LOSING SIGHT OF THE FOREST FOR THE TREES?
PRODUCTIVE CAPABILITIES AND GAINS FROM TRADE AS DRIVERS OF VERTICAL SCOPE

Michael G. Jacobides
Assistant Professor of Strategic and International Management
London Business School
Sussex Place, Regent’s Park, London NW1 4SA, United Kingdom
Tel +44 20 7706 6725; Fax +44 20 7724 7875
mjacobides@london.edu

Lorin M. Hitt
Alberto Vitale Term Associate Professor of Operations and Information Management
The Wharton School, University of Pennsylvania
571 Jon Huntsman Hall, Philadelphia PA 19104
lhitt@wharton.upenn.edu

Working Paper WP 2004/04
Leverhulme Trust’s Digital Transformation Research Programme
@ London Business School
June, 2004

We would like to thank Sid Winter for valuable insights that helped build and then sharpen the theory; Doug Duncan for his support and guidance in the mortgage banking industry; Freek Vermeulen, Julian Birkinshaw, Sumantra Ghoshal, Costas Markides and Margarethe Wiersema for useful comments; and Srikanth Kannan for research assistance. We would also like to thank Will Mitchell and two anonymous SMJ referees for their useful comments. This work was partially funded by National Science Foundation Grant IIS-9733877, the Centre for the Network Economy and the Leverhulme Trust Digital Divide Programme at the London Business School, and was made possible through generous funding by the Mortgage Bankers Association of America. Part of the work was done while the first author was a NATO Science Fellow.
LOSING SIGHT OF THE FOREST FOR THE TREES?
PRODUCTIVE CAPABILITIES AND GAINS FROM TRADE AS DRIVERS OF VERTICAL SCOPE

Abstract
Focusing on proving or disproving Transaction Cost Economics has led to a relative neglect of some key drivers of vertical scope, such as differences in productive capabilities, rather than in capabilities of governance. We consider how such productive capability differences shape vertical scope through gains from trade. Using highly detailed data on the mortgage banking industry, we find them to be a key determinant of the make-vs.-buy decision. We conclude that the distribution of productive capabilities along the value chain, catalyzed by transaction costs, ultimately drives scope.

Keywords: Productive Capabilities; Gains from Trade; Transaction Costs; Vertical Integration
What determines vertical scope? In the last three decades, transaction cost economics (TCE) has been the dominant paradigm for understanding why firms choose to “make” rather than “buy” required production inputs. The basic TCE argument is that the hazards of the market lead firms to increase vertical scope. Absence of such hazards and the presumed benefits of specialization in production lead firms to use “the market.” Numerous empirical studies on vertical scope have been produced over the last fifteen years, and in most of them transaction costs (TC) have been shown to be a statistically significant driver of the make-or-buy choice (see Shelanski and Klein, 1995; Williamson, 1999). However, the attention to validating or refuting the propositions of TCE and related propositions from organizational economics may have taken the focus away from understanding the different factors that can come into play in explaining vertical scope.

Of late, a number of management scholars have voiced their concerns with the predominance of TCE-based views of vertical scope, and there is a nascent “capability-based view” of integration, which favors the factors that relate to firms’ relative advantages, as well as other non-TCE factors (Argyres, 1996; Poppo and Zenger, 1998; Schilling and Steensma, 2002; Leiblein and Miller, 2003; Hoetker, 2004). We consider this debate to be important, and believe that our understanding of firm boundaries would be advanced by empirically examining not only the validity of these perspectives, but also their relative explanatory power.

Our paper contributes to this research stream by comparing existing views on vertical scope from economics and TCE. It also articulates and tests a “comparative advantage” perspective, which focuses on the role of particular firms’ productive capabilities in explaining scope. We define productive capabilities as the efficiency with which firms turn inputs into outputs in particular parts of the production process, similar to what Winter (2003) refers to as “zero-level” capabilities, as contrasted with dynamic capabilities. Our argument is that when a firm uses “the market,” it really uses “another firm”, and hence relies on another firms’ productive capability: firms tend to use the market to buy when their own productive capabilities are inferior to those of other firms that would be willing to transact with it.
Vertical specialization can thus be likened to the analysis of international trade: TC imposes a net “tax” to transactions that happen through the market, when compared to safer, intra-firm transfers. The magnitude of this net tax depends on how the internal governance costs (driven by bureaucracy) compare to the market procurement costs (driven by coordination difficulties, information asymmetries, or the threat of opportunism and hold-up, which is, in turn, determined by asset-specificity and uncertainty). In other words, this “net TC tax” is the difference between external and internal governance costs. Yet, much like in international trade, this net tax discourages but does not prohibit specialization. If differences in productivity between different firms along the value chain are high enough, even a high tax will not curb the tendency to specialize vertically. Conversely, if productive capabilities do not vary much, and in particular if these productive capabilities are symmetrically distributed along the value chain, vertical specialization will not occur, even if the net TC “tax” is extremely low. Therefore, by relaxing the “ceteris paribus” assumption inherent in TCE, which for analytical convenience ignores the differences in productive capabilities, we provide some fresh theoretical and empirical insights.

