisite for and consequence of platform growth. Further research is needed to understand how a platform’s early decisions, particularly regarding data, may shape third-parties’ actions, and also how third-parties’ actions may shape the data strategy of a platform.

We argue that these issues can be further complicated when the platform sponsor is an incumbent firm that develops a platform business alongside existing product businesses. The development of platform capabilities has become a central topic in the digital transformation literature (Hess et al., 2016; Majchrzak et al., 2016). The rise of platform-based businesses has led firms to re-evaluate value propositions of standalone products; introducing platform models can enable them to harness innovative capabilities of complementors (Cennamo and Santaló, 2013) and potentially build network effects (Afuah 2013; Katz and Shapiro 1994). Extant research has emphasized that a large customer base or existing relationships with suppliers can provide a good stepping stone for launching a platform business (Shapiro and Varian 1999; Park 2004). What is less present in literature, however, is that in these contexts, platform sponsors must make decisions about the aggregation, sharing, and use of data, which may require navigating internal power dynamics between new and existing business units regarding which data will be used. In addition, an incumbent firm may also need to redefine relationships with existing customers and suppliers with whom they may have done business in set ways for a long time. Extant research has fallen short of understanding how firms introducing platforms...
alongside their product or service businesses can develop strategies for managing these critical
dynamics to jumpstart network effects with an emerging platform. Understanding these
dimensions is essential for organizations to successfully adapt their business models and
capitalize on the benefits of platform-based strategies. It also sheds a much necessary light on
a path of digital transformation for many incumbent firms.

To address this gap, we conducted a qualitative, 4-year longitudinal, field-based case study
of a product-oriented U.S. firm (Fincorp - a pseudonym) as it developed and introduced
a platform-based business to operate alongside its existing financial product offerings. Fincorp
is a leading provider of financial software for SMEs. In 2012, driven by competitive pressures
and recognizing the potential for higher revenues, Fincorp began developing a platform to
connect its large SME customer base with third-party lenders. Based on semi-structured
interviews with 89 individuals who had held senior management roles within Fincorp and at
complementor organizations during the relevant period (2012–2016) as well as numerous
archival documents, we developed a process model for how firms develop and deploy a
platform business model alongside existing products or services, and in particular, how they
manage existing data and internal/external stakeholders in this process.

Our findings identify four emergent challenges in Fincorp’s platform development, each
around how data is aggregated, shared, and used by the platform sponsor, complementors, and
end customers. How the firm navigated these challenges using power internally and externally
was central to both developing platform capabilities and shaping the behavior of complementors
and customers. Our study also emphasizes the multifaceted and dynamic role for data as a
resource, with different types of data varying in degree of utility to the various actors involved.
Our findings on how an incumbent firm can manage data heterogeneity have important
implications for competitive strategy and how incumbent organizations can adapt to the
platform economy.
In the following section we outline two streams of recent research emphasizing the agency of platform sponsors, and the role of data in platform governance. We argue that these literatures provide a lens through which to explore how incumbent organizations develop a platform business alongside their existing product or service businesses.

THEORETICAL BACKGROUND

Early platform research, inspired by the industrial organization tradition, understood platforms as a form of market – typically facilitating transactions between customers and complementors (Rochet and Tirole, 2003; Parker and Van Alstyne, 2005; Rysman, 2009). Actors can benefit from the emergence of network effects, where the utility of participation in the marketplace increases with the number of users on the opposing side (indirect network effects) or the same side (direct network effects) of the market (Katz and Shapiro, 1994). This industrial organization perspective emphasized the role of market forces, such as pricing and scarcity, in accounting for the growth of both platforms and, by extension, network effects for users (Parker and Van Alstyne, 2005; Evans, 2009). By eschewing the notion of platforms as a distinct organizational form (Saadatmand et al., 2019), this literature developed a limited perspective of platform sponsors’ agency, or their role in shaping customer and complementor engagement (Gawer and Cusumano, 2014). Adjacent work has begun to develop an organizational perspective, placing a greater emphasis on the decisions made by platforms and complementors that shape platform development (Kretchmer et al., 2020). Under this perspective, platforms are the result of intentional organizing efforts both within and between firms: “platforms…do not simply emerge without deliberate, firm-driven agency or deliberate managerial decisions and actions” (Gawer and Cusumano, 2014, 420).

Platform Development
Recent research has examined strategies that platform sponsors employ to drive engagement from other actors (customers and complementors) when developing new platforms (Boudreau & Jeppesen, 2015). A key success determinant for new platform businesses is the ability for the sponsor to jumpstart network effects by creating distinctive value propositions tailored to specific user groups that benefit by connecting to each other (Eisenmann, Parker, & Van Alstyne, 2011). Encouraging users to invite contacts and incentivizing early adopters through discounts and referral programs have been suggested as key strategies for increasing numbers of users and building critical mass (Katz & Shapiro, 1985).

For digital platforms, research has begun to emphasize the role of data gathering and utilization from platform participants as critical in driving platform growth (Saadatmand, Lindgren, & Schultze, 2019). Gawer (2021) proposes that transaction platforms operate as “conduits of data” (2021, 3), facilitating interactions by underpinning algorithmic matchmaking and providing information about each side of the market to the other. Platforms can collect a vast amount of data from their users, including demographics, behavioral patterns, and user preferences, which can be analyzed to improve user experiences and enhance platform algorithms (Chen et al., 2022). Furthermore, data access can fuel innovation by complementors. Boudreau (2010) explores how opening up digital systems that underpin platforms, providing access to rich and shareable data, can facilitate innovation from complementors. For instance, by allowing third-party developers to access APIs, platforms such as Google Maps have offered new features and services to users (Gillespie, 2018).

Given the critical role of data in fueling network effects, the success of emerging platforms relies heavily on their ability to obtain and leverage shareable data from both customers and complementors. However, gaining access to such data can be challenging as it requires engaging with actors with a wide range of interests and concerns. Emerging platforms face a two-sided "chicken and egg" problem regarding participant data sharing. Customers may
hesitate to contribute their data due to privacy concerns (Lin, 2022) and uncertainty over how it will be utilized (Athey et al., 2016). Similarly, complementors may hold back data over fears it could benefit rivals participating on the platform or the sponsor itself (Bhargava et al., 2020). However, substantial data sharing is also a prerequisite for growth, as it provides the necessary inputs for improving platform services and demonstrating network effects to attract more participation (Krämer et al., 2019). This creates an interdependent dynamic where data sharing enables growth, and growth encourages further sharing. Platform sponsors of nascent platforms thus need to pursue strategies to overcome early participant reluctance.

**Platform development alongside existing product or service businesses**

While the topic of platform development and the associated strategies to create and manage platforms are gaining traction among scholars, many studies focus on firms that start out as platforms. However, established product or service-oriented firms are increasingly recognizing that the value they once derived from offering standalone products can be significantly enhanced if they develop and integrate platform-based businesses into their existing businesses and harness the innovative capabilities of complementors to provide additional value to their customers (Gawer and Cusumano, 2008).

While the role of competitive dynamics, complementor interests, and actors’ data strategy have been studied in the context of early ventures creating standalone platforms from scratch, this has not been studied in the context of firms building a platform business alongside existing products or services (Hagiu & Altman, 2017). This presents a puzzle for the literature, as developing a platform alongside an existing business has distinct implications for how relationships with both new and existing customers and complementors are developed and potentially adapted, as well as how data is aggregated, shared, and used. Simply put, for an incumbent firm, developing a platform requires dealing with customers and suppliers as well as their data in new ways.
One of the main challenges for such firms is to shift focus from product development and sales to creating and managing network effects (Cusumano & Gawer, 2002). Additionally, they may need to shift from focusing on one particular type of end user customer to balancing needs of both end users and also complementors that serve them (Hagiu & Wright, 2020). As firms make these changes, they not only need to modify product or service offerings, but also need to modify activities that support these offerings (Hagiu and Altman, 2017; Parker, Van Alstyne, & Choudary, 2016).

More research is needed to understand various dimensions of the process of launching and managing platform business models within incumbent firms (Parker, Van Alstyne, & Choudary, 2016). For instance, existing relationships and associated power dynamics of the firm, such as with established customers or suppliers (which could become complementors), need to be rethought and remanaged as the firm supplements its role as a buyer and seller with a new role as an enabler of transactions (Hagiu & Wright, 2015). Moreover, the addition of a platform business may occur within one business unit of a larger firm, in which case existing internal relationships and power dynamics with other units may come into play affecting the success of such initiatives, particularly when the platform business has the potential to add new competitive dynamics to existing businesses.