Our perspective is analogous to the resource-based view of diversification (Wernerfelt and Montgomery, 1988): If the skills and knowledge base that make a firm good upstream also make it good downstream, we expect there will be limited grounds for vertical specialization both at the firm and the industry level. If the skills and knowledge base are different, then firms are likely to be good in either one segment or another, and this will motivate the reliance on the market, and thus vertical specialization. Our approach differs somewhat from recent applications of the resource based view (RBV) which has emphasized capabilities of governance (Winter, 1988; Argyres and Liebeskind, 1999; Leiblein and Miller, 2003), rather than productive capabilities (Demsetz, 1988, Langlois and Foss, 1999; Hoetker, 2004). In addition, as our attention is focused on comparative rather than competitive advantage (as in Hoetker, 2004); issues of sustainability and rarity, central to most RBV analyses (e.g. Combs and Ketchen, 1999; Schilling and Steensma,
2002), do not play as critical a role. This paper provides a theoretical articulation and an empirical test of this productive capability / comparative advantage-based perspective on vertical scope.

Our empirical analysis is made possible by a unique panel dataset provided by the Mortgage Bankers Association of America, used for quasi-regulatory purposes, which has extensive firm-level information on transaction risks, on efficiency measures clearly laid out for each part of the value chain, and a set of continuous measures of the degree of integration at the level of a firm (integration being defined as the percentage of loans made in-house as opposed to being “bought” as qualified loan leads through the intermediate market). The clear separation between the adjoining stages of the production process and the unusually detailed data at our disposal enable us to side-step many of problems of dis-entangling productive capabilities and capabilities of governance. On the basis of this data, we find that while both the TCE and the “productive capabilities” views of vertical scope are supported, the measures of productive capabilities are not only more robust, but also explain an order of magnitude more variance in scope than transactional factors do. This suggests that productive capabilities and the resulting gains from trade are a central driver of scope decisions in this industry, and may be important in other industries with substantial productivity differences across firms in different segments of the value chain.

The contribution of this paper, then, is two-fold. First, on the theoretical level, we articulate the “comparative advantage” view of vertical scope, based on the differences in productive capability distributions along the value chain, and explain how this links with current literature. We focus more particularly on the static drivers of vertical scope, and suggest that the appropriate comparison is not between the “costs of internal governance” as opposed to the “transaction cost of using the market”, as TCE would suggest. Rather, it is between the “net TC tax” of using the market, as compared with the gains from trade that result from vertical specialization; such gains depend, in turn, on how the focal firms’ productive capabilities compare to the productive capabilities of its potential transaction partners. Second, on the empirical level, we use direct measures of productive capabilities (operational efficiency) along the value chain, which enable us
to directly test the extent to which differences between firms’ productive capabilities along the value chain drive scope. Our data also allow us to focus on more direct measures of transaction costs, and in particular on the *ex ante* costs of using the market, such as information mis-representation and information asymmetry --what Williamson calls the “measurement branch” of TCE. More important, our research design allows us to consider the *relative explanatory power* of each of the theoretical lenses. Our findings underscore the importance of productive capabilities, and point to new venues for future research.

**THEORY AND HYPOTHESES**

Several strands of research have provided insights on what drives the make vs. buy decision. Rather than providing a comprehensive guide to the drivers of vertical scope, this section summarizes some key motivations to vertically integrate, focusing on transactional explanations and on the role of resources and capabilities.

**Institutional and Information Economics: Measurement and Transaction Costs**

Perhaps the most studied set of determinants of vertical scope comes from institutional economics, and in particular from the study of transaction costs and risks of using the market. The general prediction is that the greater the hazards relating to the market exchange, the less firms will use it, since firms are considered a relatively safe haven by comparison. Such transaction costs of using the market may be due to either to the *ex ante* problems of information misrepresentation for any given exchange, or to the risks of asset-specific investment, the value of which may be expropriated *ex post* by an opportunistic party in the future.