When adding a platform business, an incumbent firm must also consider how existing and new data are handled. Choices regarding data are particularly significant as the firm may already have relevant data - an advantage for launching a platform - but may need permission from existing customers or suppliers to use it in a platform model. Further empirical research is needed to understand how these internal and external challenges are navigated in established organizations looking to develop a platform alongside existing product or service businesses. To address these critical gaps in the literature, we ask the following research questions: How
do firms start a platform alongside their existing product or service business? How do they manage existing data and internal/external stakeholders in this process?

CASE SETTING AND METHODOLOGY

The issues above invite an inductive, longitudinal approach (Khanagha et al., 2014; Tripsas and Gavetti, 2000) to explore an incumbent’s efforts to develop a platform alongside an existing product-oriented business. This single case was selected for its revelatory quality (Eisenhardt and Graebner, 2007; Ozcan, Han and Graebner, 2017), and provides two central advantages for an empirical study. First, it functions as a specific example of the challenges facing early-stage platforms, allowing us to contribute both to theory about product-to-platform transitions and integrations, but also to questions about platform development more broadly. Secondly, by focusing our research on an incumbent firm developing a platform, we were able to study a case of platform development from its earliest stages from the perspective of one of the central firms that shape its development. This overcomes one of the prevailing problems in empirical research where researchers predominantly study successful cases of emergence due to the availability of data and access, thus creating a selection bias in the literature (Aldrich and Fiol, 1994; Ozcan and Santos, 2015).

Research Setting

We conducted a multi-year field-based, qualitative, longitudinal study of an incumbent, product-oriented firm as it introduced a new platform business model in parallel with its core business. Fincorp (a pseudonym) had established itself as a global leader in financial software for SMEs. Motivated by the possibility of capturing network effects, Fincorp began the process of developing a variety of new platform efforts including a platform model in 2012 with the
aim to connect a select subset of their SME customers with outside lenders to allow the two sides to transact via Fincorp.

**Data Sources**

To improve accuracy and validity, we triangulated our data across multiple discrete sources, collected between 2012 and 2016. We prioritized semi-structured interviews with senior management of the firm and with related experts, along with content analysis of public archival records. These sources are detailed in Table 1.

We conducted interviews with key stakeholders involved in Fincorp's platform development between 2012-2016. Our informants included 82 executives from Fincorp who held senior management roles during the platform launch, as well as 7 executives from complementor organizations. The Fincorp informants represented a range of functions, including technology infrastructure, partner strategy, service delivery, pricing, and product management, and members of the senior executive team. This enabled us to gain insights into diverse aspects of the platform's development. The complementor informants provided an external perspective.

The semi-structured interviews ranged from 60-90 minutes. In the first section, we collected background information on the informant's firm and role. In the second section, we asked open-ended questions for informants to describe major events and phases in the platform's evolution from their viewpoint. In the third section, we focused on the current state of the platform, including strategic priorities, challenges, and successes. Interviews were in-person with audio recording and transcribed by a human transcriptionist. This interview approach combining open
narrative sections with more structured questions enabled us to gather rich, contextual data on the critical issues shaping the platform's trajectory over time from diverse internal and external perspectives. The insights from these informant accounts formed the foundation of our process model.

We addressed potential retrospective and informant biases through several strategies. We collected data in multiple waves between 2012-2016, enabling both real-time and retrospective accounts. The retrospective data efficiently gathered observations across the full platform timeline, while the real-time data reduced retrospective bias (Leonard-Barton, 1990). We used open-ended questioning and pressing for specifics during interviews to elicit detailed, accurate accounts. Further, we interviewed informants at multiple organizational levels and functions to triangulate perspectives. By gathering multi-year data through diverse lenses, we aimed to develop the fullest understanding of the issues shaping the platform's trajectory over time.

To ensure robust, well-corroborated findings, we triangulated data with archival sources (Jick, 1979). We reviewed a total of 76 archival documents from Fincorp, including investor reports, transcripts of speeches made by senior management, regulatory filings, and AGM transcripts. Here, we used written commentary and discourse to build an emergent understanding of the issues shaping the development and integration of an adjacent platform. We used archival third-party commentaries from the period (including analyst reviews and industry reports) to understand the contextual ambiguity in the industry (Barr, Stimpert, and Huff, 1992).

Our work combines real-time and retrospective sources to build a rich understanding of the forces that shaped the development and integration of one of Fincorp’s platform business. This builds on a growing body of research using “history informed strategy research” (Argyres et al., 2019: 345) to understand organizational decision making through incorporating retrospective data (Wadhwani et al., 2020). Historical approaches provide a broad explanatory
power in “revelatory” (Eisenhardt and Graebner, 2007) cases, allowing for salient elements of unique and rare contexts to be drawn out and supporting the development of “more informed causal inferences and theories…supporting analyses of path dependence” (Argyres et al., 2019: 345). As Burgelman (2011) notes, “historical methods are inherently concerned with longitudinal development and involve reconstructing the unfolding of individual and collective action patterns leading up to relatively unique events” (2011: 594). **Data Analysis**

While our work draws on existing literature to develop the initial research questions and identify relevant theoretical concepts; our data collection and analysis is deliberately interpretive. We followed a structured, inductive approach to analysis (Glaser and Strauss, 1967; Gioia, Corley, and Hamilton, 2012), employing techniques intended to develop theory “grounded in the views of participants in the study” (Creswell, 2012: 14).

We first organized the data collected through interviews and archival sources into first-order concepts (Gioia et al., 2012: 21). These concepts (for example, ‘codebase integration’, ‘challenges with legacy infrastructure’) were built by collating codes around topics in the data. These concepts were then grouped into second-order themes (Strauss, 1987) based on their topic and perspective (e.g. ‘technical hurdles’, ‘reluctance to transact’). The second-order themes were then inductively grouped together in the penultimate stage of analysis to form common dimensions (e.g. ‘aggregating data across organizational units’, ‘complementor hesitancy to provide transaction data’) which capture the relationships between the second-order themes (Gioia et al., 2012, 20). These dimensions were clustered around distinct challenges in the aggregation, sharing, and usage of data in the development of the platform. An illustrative overview of this data structure can be found in Table 2. By combining these levels of analysis, we were able to develop a comprehensive understanding of the data.
With the central codes established, we sought to identify the distinct phases (Giddens, 1984) in the platform's development. This process of temporal 'bracketing' (Langley, 1999) started by pinpointing key changes in the development of the platform marked by technological achievements (e.g., obtaining data) and shifts in the actions of complementors, customers, and other organizational units. These points also signified the emergence of new themes, representing changes in the platform's capability and how other actors interacted with it. In the final stage of our analysis, we returned to the existing literature to compare it to the patterns that emerged from our analysis. This allowed us to refine our model and ensure our findings were generalizable (Yin, 2014).

**FINDINGS**

Our analysis revealed that there were four distinct phases for an incumbent firm to go through in their journey to establish a platform alongside their existing product business. As an overview, these phases and their characteristics are described in Table 3 below.

**Phase 1. Opportunity recognition and initial steps towards a platform model**

*Recognizing the platform opportunity.* When our fieldwork at Fincorp began, the senior managers at the firm stated that they had already spent ‘years’ (022, 2013) informally discussing the idea of a platform to connect their corporation’s small- and medium-sized enterprise customers (SMEs) with banks and credit unions to lend money through credit products. These
discussions were largely driven by the top management vision to “find a network effects business” (022, 2013) that would grow the number of interactions between their SME customers and a variety of other actors. A senior manager noted:

“We’re one of the few companies in the world that has that opportunity to be the global center of the ecosystem for small businesses and if we were to pull that off...it’s a much bigger opportunity than anything else” (011, 2013)

Fincorp’s SME customers had a consistent and high demand for credit but when they applied they were often rejected. This observation supported Fincorp’s case for developing a platform that would enable more efficient and timely approvals and transactions:

“So, a couple of years ago we did a survey...90 percent of all of our customers operate with some type of a credit facility. In the last two years...70 percent of all those applications were declined. So, what that basically tells you is nine out of ten small businesses need credit... [but] two out of three times that you apply for credit you’ll get declined.” (008, 2013)

This high-rejection rate and uncertainty about risk also created inefficiencies for lenders. It increased the cost of underwriting, and limited the pool of customers for whom it was viable to lend money to:

“Lenders acknowledge that it’s far from optimized. They acknowledge that it’s extremely expensive for them to underwrite. They acknowledge that if you want less than $100,000, if you have anything less than perfect credit, it’s just not worth their time looking through all of your docs. It’s just too expensive. The numbers don’t work out for them. So, what they do is, if you want less than $100,000, they don’t even look at your business.” (008, 2013)

Fincorp executives expressed a belief that by combining a rich body of data from SMEs with a channel to engage with them, they could enable lenders to make credit decisions more accurately and cost effectively. And, they could support lenders who were willing to make small loans very quickly, which was often an urgent need for SMEs.

Managers identified two mechanisms through which Fincorp could capture value from such a platform. The most direct way would be an “acquisition play” (013, 2013), allowing the
wider firm of Fincorp to attract more customers to their software product: “if we gain a reputation as the single best place to get a small business loan, but you have to use our software, obviously, that helps us acquire more customers” (011, 2013). Secondly, it would allow for direct monetization of matchmaking on credit products via the new platform unit. By reducing the level of risk in credit products through better data, it would also become possible to share the savings with lenders:

“Let’s say an average APR is six percent. They know they’re going to lose two and a half percent, so their spread is three and a half percent. But if we can take that two and a half percent just down to two percent, you’re talking billions of dollars in increased revenue for them; and so, we would want to share in that” (008, 2013)

Several managers who had previously worked in other product management roles elsewhere in Fincorp were transferred into a new platform unit starting in Q1 2012. By Q4 2012, this group had begun the process of signing up lenders with provisional agreements to use the Fincorp platform. In developing these early partnerships, the platform unit managers prioritized established lenders who could provide a reputational boost to the nascent platform, providing legitimacy and attracting customers and complementors. As one manager explained, “quality is everything” (008, 2013). The platform unit was able to develop initial interest amongst traditional lenders fairly quickly. In our interviews in Q1 2013, one manager explained: “We’re in the process of building this network of lenders. We have four on today, and another five or six committed” (008, 2013).

By the beginning of 2013, a nascent website was live that allowed customers to search for lenders, but provided no matching or on-platform transaction scope. While limited in functionality, this early site demonstrated the demand from customers:

“We’ve got a website…right now we have 25,000 customers a month coming to us looking for credit…We’re now in the process of building this network of lenders…we understand those two constituents, we want to connect them.” (008, 2013)
Challenges in aggregating customer data. While managers understood there was latent demand from both customers (SMEs) and complementors (lenders) for a more efficient credit market, facilitating this through a matching platform required lenders to have access to customer data: “the secret sauce is the data which makes it a durable competitive advantage” (008, 2013). With this data, lenders would be able to make decisions “more efficiently” (031, 2014) on the platform. Early efforts by the platform unit to achieve this ran into two related challenges to consolidate shareable data, first from existing customers (Data Challenge 1) and second, between business units (Data Challenge 2).

First, while Fincorp’s software collected large amounts of customer data, this data was not immediately accessible to Fincorp as it resided on customer’s own computers, and even in the small number of cases where it was available, it was not available to share with lenders (Data Challenge 1). For Fincorp to access individual customer’s data, customers had to regularly back up their data from their stand alone computers into Fincorp’s cloud storage service (at this time, the systems were not fully cloud-based; the customers needed to synchronize their data manually). Without this, the data remained out-of-reach: “that’s a big, heavy lift that my team’s been working on...when all of that data is locked on someone’s hard drive on their desk somewhere, who cares, right?” (013, 2013). Even when data was backed-up from desktop products, it was often done at infrequent intervals, resulting in a patchy and incomplete picture of the financial health of the SME: “…not all of our customers’ data is accessible” (002, 2014).

Even once the platform unit gained access to the data, it would only be able to grant data access to complementors once the Fincorp customers had consented to share their data. While Fincorp sought permission through in-product messaging, customers only granted permission in a minority of cases. This appeared to take platform unit managers by surprise, as many
believed that being connected to lenders would give sufficient motivation to SMEs to engage with the platform:

"They’re coming to us with a problem, we’re saying, ‘We think we can help you, but first we need to share some of this (data) with lenders and we need to get your consent for that’ and a minority say ‘Okay.’ That clearly has taught us that while on the one hand it’s a no-brainer to us; on the other hand, there is a sensitivity to people’s business files, and despite the fact that they’re coming to us, in spite of the fact that we’re trying to explain, ‘This is how we’re going to help you,’ there is a bridge there. We cannot assume that that bridge is easy for everyone to cross.” (008, 2013)

A second, parallel issue concerned the siloing of data between distinct business units within Fincorp (Data Challenge 2). In most cases, an individual SME worked with several services across the business, sharing unique data with each. Much of this data was relevant to credit decisions, but was neither consolidated centrally nor accessible to complementors. Interviews identified that this problem was fundamentally an organizational issue: “we get a bit siloed” (008, 2013) “it’s…not a technical one” (022, 2013). It was feasible for the platform unit to combine data from distinct architectures, but the reluctance of business unit managers to cooperate in any way that threatened their ownership of customer data slowed it down. Business unit managers seemed to be concerned about opening up data to complementors who could eventually offer services competitive to Fincorp. As one manager put it:

“...[our colleagues are] out fighting for their lives competing against other businesses...And we’re now saying, ‘We’re going to open this up to help our customers’ ...that’s a very uncomfortable conversation for a lot of people because they have two competitive advantages—a data advantage and channel advantage, that no one else has. Now, we’re going to potentially lose those advantages...I can’t prove what the upside is on the other side.” (013, 2013)

This concern was echoed by others across the group. One Fincorp software developer noted that “we cannot be a successful platform if, as a developer, we have nine different groups [with] data and I’ve got to go to each” (013, 2013). The challenge of working with other business units to access their data was exacerbated by the relative disparity in size between the platform unit and other business units within Fincorp. Platform unit managers characterized their unit as a
“startup, scrappy group…solving a gnarly problem” (031, 2015), operating “without a lot of resources” (002, 2013), to align far larger units “with their own priorities” (038, 2015) within Fincorp. This problem persisted throughout the first phase, in the words of a platform unit manager, “you cannot assign a king to the Balkans” (001, 2013).

These issues functioned as two distinct, interrelated challenges in data management. While a large customer dataset was necessary as input for complementors to transact on the platform, the Fincorp platform unit struggled to consolidate shareable data, both from customers (Data Challenge 1) and between business units (Data Challenge 2). Combined, these dynamics restricted the supply of data they could provide to lenders.

While the platform unit struggled to provide customer data to lenders, a pool of data was available to Fincorp internally. For customers who were actively considering credit products, the platform unit administered a 15-point customer questionnaire to supplement their existing data and better assess creditworthiness. This survey data made it possible for the platform unit to develop an early model for predicting which SMEs were good candidates to be matched with lenders. A platform unit manager explained:

“And so today, you answer a few questions and then we connect you to a lender that we think makes the most sense… then from that point on you go and kind of formally apply with the lender. The vision is the data that we have on the customer... all of that we can use behind the scenes to match customers” (002, 2013)

While Fincorp’s limited ability to share customer data with complementors prevented it from functioning as a platform, the data they had access to internally allowed them to operate in those initial months as a form of matchmaker - providing each lender with leads from their customer pool, while promoting credit products to screened clients.

“It's kind of a cold hand-off because we don’t pass data to the lender and so the customer has to start all over with the application there. The only thing is that the lender knows that this is somewhat of a qualified lead that they’re getting. Where we’re trying to head to is that being a much warmer transfer and us potentially owning more of that end-to-end experience.” (002, 2013)
This meant that matchmaking between customers and complementors could slowly get started in this transition period despite the data sharing challenges that prevented an actual platform from materializing.

**Phase 2. Resolving data access challenges to enable platform transactions**

*Streamlining data access.* Phase 2 was marked by concerted efforts by the platform unit to move away from this “cold leads” model and begin to function as a transaction platform. This was underpinned by two changes in the way that Fincorp stored and accessed their customers’ data to address Data Challenge 1. First, Fincorp launched a cloud-based browser version of its long-standing financial software. This online “native” (038, 2015) product bypassed local desktop storage and instead synced customer financial and company data (including data often used in underwriting, such as company revenue and EBITDA) directly to the cloud. The aim of this product revision was to create a repository of “live” (007, 2013) customer data, without any lags. For lenders, this data would more accurately reflect the current status of the potential borrower SME’s underlying business.