Inasmuch as it is difficult to measure and assess intermediate goods, it is also difficult to set up market exchange (Barzel, 1982). In the absence of a standardized “grammar and syntax” (Argyres, 1999), market transactions can be costly or impossible altogether (Baldwin and Clark, 2003; Jacobides, 2004). More important, information asymmetry can pose risks when using the market, due to fear of strategic misrepresentation. Akerlof (1970) provides the classic illustration of this
problem, in his analysis of the market for “lemons” (used cars). When the quality of the car is known to the seller, but uncertain to the buyer, it may be impossible to transact, except perhaps at a deep discount to actual value. The risk that the buyer will receive a “lemon” due to seller misrepresentation discourages market trade, especially where the quality of the underlying good or service (or, in our analysis, loan) cannot be verified. Integrated firms, in contrast, have advantages of lower measurement costs or a greater ability to utilize subjective incentives that do not rely on precise measurement (Masten et al., 1991; Jacobides and Croson, 2001). Poppo and Zenger (1998) recently provided evidence on such measurement costs as they affect both the choice of vertical scope and the resulting satisfaction with the governance choice. This leads us to predict that

**Hypothesis 1:** The greater the information asymmetry and the ex ante costs of assessing the quality of what is traded, the greater the extent of vertical integration.

While uncertainty about product quality gives rise to the lemons problem, there are additional transaction risks created by uncertainties in the business environment. In particular, firms may be reluctant to make long-term relationship-specific investments when there is the possibility that these investments can be expropriated in the future (Williamson 1975, 1985; Klein, Crawford and Alchian, 1978). The central problem is that effective trade and efficient production often require investment in assets unique to a relationship by the buyer, the supplier, or both parties. These assets often have little value if the relationship is terminated. This creates the possibility that their value can be expropriated in future negotiations by the threat to withhold trade. Since contracts are incomplete in that they cannot foresee future contingencies, it is not possible to create contracts to mitigate such a risk (Hart, 1995). Absent the complete contracts, firms will under-invest in relationship-specific assets in uncertain environments, reducing the potential value of market interaction and encourage vertical integration where decisions to invest can be made unilaterally. This line of argument, based on the threat of *ex post* opportunism, in the context of a long-lived buyer-supplier relationship, has been extensively researched and documented (Shelanski and
Klein, 1995; Williamson, 1999; Boerner and Macher 2003), with empirical studies generally supporting the notion that asset specificity and uncertainty correlate with integration.

**The Proposed “Comparative Advantage” View**

An alternative line of reasoning, which has emerged more recently, has to do with firms’ particular resources and capabilities. This sub-section summarizes the relevant background and related literature, before developing our own comparative advantage perspective on integration.

*Existing Research linking Capabilities and Vertical Scope.* Over the last few years, it has become clear that TCE is not a self-sufficient theory of vertical scope. Williamson, for instance, recommends that the traditional TCE query “‘What is the best generic mode (market, hybrid, firm) to organize X?’ be replaced by the question ‘How should firm A -- which has pre-existing strengths and weaknesses (core competences and disabilities) -- organize X?’” (1999: 1003). This question has been recently pursued by Madhok (2002), who suggested that an individual firm’s choice must depend not only on the characteristics of the transactional conditions, but also on its strategic objectives, the attributes of its own capabilities, and the governance context it has created. On this broader view of firm scope, there has been significantly less empirical evidence. That being said, Walker and Weber (1984) found that the most important predictors of sourcing were cost differences between the focal firm and outside suppliers in producing a specific component.

Despite this strong evidence, most of the subsequent research in firm boundaries did not include firm-specific cost or capability measures, or use capability differences as controls. Argyres (1996) was one of the first to provide qualitative evidence on the role of firm capabilities in integration decisions, observing that in the cable manufacturing business capabilities were a significant driver of vertical scope, in addition to transaction costs. The growth of the capability- and resource-based view of the firm further led scholars to consider the role of firm heterogeneity: Poppo and Zenger (1998) considered skill sets and scale as determinants of the decision to outsource IT services; Leiblein and Miller (2003) found a strong correlation between insourcing the production of a
specific component and past experience producing the component. Hoetker (2004), using a
different research design, also found capability differences to be critical determinants of supplier
choice. These studies support the existence of path-dependencies in boundary choices, and
specifically the role of firm capabilities (Argyres and Liebeskind, 1999; Madhok, 2002). We
consider this research as pointing to the right direction, and try to extend it, by providing a clearly
articulated framework which links productive capabilities (i.e. the efficiency of turning inputs into
outputs) and vertical scope.