Second, Fincorp simplified the process through which Fincorp gained access to sharable SME data from the desktop product. Permission to sync locally stored data with the Fincorp cloud was now managed through an opt-out system. Under the new system, “the default [was] synced” (002, 2013). Fincorp’s ability to switch to this approach was partly due to its strong market share and thus relative power over a “completely fragmented market” (013, 2014) made up of small businesses. For these businesses, changing suppliers would require substantial changes to their financial administration, giving Fincorp a relatively high level of power to introduce changes: “there is some lock-in…if I’m already storing the documents you need in case you get audited, if I’m already creating payments for you, if I’m already sending your 1099’s out, it’s very hard for you to change” (013, 2015).
In parallel, the platform unit continued to work on consolidating SME data across various business units to address Data Challenge 2. During Phase 1, the platform unit had addressed this challenge through diplomatic “uncomfortable conversations” (013, 2013) with managers from the largely autonomous business units. The slow process of this approach led the platform unit managers to focus on influencing the top management instead:

“Literally, two hours ago [the CEO] and I were in a meeting with our CTO and we said we need to focus on that this spring with our engineers. We need to keep coming back to the network effect platform…” (003, 2013).

“We've been on this for years in terms of trying to influence the business units and...we’ve tried a couple of different routes. We started off with bottoms up, where we as individuals go work with the heads of the business units and try to convince them that an idea is a good idea...with not so good results. Probably our greatest success was convincing [CEO] and their staff that something was important. Then it became corporate strategy.” (016, 2013)

Platform unit managers expressed that one critical strategy they used in this phase was to demonstrate the potential impact of the platform to top management by using relatively small amounts of data. In their words, this led to the top management team - including the CEO - being “aligned” (016, 2013) and therefore taking a larger role in forcing business units to cooperate. A manager explained:

“So we showed them...we actually have been running experiments to prove that it’s predictive, and we’ve been doing matchmaking between our banks and our small businesses; and we’ve, already, through just a small pilot, done three million dollars’ worth of loans that are very positive, and that’s a way to use data that starts to create opportunities.” (003, 2014)

“We take that to the top management...show that it’s a source of durable advantage that has remained pretty elusive to us to this day, in fact, we don’t have a business that’s really capitalized on this today and we’re going to talk about how we’re going to do this...[CEO] helped evangelize this” (003, 2014)

One result of these demonstrations of the predictive utility of customer data was that top management commissioned the Chief Technology Officer to lead the development of shared technology projects. Much of this work focused on the development of a Financial Data Platform (FDP), integrating technologies from each of the business units at both a data and
service layer (004, 2014). This began a process of coordinating efforts across business units to engage consistently with third-parties in general. A business unit manager described:

“Our CTO mandated that we need to have standardized services that are exposed to both internal users as well as third-parties...[for example], if I go to developer.facebook.com I can get everything I need to build a Facebook app, and we wanted to make it that easy that any third-party developer should be able to go in, read the documentation, try it out, and then start culling those APIs.” (027, 2014)

**Interoperability issues in existing data.** While Phase 2 saw a marked increase in the amount of live SME data that Fincorp was able to share with lenders, technical challenges complicated efforts to create a single, accessible dataset for lenders. Managers found that the data from the old systems were not immediately interoperable with data uploaded through the newer cloud-based product. While both of these data sources included the same types of information (e.g. financial records, tax codes, and payments history), they differed in the ways the data was structured: “[the systems that collect the data] are distinct at a foundational level...the codebase is different” (002, 2014). This difference slowed efforts to attract complementors, as data from many long-term customers, who predominantly used the older desktop software, remained largely beyond reach.

“The flawed notion we kind of carried for a long time was this idea of the Desktop and Cloud products being equal, allowing us to say to complementors ‘the data is actually interoperable between them and we’ll allow you to use it.’ It did not work. So, Desktop is now off in kind of a separate box right now...it’s not a seamless integration.” (030, 2014)

Platform unit managers initially attempted to overcome interoperability challenges through workarounds. In these early stages, they formatted and pooled desktop customers’ data into a “data dump” every four to five months (027, 2014). They could then extract from here some of the missing information required by lenders to evaluate loan applications:

“We should be able to read their company ID...and then in our backend we had a dump of one million customers and their data like their company information, their annual sales, and stuff like that. So what we would do is if somebody comes in from the Desktop product...we would link them up in our backend and then do some
While this workaround provided a supply of some desktop customer data to lenders, the data was not live, requiring snapshot uploads. It also failed to support the end goal of the platform to match customers with credit products without having to begin an application: “we want to get all this data…match it with the lender criteria, and then show it to our customers without actually asking them any questions” (027, 2014).

**Complementor hesitancy to provide transaction data.** In this second phase, managers also began to recognize a reluctance from lenders to complete transactions on-platform. While lenders had begun using the platform to source leads, most of them completed the transactions on their own websites. This was in part due to technical challenges integrating on-platform transactions with their older “legacy” infrastructure and risk assessment tools (035, 2015) that many banks and credit unions used. In addition, many larger lenders showed a general reluctance to step away from their well-established processes.

“They were still manually underwriting it and it’s hard to change that process. I think banks also got very burned in the consumer meltdown phase for mortgages. So to say ‘I’m going to use technology and more boldly decision these small businesses’ is hard for them...it’s a little bit slower.” (031, 2015)

This reluctance from lenders to complete transactions on the platform created both commercial and technical challenges. From a commercial perspective, it meant that the platform still operated as a form of informal ‘matchmaker,’ rather than as a transaction platform. Platform unit managers’ vision of a seamless credit application from within the Fincorp financial software remained far distant (002, 2014). From a technical perspective, the lack of transactions on the platform meant that the platform unit had no access to data to understand the effectiveness of the matching they provided: “the data work was really geared at ‘can we build a predictive score based on the customer data that we have?’…The thing that we were lacking was a lot of outcome data.” (002, 2014). Completing transactions on-platform would
give Fincorp complete “transaction data”: names of customers accepted for credit products, their repayment timelines, default rate, etc. This dynamic marked a new obstacle in the platform’s development, in which a lack of transaction or outcome data limited the platform’s ability to refine its predictive model for matching customers and complementors (Data Challenge 3).

Without transactions taking place on the platform, Fincorp had to request post-transaction “outcome data” from lenders. This request was largely unmet, which revealed to the platform unit that lenders were not only being inertial but also competitive: they were hesitant to share information that might be useful to competitors such as the amount lent, structure of the arrangement, and repayment plan (008, 2014). For many lenders, sharing data that “could help competitors…wasn’t an arrangement that they wanted to be part of” (031, 2015). They were happy to use the large repository of SME data to pre-select customers, but preferred to complete the tractions off-platform. In the interest of growing the complementor side of the platform, Fincorp dropped sharing outcome data as a requirement for lenders wanting to use the platform:

“In the first couple of iterations of this we asked for quite a lot of integration with them. We now ask for no integration. In the first couple of iterations of this we asked for quite a lot of data sharing. We now ask for no data sharing. [We] realized that the thing we need to do differently is make the bar to entry just ridiculously low or, potentially, just zero; because we were creating objections by asking for things that made perfect sense to us that caused a different internal group at a bank to be involved.” (008, 2013)

However, this meant that Data Challenge 3 persisted: without outcome data, Fincorp could not refine its predictive model. This, in turn, restricted the platform’s value to new lenders: “without that data…we have no burden of proof to persuade them to use us” (008, 2013).

Phase 3. Improving platform data
Phase 3 was defined by a focus on encouraging complementors to more actively and extensively engage with the emerging platform to resolve existing data challenges. This required addressing gaps in SME data, and refining the predictive matching model even though they did not have (or had minimal) transaction / lender-provided outcome data.

**Building monolithic customer datasets.** At this point in time, the main financial software that Fincorp was selling still had both desktop and cloud versions. Fincrop’s management team decided that by late-2014, the base desktop product would no longer be available to new customers. To encourage a full transition to the cloud product, they would discontinue support for the Desktop product: “So we've actually sort of cut the cord, right… we will stop support of the Desktop sync manager” (007, 2015).

Fincorp’s transition to a cloud-only product helped the platform unit produce a single, monolithic customer dataset that complementors could use. This “critical mass” (028, 2014) of sharable SME data would provide a basis on which lenders could make decisions and transact with customers on-platform: “We’ve crossed a bunch of commercial milestones. And part of the secret is that…We’re now prepopulating applications for 95% of the population…allowing lenders to make decisions” (031, 2014). This provided a more sustained and satisfactory solution to the interoperability problems that had persisted in Phase 2, and helped to resolve Data Challenge 1.