The “Comparative Advantage” Theory of Vertical Scope. While it has recently become clear
that heterogeneity in firms is related to the choice of governance, the discussion in the literature
still tends to neglect that “the market” is, in essence, “a firm.” The “market” does not, in itself,
produce anything. “The market” is an organizational interface, behind which is another firm -- a
firm which finds it advantageous to sell on the basis of its own productive capabilities. In the TCE
calculus, at least as traditionally construed and as empirically tested, the question of what lies
behind “the market” was not a relevant one, since, per the tenets of economic theory, firms should
have access to the same technology (Nelson, 1991). However, if we consider the differences in the
productive abilities of different firms, the relevant choice is between the firm, and its cost levels (if
the input is produced) as opposed to the cost levels of another firm, including its profit margin,
after deducting transaction costs.

To make headway in a theory of capability-based analysis of vertical scope, we also must
further refine the concept of capability. An important distinction is between productive capabilities
of a firm in each of the vertically related stages, and the capabilities of governance, the ability of a
particular firm to use a governance interface (integration or the market) to create value by linking
these two stages. To be precise about this distinction, we introduce some notation. Consider a firm
that performs an upstream activity (e.g., production) and is considering whether to insource or
outsource a downstream activity (e.g., sales). Let the efficiency of the downstream activity for the
focal firm (i) be represented by \( p_i \), and the bureaucratic cost of integrating both activities
internally, including the cost of muted incentives, as \( B_j \). Similarly, let the productivity of a potential outside vendor \((j)\) be represented by \( p_j \) and the transactions cost for firm \( i \) procuring product from that vendor \( j \) through the market be represented by \( TC_{ij} \). In this framework, the vertical integration decision depends on the comparison of whether: 
\[
p_i - B_i > p_j - TC_{ij}.
\]

Historically, the TCE literature has excluded productivity differences of the focal firm and the outside firm (vendor), that is, \( p_j - p_i \), from its direct purview, and has considered internal governance costs \( B_j \) as a constant, non-negative term (Williamson, 1985: Ch 6). Instead, it has focused on the drivers of variation in transaction costs. The more recent TCE literature has, in addition, recognized the impact of firm-specific factors by considering additional variation across firms in \( TC_{ij} \). Such research has argued that transaction costs vary either across suppliers \((j)\) or across buyers \((i)\). For instance, Dyer (1996) argues that using particular suppliers consistently across time mitigates the risks of procurement from that set of suppliers \((j)\). Winter (1988) argues that firms develop particular capabilities of transacting, so that some become better than others in using the market, thus pointing to lower TC for some set of firms \((i)\). Similarly, Argyres and Liebeskind (1999) argue that history of a focal firm \((i)\) or the other governance choices a firm makes in other lines of business affects the costs of using the market. Leiblein and Miller (2003) recently observed that using market procurement by a firm \( i \) in the past is associated with using the market in the future, presumably because of the achievement of low TC. Such studies, though, have typically not focused on productivity differences. To the extent firms’ specific capabilities are considered, they have been focused on the capabilities of governance, instead of variations in productive capabilities (Demsetz, 1988; Langlois and Foss, 1999; Hoetker, 2004).

Productive capabilities, then, can be defined as the efficiency with which a set of inputs can be turned by a firm into a set of outputs in a specific part of the value chain. For instance, productive capabilities could be the efficiency with which a marketing division can turn marketing resources
into market penetration or, in a mortgage banking context, the relative efficiency of turning IT infrastructure, branches, and human capital into qualified loan applications. These are what Winter (2003) recently termed “zero-order capabilities”; that is, the operational efficiency with which a firm conducts its business in any one part of the value chain.

A focus on productive capabilities as drivers of vertical scope, though, only becomes meaningful if there is reason to believe that there is variation in productive capabilities across firms. We believe that this is the case, and the limited evidence on intra-industry productive capabilities or cost structures seems to corroborate this view (Lieberman and Dhawan, 2000). We further argue that there are good theoretical grounds to expect that this should be so: Capabilities, and in particular productive capabilities, rest on the firm’s general and specific knowledge of how to do things (Richardson, 1972; Teece, Pisano and Shuen, 1997) as well as the specific investments and complementary assets (equipment, training, retention of key personnel, etc.) required to put that knowledge to work (Barney, 1994, Winter, 1995). In this perspective, heterogeneous capabilities can arise as a result of a path-dependent learning process, in which there is abundant opportunity for various contingencies to shape the way of doing things that ultimately emerges (Winter, 1990, Levinthal, 1997). This is true even if process knowledge diffuses across firms, since process complexity and interaction among activities limit imitation (Porter 1996, Rivkin 2001, Siggelkow 2001) and some types of investments in capabilities are at least partially irreversible (Ghemawat, 1991; Winter, 1995). In addition, in a volatile environment, differences in capabilities can be further accentuated by differences in the ability of organizations to learn and adapt (Teece, Pisano and Schuen, 1997). Collectively, these arguments suggest that even in an environment where primary resources are homogeneous, different organizations are likely to display significant variations in processes leading to differences in productive capability.

The more pertinent question for our context is how the productive capabilities of a firm are distributed along the value chain. Firms have different capabilities in each part of the value chain; and these relative capabilities influence their decision as for whether they will be integrated or
specialized. In an industry with two vertical segments (upstream and downstream) a firm which has strong capabilities upstream and downstream will be integrated. If a firm is good upstream and not as good downstream, then it will have a good reason to explore the use of the intermediate market as a net seller of intermediate good; conversely, if a firm is good downstream and not as good upstream, it will have the incentive to be a net buyer in the intermediate market which links the up- and down-stream segments. Specialization (use of the market) will happen to the extent that some firms are good upstream and some downstream. For the market to be economically sensible, it must be the case that there are grounds for co-specialization; some must be better (or only active) upstream, and some must be better (or only active) downstream.

This observation leads to two specific predictions about vertical scope. First, at the firm level, a firm will tend to be integrated when it is equally capable (on a relative basis) upstream and downstream. Thus, the extent to which capabilities are positively correlated within a firm will govern the degree of integration independent of transactions cost issues.

Second, this implies a further prediction at the industry level. To the extent that there is no variation in capability across firms in each value chain segment, there is no opportunity for gains from trade (irrespective of transaction costs), and therefore firms will likely remain integrated. The existence of gains from trade at the firm level requires a limited (or negative) correlation between upstream and downstream capabilities at the industry level. If the upstream and downstream segments’ capabilities are very strongly correlated at the level of the industry, then no firm will be good in one segment and not good in another -- which is, as we saw, the necessary precondition for the market to work. So dissimilarity along the value chain (lack of correlation between the up- and downstream segments) leads to capability differences, which in turn lead to the desire of individual firms to reap gains by engaging in vertical specialization.

This second point has a natural parallel to the comparative advantage theory of international trade (Ricardo, 1817). Despite the “taxes” that the market imposes through transaction costs, inter-
firm transactions will still occur in the pursuit of gains from trade – much like international trade, which is driven by differences in productive capabilities of different nations, and which may or may not be curbed through international taxation.

These two observations about the distribution of capabilities within and across firms lead to the following two predictions:

**Hypothesis 2:** Superior productive capability in a particular vertical segment will be positively associated with integration in that same vertical segment.

**Hypothesis 3:** Superior productive capability in a particular vertical segment will be negatively associated with integration in its downstream segment.

Hypothesis 2 argues that the better a firm is in a segment—say, downstream, the less it will tend to use the market as opposed to its own downstream operations, as doing so would be relying in inferior capabilities for that segment (since \( p^D_i > p^D_j \)). An upstream firm is more likely to be integrated into the downstream segment, if it has a superior downstream productive capability.

Hypothesis 3 argues that superior capabilities in one segment will be associated with a smaller degree of integration into the other vertically linked segment. Simply put, the better an upstream firm is in the upstream segment, the smaller the probability that it is vertically integrated into the downstream segment, all else being equal. That is, the better a firm is upstream (i.e., \( p^U_i > p^U_j \)) the greater its use of the intermediate market linking the up- and down-stream segment, ceteris paribus: Firms capable upstream will want to use the intermediate market as a net seller of upstream goods, and as such are more likely to be less vertically integrated. This means that a comparative advantage in one segment leads, other things being equal, to greater tendency to be vertically specialized. The objective is to leverage the productive capability \( p^U_i > p^U_j \). Hypothesis 3 further predicts that firms may want to use both their own integrated operations and the market concurrently, as they try to leverage their productive capabilities. This, we argue, is the reason why in may settings firms (e.g. in financial services or manufacturing) use both owned and other firms’
distribution channels; their superior upstream capabilities make them want to use the market, in addition to their own distribution networks, to leverage their strengths.