**Acquiring proxy data to fill gaps.** To address Data Challenge 3 (i.e., lack of outcome data from lenders), the platform unit made a decision in 2014 to acquire third-party ‘proxy’ data on firm creditworthiness from a global consumer credit reporting firm. This data included credit scores and business-by-business breakdowns on the levels of credit each firm held, and of any missed payments. This data had the effect of “closing the loop” (006, 2014) in the platform’s predictive model. With proxy outcome data in place, they refined the predictive matching:
“We don’t know the outcomes of decisions, and so we turned to get a download of outcome data from decisions that other lenders had made over the past few years. We have used that data from [credit reporting firm] to build the predictive model.” (002, 2015)

This supply of third-party data allowed the platform unit to model which characteristics in the SME data were predictors of creditworthiness and future financial performance, therefore improving the accuracy of matchmaking on the platform. It also provided the ability to “prove” (028, 2014) to existing and new lenders that the model was accurate and insightful. The platform unit began to use this evidence to encourage lenders to engage directly with the platform, completing transactions exclusively using data from the platform.

**Offering new predictive capabilities to complementors.** Fincorp’s increasingly large dataset not only included data on the business operations of its SME customers, but also behavioral data, e.g., how often they used the software, which other products they bought, as below:

“I know which version you bought. I know how many times you logged in. I know how many vendors you’ve paid over the last six months. I know from a trending standpoint how many customers you’ve either increased or decreased over a five-year period… There is so much other data that I know about you that’s not in that … file.” (101, 2014)

For Fincorp, this data proved to be highly useful in predicting SME creditworthiness:

“How many times you have logged in a week for the last three years is predictive of your risk on default…you’d be surprised. There’s an inflection point. So, logging in a reasonable amount is a good thing relative to people that don’t log in very much. But logging in too much makes you more risky…there’s a correlation there with having an extremely limited cash flow…we can partner that with what product do you buy from us? What other products do you buy from us? How frequently do you upgrade? Believe it or not, whether you started with a trial version or not is actually predictive of risk.” (008, 2014)

The platform unit’s discovery of these patterns also came with the realization that this richer data did not fit lenders’ standard, traditional underwriting models. Instead, lenders mostly made decisions based on revenues and consumer credit bureau scores of founders:

“(There is) data such as revenue, EBITDA…the specific data that they need. There’s a lot more data in the platform…(Lenders) just want to focus on the
This underpinned a fourth data-related challenge: existing complementors could not integrate the different types of newly available platform data into their legacy decision making frameworks (Data Challenge 4). Lenders’ inability to integrate this data frustrated platform unit managers, who felt that the nascent platform’s true potential was not being realized by lenders, limiting their engagement with it: “we want to drive them to use all the business data. (We) truly feel like it’s more predictive” (031, 2015). Fincorp managers expressed concern that lenders’ inability and lack of motivation to use business-level data impacted the ability of the platform to deliver on its value proposition: more competitive prices and faster transactions than traditional SME-lender interactions:

“There is an unmet need, for a low-price product with low effort…by low-price we are thinking less than 12% interest, and low effort meaning they can get it within a day. In order to really achieve that vision, we know that decision has to be primarily made on the business, and that involves leveraging all the data that we have access to...we want to take the opportunity to make decisions using business data, which is historically not in the case for small business lending.” (002, 2015)

This inability amongst lenders to integrate business and behavioral data also had implications for some groups of customers on the platform. For more established customers, the offers provided by lenders on the platform - using a more limited set of data - were less attractive than pursuing credit with their existing banks, as these banks could provide more competitive rates to customers with which they had worked for years. As one Fincorp leader noted “our best customers are still having the hardest time on the platform” (031, 2015), restricting this segment’s interest in transacting with lenders on the platform.

Data Challenge 4 was further exacerbated by a distinction in the types of credit products that larger, more established, customers needed. While the platform model had been focused on small business lending (providing lump sum payments to small businesses), many larger customers needed drawdown or “line of credit” (002, 2015) facilities. Typically offered by
banks with longstanding relationships to a business, these products allowed ad hoc withdrawals against a pre-approved line of credit to help manage cash flow issues. From the interactions with customers and an overview of the business-level data, platform unit managers discovered that this was an unmet and growing need:

“When one business does business with the other, what they typically do is ship, send an invoice, and get paid in 15 or 30 days. A lot of bigger companies abuse that... We were able to look at the data we have on customers and see that over time small businesses are getting more and more squeezed. Their invoice volume is growing and the terms outstanding are longer. So there’s even more of a need for them to have working capital.” (031, 2015)

The next (fourth) phase was thus characterized by an effort to resolve this data challenge to fuel platform growth.

**Phase 4. Leveraging under-utilized platform data to boost platform activity**

*Becoming a complementor to leverage under-utilized data.* The platform unit’s dissatisfaction with the price and type of credit products that lenders offered to Fincorp’s larger customers presented the platform unit with an opportunity to leverage the business-level and behavioral data they had acquired through phases 1 and 2. Combined with the outcome data acquired in phase 3, the platform unit came to a position of making accurate predictions about risk and pricing for drawdown credit products, which would allow them to underwrite and offer these forms of credit themselves.

“There was a sense that we need to put our money where our mouth is and lend. And so we took this project to the CEO and CFO, and got approval to pursue this. We initially were looking to invest about $10 million dollars of our own money...[so] you have almost two businesses now, we have the platform, and then you have this integrated product” (002, 2015)

To facilitate this early step into providing their own products, Fincorp contracted a third party to run the administration of the product, such as “service and collection” of loans (031, 2015), while underwriting the product themselves. This provided Fincorp with another pathway for direct monetization, and a mechanism to attract and retain customers on the platform.
“it's a lot better in that regard; it's faster and easier. So, that is branded Fincorp line-of-credit. We started the launch. It's an invitation-only line of credit, so we're going to our more established businesses and offering it out...a lot of customers need to have term loans, a lot of customers need invoice receivable, a lot of customers aren't established enough for the line of credit...the more we can learn in the line of credit, the more we think we can grow the product, overall.” (031, 2015)

Fincorp built these line-of-credit products and priced them using data that was not being used by other lenders: “It does require deeper data. And in our application, in addition to the [Fincorp data we have] we're actually asking the customer, ‘Can you share information about your banking?’...’What has your business account looked like over the last three months?’ And there is deeper integration within that regard.” (031, 2015). The accuracy of this model increased with large supplies of recent data: “If they continue to put a lot of data in the cloud, we can help them even more and say, look, your business has grown even more. Congrats. We’ve upped your line to $75K or something like that. You could build a stickiness in that regard, similar to what Amazon is doing.” (031, 2015)

**Complementors’ response to platform sponsor’s new product.** A central concern for the managers at the launch of the platform had been that launching a credit product would deter complementors. As one senior Fincorp executive noted in 2013: “one of the keys is we never become a lender because the minute we become a lender we compete on our own platform, we put ourselves at odds with our lenders” (008, 2013). However, Fincorp executives found that by using data that lenders could not use to offer a product that lenders did not offer on the platform, they could limit resistance from complementors:

“It's a gap that's not offered, and in the marketplace, we've had very good success with the less-established businesses. In some of the younger businesses, the very high end of the credit spectrum, there wasn't this perfect match of speed and ease with a very low fee. So, in some ways, this is almost a new channel for us, if you will” (031, 2015).

Furthermore, a more diverse range of credit products would attract a larger pool of customers, enhancing network effects and providing promising leads for complementors.
“A lot of lenders on the platform side...are equally excited and don’t view it as a threat, but view it as a huge opportunity to work together. They get that there was a problem for many of our customers that come to the platform. We have kind of separated it into two, and...our partners are perfectly fine with it and understand the need that we are going after.” (002, 2015)

Over time, the network effects created by the new credit products also attracted new complementors. Throughout 2015, a “growing ecosystem” (018, 2015) of third-party developers began using the business data on customers to build products on the platform.

“(It is) delivering an awesome user experience, it is really around this data because both—like I said, because it makes it so much easier and it’s completely personalized... Every application becomes custom to you. That’s an awesome experience, and you can only do it with data.” (045, 2015)

Fincorp managers likened the platform to an “operating system” (018, 2015) for small businesses that would attract innovative third-party developers producing various other financial applications for customers across services.

“It helps us build the ecosystem. That’s the big thing with platforms, is that end users may come to your ecosystem for one of your apps, then stay because of the network of applications you have around it... And when you get more users, there’s more opportunities to make money. This is what Google, Facebook, do because you can then present them with other opportunities. So that’s the ultimate, or one of the benefits, from a business perspective.” (045, 2015)

Fincorp’s credit platform saw sustained growth over the following years. By 2022, it had grown to become a larger portion of the organization’s revenue (Annual Report, 2022). The ecosystem of third-party developers had expanded to include organizations providing invoicing, forecasting, and inventory services.

DISCUSSION

a. Synthesis: A model of data and actor interests in platform development

Before delving into the contributions our study makes to extant knowledge on platforms and firm strategy, we find it useful to take a bird’s eye view on our findings, which outline the dynamic interplay of actor interests and data strategy, and highlight their role in shaping
platform development. Our case shows that platform development in established firms is shaped by the interests of a diverse range of actors, including the platform sponsor (the organizational unit that leads platform development), other organizational units dedicated to existing product businesses, distinct coalitions of complementors, and both new and existing customers. These interests shape, and are shaped by, the data strategy of these distinct groups of actors, resulting in a dynamic process model.

The four phases of our case are defined by shifts in platform capabilities, underpinned by changes in data strategy of the actors involved. These changes allowed the platform sponsor to address distinct data challenges that emerged through the nascent development of the platform. These data challenges - defined by issues in the aggregation, sharing, and utilization of data to develop the platform - are visualized in Figure 1.

---

**INSERT FIGURE 1 ABOUT HERE**

---

Phase 1 was centered around the platform sponsor attracting large and reputable complementors to its platform by promising access to its large customer base and developing matchmaking capabilities using its existing customer data. In this process, two data challenges arose. The first concerned how to access shareable data from existing customers (Data Challenge 1). In our case, the power imbalance between the large platform sponsor firm and the smaller, fragmented customers made it easier for the sponsor to force the implementation of an opt-out, rather than opt-in, system for data sharing. This fast-tracked the process of gaining consent from existing customers to share data.

The second, related challenge was more internally organizationally based and concerned how to consolidate existing customer data that was siloed in each of the distinct business units within the firm (Data Challenge 2). The platform unit within the large firm understood overcoming this internal siloing of customer data to be a strategic challenge as it required the
unit to negotiate access with larger, more powerful units. In our case, engaging the firm’s top management to encourage other units to share data alleviated the problem considerably. We observed that this was possible for two reasons. First, the platform unit, which was commissioned by top management in the first place, benefited from top management sharing its ‘platform vision.’ In addition, the platform unit strategically used early evidence of the platform’s matching capabilities (enabled initially through manual data dumps) to entice top management to take an active role in facilitating data sharing among business units. Overcoming the strategic challenge in getting access to siloed data, however, revealed another, this time technical, aspect of Data Challenge 2: data from various business units were not interoperable. The platform unit worked around this challenge by using periodic, manual data dumps from other business units into the platform. Until there was enough data to justify efforts to solve interoperability issues and fuel live transactions, the platform sponsor operated as a high-level matchmaker, providing cold leads to both sides of the platform without further data.

As data became increasingly available on the platform, the platform sponsor focused on enabling and encouraging transactions between complementors and end customers on-platform. However, this process was hampered by complementors’ preference to receive from, but not give back data to the platform. This was a difficult challenge to overcome in our case as the power dynamics worked against the platform sponsor. We observed that the well-known strategy of approaching large, established complementors to jumpstart network effects (Eisenmann, Parker, & Van Alstyne, 2011) had a negative consequence for data sharing with complementors. In our case, large reputable lenders supported the legitimacy of the nascent lending platform, but their rigid processes left little room for the integration of new data sources. In addition, there was also a strategic component in that complementors were hesitant to share valuable transaction data with potential competitors on the platform. Lack of complementor data presented a new data challenge for the platform sponsor: without on-platform transactions,
the platform sponsor did not receive the transaction or outcome data needed to refine the platform’s matching capabilities (Data Challenge 3). We observed that the platform sponsor addressed this challenge again through a workaround; this time by procuring external proxy data to fill in the gaps in the database and ‘close the loop’ in the predictive model.

As input and output data became increasingly available on the platform, and Data Challenges 1-3 were addressed, the platform sponsor examined more closely whether all of the available data was being used by complementors. Using the external outcome data, the platform sponsor discovered that even when existing complementors used the platform, they underutilized the available data. This was, again, largely due to the platform sponsor's earlier strategic decision to invite larger, more established complementors, who came with rigid processes, to the platform.

Existing complementors’ reluctance to utilize the available data created a gap in the range of products offered on the platform, subsequently limiting engagement from certain customer groups whose needs were unaddressed by the offerings (Data Challenge 4). Understanding that overcoming this challenge through a platform business model with existing complementors would be impossible due to the existing power dynamics, the platform sponsor took a two-pronged approach to further grow the platform. First, they launched their own products to fill gaps in offerings, and second, they gradually incorporated new (less established and more flexible) complementors that were better able to leverage the novel data types available on the platform. This strategy of onboarding smaller, but more agile complementors later on, once rich data was available, helped grow the platform by fueling network effects through more diverse and innovative products.

Overall, our account of how Fincorp launched and successfully grew a platform alongside its existing product businesses highlights the importance of addressing various external and internal data challenges that arise in the process. We find that this can be done first
by skilfully managing existing and new stakeholders to open up channels for data flow and second, by using workarounds to fill data gaps whenever needed to avoid hampering network effects. The impact of these activities on each of the data challenges encountered by the incumbent firm is visualized in Figure 2.

----------
INSERT FIGURE 2 ABOUT HERE
----------

b. Contributions to literature

Our findings provide several contributions to the literature on platforms. Previous studies have called for more research to understand various dimensions of the process of implementing a platform business model in incumbent firms (Parker, Van Alstyne, & Choudary, 2016). They have suggested examining how existing relationships outside the firm (e.g., with existing customers or suppliers) need to be reevaluated as the firm shifts to a new role as an enabler of transactions rather than a classic buyer/supplier (Hagiu & Wright, 2015). In addition, scholars have hinted at how incumbent firms with multiple businesses may see existing internal relationships and power dynamics with other departments come into play as critical factors affecting the success of platform initiatives (Nambisan et al., 2017).

Our study answers these calls through an in-depth case study of how one organizational unit inside an incumbent firm created and integrated a platform business model into the firm’s existing product business. In this process, our findings highlight an aspect of platform development previously unexplored, *the interrelated dynamics of data and power*.

We find that although an incumbent firm may enter the process of platform development with existing data and relationships they can leverage to attract complementors or customers, incumbent platforms cannot have all data on the platform at the beginning as existing relationships need to be renegotiated. In incumbent platforms, at the beginning, including
existing user/complementor data on the platform will require negotiations internally and externally. In addition, compatibility issues with internal and external stakeholders may make the platform creation and integration more challenging. We find that challenges in compatibility also provide an excuse for various stakeholders to strategically withhold data.

Our findings also highlight the various strategies to navigate the data and power dynamics with diverse stakeholders during platform development. First, we find that incumbent firms must facilitate both a sustained equilibrium of data from buyers and sellers (end users and complementors), but also access to usable data between business units. While prior research has identified agile organizational units serving as platform sponsors within incumbent firms (e.g., Gawer & Cusumano, 2014), we posit that these units need to adopt multifaceted strategies to elicit the necessary cooperation from the larger organization. In firms focused on product development, achieving this collaboration entails engaging with stakeholders whose objectives may not be congruent with those of the platform unit (and, in fact, might be at odds with them). Overcoming this requires the unit to get buy-in and assistance from top management in the wider firm by demonstrating initial engagement from complementors and customers, and the potential for data-driven matching in the market. Using existing data to show the power of the predictive capability of the platform will help the platform unit get buy-in from the wider firm’s leadership to help change data sharing norms and practices inside the firm through a top down approach.

Regarding external relationships, our study provides a nuanced view of how firms can navigate these over time during platform development. First and in line with theories of power dependence (Pfeffer et al, 1976), incumbents can exploit relative power over (fragmented, small) customers to change terms and get user data onto the platform. Consolidated, larger complementors, however, are a different story. Large, well-known partners may be necessary at the beginning to boost the reputation and trustworthiness of a new initiative (Baum et al,
2000; Stuart, 2000; Ozcan and Eisenhardt, 2009). But these partners are a double-edged sword as they may be harmful for fueling network effects through data. These large firms are likely to be more competitive and less open to data sharing both for strategic and competency related reasons. We find that in these cases, incumbents need to compensate for data scarcity with externally acquired data to fuel network effects and jumpstart the platform.

Finally, our findings show that a platform sponsor can strategically change its complementors over time. As more data is generated, the platform can add innovative complementors (typically smaller, more fragmented) that can harvest generated data more effectively compared to large, reputable but less cooperative complementors.