Our theoretical perspective also provides a consistent explanation for why firms may be active in both vertical segments, as opposed to simply specializing in the segment where they have an advantage. Consider a firm with a better-than-average upstream division, and a worse-than-average downstream division. The fact that it has a downstream division at all may be due to historical reasons, perhaps reflecting different competitive conditions, or even past mistakes. Yet it may have an economic incentive not to redress them, by abandoning the relatively inefficient downstream division: The irreversibility of prior commitment downstream (Ghemawat, 1991; Staw, 1981) or the fact that the costs for operating downstream are sunk makes it economical, on the margin, to maintain the original downstream operation as a going concern, despite its average inefficiency. The marginal benefit of shutting it down is negative, as some of the relevant costs have already been sunk (Sutton, 1991; Ghemawat, 1991). We also expect that such a firm will grow its upstream operations, and shrink its downstream segment. Thus, the existence of sunk investments and limits to the speed of entry or exit in a segment underpins Hypotheses H2 and H3, and also explains mixed governance and the gradual changes in firms' scope often observed in practice.²

Our perspective also focuses on comparative, and not absolute advantage within each part of the value chain. For specialization to occur, the benefits (i.e. gains from trade) from not relying one one’s own productive capabilities must simply exceed the costs of using the market (the “net tax” of TC). For instance, if firm A has a comparative (but not absolute) advantage upstream, and a disadvantage downstream, and firm B has the inverse, i.e. a comparative (not absolute) advantage downstream and a comparative disadvantage upstream, vertical specialization may still happen. When neither of the two firms has an absolute advantage (none of these firms is best-in-class) either up-stream or downstream, specialization will happen inasmuch as the gains from trade from these comparative and asymmetrical advantages is large enough to compensate for the costs
of engaging in a market transaction. Indeed, we predict that all firms for which gains from specialization surpass transaction costs are likely to be specialized. This observation has two corollaries. First, vertical specialization requires gains from trade; and such gains from trade can only happen if productive capabilities are not symmetrically distributed along an industry’s value chain. Thus, industry-level distribution of productive capabilities is a critical predictor of vertical scope. Second, in order to explain vertical scope, we can look at comparative advantage, rather than the more exacting criterion of competitive, absolute or best-in-class advantage.

The Comparative Advantage vs. the Resource-Based View. Finally, for the sake of theoretical completeness, we should note that our approach is subtly different from what the Resource-Based View (RBV) implies. Admittedly, the RBV does not lead to a simple, across-the-board prediction about firm integration (Barney, 1999). That being said, the RBV suggests that integration is broadly related to the stock of resources under control (Combs and Ketchen, 1999, Leiblein and Miller, 2003; Leiblein, Reuer and Dalsace, 2003), especially if these are rare and non-tradeable. Thus, scope is set with an eye to leverage resources so as to yield competitive advantage (White, 2000; Afuah, 2001; Schilling and Steensma, 2002). Integration, for instance enables a firm to access complementary assets so as to reach the final customer to exploit an upstream advantage, especially when it is difficult to contract for downstream (sales) capabilities (Teece, 1986, 1996; Williamson, 1999) or when the asset is short-lived and rapid market access is essential (Schumpeter, 1911). We do not explicitly consider these resource-specific (as opposed to capability-based) and dynamic motivations to integrate to protect or exploit a firm’s particular resources; instead, we argue the “simpler” view that even a potentially fungible and short-lived competitive advantage in terms of the productive capability of a firm in an industry may explain a significant part of the variance in the way firms choose their vertical scope.

Finally, the theory we develop in this paper focuses on a fairly narrow, static picture. We do not explicitly consider the cost of developing capabilities or at whether the capabilities are best formed organically or through mergers and acquisitions; all we argue is that at any point in time, firms
decide their scope on the basis of their relative, *comparative* and not necessarily absolute advantage. We return to the limitations of our argument, and to the necessary extensions of this research approach in the discussion.

To recap, different theories have shed light to what drives vertical scope. Table 1 summarizes the major approaches, identifying their focus and predictions as for when we expect to see integration, as a function of the attributes of the transaction (information asymmetry and impactedness; threat for opportunism in a re-negotiation under asset-specificity), or of the transactors (governance capabilities in carrying out transactions; productive capabilities).

**METHODS AND DATA**

**Research Design and the Choice of our Setting**

*Research Design.* Our general analysis approach is to consider the relationship between integration and the factors driving integration as predicted by the measurement branch of transaction cost and information economics (H1) and our “comparative advantage” perspective (H2, H3) that relies on gains from trade, as a function of differences in productive capabilities. In addition to examining the significance of constructs provided by these theories, we also consider these theories’ explanatory power in terms of their relative ability to explain the variance in integration across firms.