These findings provide necessary insights into platform evolution, particularly in the less studied settings of incumbent-sponsored platforms. They tell a story of data and power in platform development, highlighting the various challenges and strategies for incumbent firms in this process. They contribute to previous studies that have argued that different actors and different relationships are needed over time to grow a platform (Hagiu & Wright, 2015; Parker, Van Alstyne, & Choudary, 2016; Rietveld, et al. 2019).

Asymmetric Value of Data in Platform Development

Our findings also tell a broader story of data management in platform development. We show that when adopting a platform business in an incumbent firm, an important consideration is the way existing and new data is handled. Choices regarding data are particularly significant as the incumbent firm may have relevant (e.g., customer) data but may face various challenges in gathering it from across the wider firm and/or convincing existing customers to allow access to data (Tilson et al., 2010; Barrett, et al., 2015).

We show that data has asymmetric value across stakeholders and across phases of platform development. While previous research has shown that data access drives platform
engagement, we demonstrate that data is not homogenous, but rather varies along several dimensions. This means that data shapes actor interests, but not uniformly. In our case study, we have observed that some types of data are highly valuable to both the actors that own or generate the data, but also to other actors (e.g., lenders’ outcome data). This data ends up being closely defended by the owner/generator to protect their own interests in the platform ecosystem. Other forms of data may be valuable to other actors but provide less of a competitive advantage to the owner/generator (e.g., SME behavioral data). This data is more freely supplied, but may not immediately be utilized by other actors on the platform. Data can also be distinguished by the degree to which it can be assimilated and leveraged by other actors. Not all actors have the same degree of absorptive capacity (Cohen and Levinthal, 1990), limiting their ability to accommodate and utilize novel types of data. This can be driven by organizational attitudes to risk (e.g., with novel behavioral data), but also by technical challenges of integrating data with different architectures and source codes.

Our findings on data heterogeneity also have important implications for competitive strategy. Previous studies have emphasized the trade-offs facing a sponsor competing with complementors on its platform (Zhu and Liu, 2018; Wen and Zhu, 2019; Rietveld and Schilling, 2021). Extant theory anticipates that there are very few pathways in which a platform sponsor firm could act as a complementor on its own platform without cannibalizing the value extracted from the platform. The distinctions we uncover in the value and usability of the data on a platform suggest, however, that incumbent firms can provide products and services on a platform by leveraging data that cannot be used by existing complementors due to their priorities and technical competencies. In other words, data heterogeneity allows platform sponsors to capture value from a platform in different ways simultaneously.

Overall, by showing multiple types of data and their asymmetric value to different stakeholders over time in the platform development process, we hope our study will advance
literature regarding a particularly important but under-researched phenomenon today: incumbents creating and integrating platforms into existing product or service businesses.

**CONCLUSION**

This paper aims to address critical gaps in literature on platform development through an investigation of how large, multi-business unit firms can start a platform alongside their existing business, and how they manage internal and external relationships and data in this process. Our findings identify four distinct phases in the development of the platform, bracketed by changes in the way that data is aggregated, shared, and used by the platform sponsor, complementors, and end customers. This research contributes to a better understanding of the role of data in platform development and sheds light on the strategies incumbent firms can adopt in the platform economy.

**REFERENCES**


TABLES AND FIGURES

Table 1 - Data Sources

<table>
<thead>
<tr>
<th>Source</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interviews</strong></td>
<td>82 interviews with Fincorp Senior management between 2012 and 2016</td>
</tr>
<tr>
<td></td>
<td>• Interviewees selected with input into firm platform and data strategy.</td>
</tr>
<tr>
<td></td>
<td>• Responsibility for product management, infrastructure development, service delivery, pricing, operations, and partner strategy.</td>
</tr>
<tr>
<td></td>
<td>7 interviews were conducted with managers in complementor organizations.</td>
</tr>
<tr>
<td></td>
<td>• Senior management in large, legacy organizations familiar with Fincorp</td>
</tr>
<tr>
<td></td>
<td>• Founders and senior management in agile fintech organizations that engaged with the platform in latter stages of research</td>
</tr>
<tr>
<td></td>
<td>Interviews used a flexible interview protocol to reveal topics and patterns most salient to respondents. Protocol drew on both real-time and retrospective topics, capturing linkages and patterns across time.</td>
</tr>
<tr>
<td><strong>Archival documents</strong></td>
<td>76 archival documents from Fincorp, which included:</td>
</tr>
<tr>
<td></td>
<td>• 20 Investor reports</td>
</tr>
<tr>
<td></td>
<td>• 7 Transcripts of speeches made by senior management</td>
</tr>
<tr>
<td></td>
<td>• 43 Regulatory filings</td>
</tr>
<tr>
<td></td>
<td>• 6 AGM transcripts</td>
</tr>
<tr>
<td></td>
<td>Archival document review process drawing on written commentary to build an emergent understanding of the issues shaping the development and integration of an adjacent platform.</td>
</tr>
<tr>
<td><strong>Secondary documents</strong></td>
<td>26 third-party commentaries from the period</td>
</tr>
<tr>
<td></td>
<td>• 7 Analyst Reviews</td>
</tr>
<tr>
<td></td>
<td>• 19 Industry Reports</td>
</tr>
<tr>
<td></td>
<td>Used to triangulate accounts from primary sources (interviews and archival documents) and develop an understanding of the contextual ambiguity in the industry.</td>
</tr>
</tbody>
</table>
| Dimension                              | Themes                                      | Illustrative quotations                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |}
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregating data across organizational units</td>
<td>Organizational tensions</td>
<td>• “that’s a big, heavy lift that my team’s been working on… It’s one thing to say, ‘Well, we’ve got four million small business customers,’ but when all of that data is locked on someone’s hard drive on their desk somewhere, who cares, right?’” (013, 2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “we need to make underwriting decisions faster and more efficiently…we need to convey that” (002, 2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “infrequent and irregular” (002, 2013) intervals, resulting in a patchy and incomplete picture of the financial health of the SME</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “fundamentally an organizational issue…not a technical one” (022, 2013).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “…[our colleagues are] out fighting for their lives…And we’re now saying, ‘We’re going to open this up to help our customers’ …that’s a very uncomfortable conversation for a lot of people because they have two advantages--a data advantage and channel advantage, that no one else has. Now, we’re going to potentially lose those advantages… I can’t prove what the upside is.” (013, 2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “we cannot be a successful platform if, as a developer, we have nine different groups of data and I’ve got to go to each” (029, 2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “we need to…centralize things so that we don’t have to duplicate things three, four, five times and there will actually be efficiencies… that actually is the fundamental challenge in executing” (029, 2013).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “we just ended up having the same functionality being developed by multiple teams” (027, 2014).</td>
</tr>
<tr>
<td>Intra-organizational power</td>
<td></td>
<td>• “literally, two hours ago—[the CEO] and I were in a meeting with our CTO and we said we need to focus on that this spring with our engineers. We need to keep coming back to the network effect platform…” (003, 2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “Our CTO mandated that we need to have standardized services that are exposed to both internal users as well as third parties…[for example,] if I go to developer.facebook.com I can get everything I need to build a Facebook app, and we wanted to make it that easy that any third-party developer should be able to go in, read the documentation, try it out, and then start culling those APIs.” (027, 2014)</td>
</tr>
<tr>
<td>Technical hurdles</td>
<td></td>
<td>• “they are distinct at a foundational level…the codebase is different” (002, 2014).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “The flawed notion we kind of carried for a long time was this idea of the Desktop and Cloud products being equal, allowing us to say to complementors ‘the data is actually interoperable between them and we’ll allow you to use it.’ It did not work. So, Desktop is now off in kind of a separate box right now…it’s not a seamless integration.” (030, 2014)</td>
</tr>
<tr>
<td>Data-dumps as an intermediate solution</td>
<td></td>
<td>• Compromise: “data dump” every four to five months (027, 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “We should be able to read their company ID…and then in our backend we had a dump of one million customers and their data like their company information, their annual sales, and stuff like that. So what we would do is if somebody comes in from the Desktop product… we would link them up in our backend and then do some prefilling on the website so that things like what was your revenue, how many employees do you have, and things like that can be prefilled.” (027, 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “we want to get all this data…match it with the lender criteria without actually asking them any questions” (027, 2014).</td>
</tr>
<tr>
<td>Longer-term development</td>
<td></td>
<td>• “So we showed them…we actually have been running experiments to prove that it’s predictive, and we’ve been doing matchmaking between our banks and our small businesses; and we’ve, already, through just a small pilot, done three million dollars’ worth of loans that are very positive, and that’s a way to use data that starts to create opportunities.” (003, 2014)</td>
</tr>
<tr>
<td>Dimension</td>
<td>Themes</td>
<td>Illustrative quotations</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Aggregating sharable data from existing</td>
<td>Challenges gaining permission</td>
<td>• &quot;They’re coming to us with a problem, we’re saying, ‘We think we can help you, but first we need to share some of this (data) with lenders and we need to get your consent for that’ and a minority say ‘Okay.’ … there is a bridge there. We cannot assume that that bridge is easy for everyone to cross.” (008, 2013)</td>
</tr>
<tr>
<td></td>
<td>Workarounds</td>
<td>• “And so today, you answer a few questions and then we connect you to a lender that we think makes the most sense… then from that point on you go and kind of formally apply with the lender. The vision is the data that we have on the customer... all of that we can use behind the scenes to match customers” (002, 2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “It’s kind of a cold hand-off because we don’t pass data to the lender and so the customer has to start all over with the application there. The only thing is that the lender knows that this is somewhat of a qualified lead that they’re getting. Where we’re trying to head to is that being a much warmer transfer and us potentially owning more of that end-to-end experience.” (002, 2013)</td>
</tr>
<tr>
<td></td>
<td>Customer shift to cloud</td>
<td>• “So we’ve actually sort of cut the cord, right…we will stop support of the Desktop sync manager” (007, 2015).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “We’ve crossed a bunch of commercial milestones. We’re prepopulating for 95% of the population.” (031, 2014)</td>
</tr>
<tr>
<td></td>
<td>Power to launch opt-our agreements</td>
<td>• “critical mass” (028, 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “there is some lock-in…if I’m already storing the documents you need in case you get audited, if I’m already creating payments for you, if I’m already sending your 1099’s out, it’s very hard for you to change” (013, 2014)</td>
</tr>
<tr>
<td></td>
<td>Complementor hesitancy to provide transaction data</td>
<td>• “legacy” infrastructure and risk assessment tools (035, 2015)</td>
</tr>
<tr>
<td></td>
<td>Reluctance to transact</td>
<td>• “They’re still manually underwriting it and it’s hard to change that process..to say ‘I’m going to use technology and more boldly decision these small businesses’ is hard for them…it’s a little bit slower.” (031, 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Results in “‘matchmaker…rather than platform” (002, 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “work was really geared at ‘can we build a predictive score based on the customer data that we have?’ And that was the hypothesis, and early results said we could…(however) we were lacking was a lot of data from lenders.” (002, 2014).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “In the first couple of iterations of this we asked for quite a lot of integration with them. We now ask for no integration. In the first couple of iterations of this we asked for quite a lot of data sharing. We now ask for no data sharing. [We] realized that the thing we need to do differently is make the bar to entry just ridiculously low” (008, 2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “Without that data…we have no burden of proof to persuade them to use us” (008, 2013).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Sharing data that “could help competitors…wasn’t an arrangement that they wanted to be part of” (031, 2015)</td>
</tr>
<tr>
<td>Proxy data</td>
<td></td>
<td>• “closing the loop” (006, 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “We turned to get a download of outcome data from decisions that other lenders had made over the past few years. We have used that data from [Credit Rating Agency] to build the predictive model.” (002, 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• ability to “prove” (006, 2014) to lenders that the model was accurate</td>
</tr>
<tr>
<td>Dimension</td>
<td>Themes</td>
<td>Illustrative quotations</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Leveraging proprietary data | Offering new capabilities                       | • “I know which version you bought. I know how many times you logged in. I know how many vendors you’ve paid over the last six months…There is so much other data that I know about you that’s not in that … file.” (101, 2014)  
• “How many times you have logged in for the last three years is predictive of your risk on default…you’d be surprised. So, logging in a reasonable amount is a good thing relative to people that don’t log in very much. But logging in too much makes you more risky…there’s a correlation there with having an extremely limited cash flow…Believe it or not, whether you started with a trial version or not is actually predictive of risk.” (008, 2013) |
|                         | Resistance                                      | • “(there is) data such as revenue, EBITDA…the specific data that they need. There’s a lot more data in then platform … (Lenders) just want to focus on the data that goes into the underwriting decision.” (031, 2015)  
• “we want to drive them to use all the business data. (We) truly feel like it’s more predictive” (031, 2015).  
• “our best customers are still having the hardest time on the platform” (031, 2015) |
|                         | Identifying unmet needs                         | • “There is an unmet need, for a low-price product with low effort… In order to really achieve that vision, we know that decision has to be primarily made on the business, and that involves leveraging all the data that we have access to…we want to take the opportunity to make decisions using business data, which is historically not in the case” (002, 2015).  
• “We looked at the data we have on customers and see that over time small businesses are getting squeezed. Their invoice volume is growing and the terms outstanding are longer… there’s a need for them to have working capital.” (031, 2015) |
|                         | Becoming a complementor to boost platform activity | • “There was a sense that we need to put our money where our mouth is and lend. And so we took this project to the CEO and CFO, and got approval to pursue this. We initially were looking to invest about $10 million dollars of our own money…[so] you have almost two businesses now, we have the platform, and then you have this integrated product” (002, 2015).  
• “I’d love for a party to step up and provide low-cost financing to high-credit quality individuals, and do same-day funding. We’d love for it to happen, and the customers need it.” (031, 2015)  
• “we’re looking at information about banking…there is deeper integration…we’re including behavioral data too” (031, 2015).  
• “If they continue to put a lot of data in the cloud, we can help them even more and say, look, your business has grown even more. Congrats. We’ve upped your line to $75K or something like that.” (031, 2015) |
|                         | Complementor response                           | • “we never become a lender because the minute we become a lender…we put ourselves at odds with our lenders” (008, 2013)  
• “There are two parts to this. One, these customers that we’re targeting aren’t going to lend. We see them coming through the marketplace…The other thing is, for the lenders, this does a good job of providing them qualified leads.” (031, 2015)  
• “A lot of lenders on the platform side…are equally excited…They get that there was a problem for many of our customers that come to the platform…our partners are perfectly fine with what we are going after.” (002, 2015)  
• “It helps us build the ecosystem. That’s the big thing with platforms, is that end users may come to your ecosystem for one of your apps, then stay because of the network of applications you have around it…And when you get more users, there’s more opportunities to make money.” (045, 2015) |
Table 3 - Summary of Phases

<table>
<thead>
<tr>
<th>Phase</th>
<th>Description</th>
<th>Details</th>
</tr>
</thead>
</table>
| 1     | Initial steps towards a platform model           | • Recognizing the platform opportunity  
|       |                                                 | • Data Challenge 1 emerges: aggregating shareable data from existing customers  
|       |                                                 | • Data Challenge 2 emerges: aggregating data across organizational units  |
| 2     | Resolving data access issues to enable early platform transactions | • Streamlining data access to address Data Challenge 1  
|       |                                                 | • Interoperability issues with internal data perpetuates Data Challenge 1  
|       |                                                 | • Demonstrating utility of data to engage TMT and drive firm-wide efforts to address Data Challenge 2  
|       |                                                 | • Data Challenge 3 emerges: Complementor hesitancy to provide transaction data  |
| 3     | Closing the loop to create predictively useful matching | • Building monolithic customer datasets to address Data Challenge 1  
|       |                                                 | • Acquiring third-party proxy data to fill gaps and address Data Challenge 3  
|       |                                                 | • Data Challenge 4 emerges: Leveraging proprietary data  |
| 4     | Leveraging under-utilized platform data to boost platform activity | • Becoming a complementor to leverage under-utilized data, addressing Data Challenge 4  
|       |                                                 | • Complementors’ response to platform sponsor’s new product  |
Figure 1 - Data Challenges (DCs)

DC 2: Aggregating data across organizational units

DC 1: Accessing sharable data from customers

DC 3: Transfer of outcome data from complementors

Key:
- External Actors
- Data
- Internal Actors
- Data Flow
Figure 2 - Approaches to Prevailing Data Challenges

**DC 2: Aggregating data across organizational units**

- Internal Business Units
- Existing Customer Data
- Cross-Unit Data Repository
- Showcase Platform Capabilities w/ Early Data
- CTO
- TMT
- Platform Operator
- Additional Predictive Data

**DC 1: Accessing sharable data from customers**

- Customers
- Changed Data Sharing Terms w/ Customers
- New Customer Data
- Transaction Data

**DC 4: Leveraging proprietary data**

- External Actors
- Data
- Data-related Decisions

**DC 3: Transfer of outcome data from complementors**

- Operator’s own Platform Products
- Legacy Complementors
- New, Smaller, Innovative Complementors

**Key:**

- Data Flow
- Influence

6