In order to explore this issue analytically, we would need a setting whereby we can distinguish between the productive capabilities, and the capabilities of governance. That means that we need a setting that has two distinct parts of the value chain, where we can at least discern the productive capabilities in each stage, as well as assess the nature of the TC “tax” between these two segments. We would also want variance in that “tax”, so that we test the relative explanatory power of transactional arguments.
Also, to make this exercise analytically tractable, we would also need a setting whereby we can measure the firms’ productive capabilities in different parts of the value chain. In the event that the product or service in question (in each of the two adjoining parts of the value chain) is commoditized, so that there is not much of a difference in the output side, we can measure productive capabilities as the efficiency of using the inputs, or alternatively as a measure of operational efficiency – as an operating measure of this “zero-level” capability (cf. Winter, 2003). We could alternatively use a setting where there is heterogeneity in terms of quality or differentiation of the output, but this would greatly complicate the analysis as it would require doing a quality-adjustment in terms of the outputs. For all these reasons, we chose mortgage banking- a setting which conveniently offers itself for the analysis set forth in this paper.

**Our Setting: Mortgage Banking.** Mortgage banks are non-depository financial institutions that originate, process, approve, and then (in most cases) sell mortgage loans to the secondary market. The mortgage banking industry is a good setting to study vertical scope, as it has a complex value chain which has increasingly been fragmented into quasi-independent parts, which are performed both by integrated firms as well as narrowly targeted specialists. Mortgage banks originate loans, which they also service, but typically do not hold the loans as assets; rather, they sell them to the secondary market through large securitizers (quasi-public such as Fannie Mae and Freddie Mac or private such as Citibank Mortgage). The structure of the mortgage banking industry is provided in Figure 1A, which shows the vertically co-specialized world of mortgage banking (which itself competes with more integrated Savings & Loan institutions of banks). Mortgage banks generated more than 56% of the total mortgage loan production in 2002, about $800 billion in new loans. Mortgage banks, therefore, are a very important sector, despite the dearth of related research.

Mortgage banking is one of the least regulated segments of the financial services sector, perhaps because the mortgage origination process carries little systemic risk for the financial services sector as a whole. As a result, mortgage banks are free to make their own choices of scope
unaffected by regulatory pressures, and thus our dependent variable, the degree of integration, is not contaminated. However, while there is no formal regulatory framework for the industry, mortgage banks are called upon by their Association and the securitizers to disclose key measures of financial and operational health through a survey (the MBFRF). This database offers unusually detailed measures in different parts of relative efficiency in each part of the value chain.

Our particular focus is on the most upstream part of the mortgage loan production process, which is illustrated in Figure 1B. Specifically, we look at one of the most important decisions of vertical scope for mortgage banks: the question of whether they are integrated in the production of loans, or whether they use correspondent banks or loan brokers from whom they procure loans. Mortgage banks can outsource the generation of the lead, the gathering of the relevant paperwork and the filling of the relevant forms, as well as the hand-holding of the loan applicant to a mortgage broker or to another bank (a “correspondent bank”). The broker or correspondent would then “sell” the qualified loan to the bank (receiving a commission ranging from 80 to 150 basis points depending on how close the loan is to funding, and how valuable it is). This choice, then, is a clear make-vs.-buy: The bank can make the loan itself or buy it from another entity.

Figure 1B also provides an illustration of what exactly we mean by the term “productive capabilities” in our setting. We focus on two vertically linked segments: Loan origination, and loan warehousing. Productive capability in loan origination is the efficiency with which a mortgage bank or a mortgage broker undertakes the steps described in the steps described in Figure 1B: Pre-qualification, document generation, application processing, credit analysis and underwriting, and approval. Productive capability in warehousing consists of the ability in engaging in pipeline / warehousing analysis and the transfer to the secondary market (ability to manage interest rate risk, prepare the loan bundles for sale to secondary market investors, etc). Governance capabilities consist of linking one’s (own or outsourced) loan origination process to the loan warehousing process. Even if these two types of capabilities do intertwine over time, they can be assessed separately at any one point in time. So in our setting, we can analytically
distinguish productive capabilities, and also measure them as “efficiencies”, since the outputs (mortgage loans) are measurable.

The mortgage banking setting is also convenient in terms of the demographics. Both the upstream and the downstream part of the value chain are quite competitive, with a substantial number of participants, and this tends to alleviate the small-numbers bargaining considerations. While consolidation has been growing, the concentration in the industry is quite low, especially in the loan origination segment. More than 850 medium and large firms were active, in addition to a host of smaller loan brokerage firms, specializing only in origination. Furthermore, origination does not appear to require economies of scale, but we did control for this by including several scale measures, as we note below.

**Data: Mortgage Banks’ Integration into Retail Loan Production**

_Retail Integration as a Dependent Variable._ We take the perspective of all banks which warehouse a loan (which, as Figure 1B suggests, consists of the activities necessary to take that loan to the secondary market and sell it), and consider whether these banks make their own loans, whether they buy them, or do both. Another feature of our setting, then, is that banks can either be fully integrated (never using outside agents); fully specialized (a type of bank that became known as a wholesaler); or use a mixed mode, that is both buy and make loans. Most banks in our sample are indeed users of the mixed mode: They buy some loans, and use the market for some other loans. This multiple sourcing structure is probably at least as common (if not more so) than the binary make-buy decision analyzed in the TCE literature (cf., Harrigan, 1985). A useful feature of our setting is that we can measure integration as a continuous variable, representing the aggregation of a large number of integration decisions at the loan level. This avoids the potential classification problems and statistical inefficiency that arise when a naturally continuous decision variable is forced into a discrete choice (e.g., probit) modeling framework.
Data: The Mortgage Banking Financial Reporting Form. This study was made possible through a large-scale collaborative project with the Mortgage Bankers Association of America (MBAA), that enabled us access to the detailed confidential and quasi-regulatory database that the MBAA maintains on the activities of the mortgage banking sector, called Mortgage Bankers Financial Reporting Form (MBFRF). Each year, a sample of mortgage banks is selected, and efforts are made to ensure responses are received by a wide variety of respondent banks, with the explicit aim being to maximize variation in the respondent profile while covering a reasonable share of the total market. The sample is an unbalanced panel of all firms which have warehousing activities, and the yearly observations range from 79 to 285 per year; each firm is sampled for an average of 3.1 years, out of a possible maximum of 10 years, and the total number of firms in the sample is 685. Re-sampling is done in a way that all the types of firms are equally represented: The MBAA executives who finalize the sample informed us that they take pains to ensure that all different types of firms are represented in the sample; that is, the sample always corresponds to a representative cross-section of the industry. Thus, we expect that there is little, if any chance of the results being driven by sample selection. On average, the respondents of the MBFRF are responsible for about 25% of the total loans produced in the United States, and as such this is a highly representative sample. The database we utilize covers the period 1988-1998, and contains 1,792 usable observations.

Due to its quasi-regulatory and confidential nature, these data contain a much richer and finer level of detail than is typically found in public regulatory filings databases (e.g., FDIC call reports, Compustat). In particular, these data cover the costs, revenue, output quantity, and employment in each part of the production process enabling direct measures of productive capabilities by value chain segment, degree of integration, and firm scale. In addition, characteristics of the loans produced or managed are also captured, which enable the measurement of transaction costs.
Measures: Integration, Transaction Costs, Comparative Advantage Metrics and Controls

**Dependent Variable: Integration in Retail.** Our dependent variable is integration in retail loan production; more specifically, the percentage of total loans produced through a bank’s own retail branches (as opposed to those purchased from brokers or correspondents) in terms of dollar value. We have also used the percentage of loan numbers (as opposed to dollar value) made in-house as a dependent variable, and the qualitative results did not change.

**Transaction Costs.** Previous theoretical and empirical research in TCE has recognized the empirical difficulties of measuring transaction costs directly. Williamson (1975), recognizing this problem, suggested focusing instead on conditions that drive transactions costs such as asset specificity or uncertainty. While this is useful if the objective of research is to validate these components of the TCE framework, it is necessarily limiting in that other types of TC, including frictional costs (Coase, 1991) and information / measurement costs (Akerlof, 1970; Barzel, 1982) are not related to these factors. Our paper focuses on these *ex ante* TC, given that we do have good direct measures.

From our industry interviews we determined that the principal risks of using the market for the procurement of loans surround the risks of trading loans of unknown quality. Borrowers can vary in ways that are difficult to observe or objectively measure that influence their probability of delinquency or default – a very real cost to the mortgage banks. Moreover, there is a strong incentive for firms producing the loans to “disguise” the true riskiness of the loan by the way data are reported or to sell loans which they believe are more likely to default but have the same observable characteristics. This is compounded by lag times between origination and default, limiting the ability of the firm to learn about the risk of their trading partners. The ability to obfuscate risk and the amount of unobserved risk is likely correlated with the inherent risk level of each type of loan. Government-sponsored loans (FHA/VA) are known to be the highest-risk type of loan due the large heterogeneity of the eligible pool of borrowers, and the fact that the credit
standards are more generous as a matter of social policy. “Plain vanilla” fixed-rate conventional mortgages (FRMs) have intermediate levels of risk -- their long-term and (relatively) high interest rate raises the possibility of default due to unexpected events. Finally, adjustable rate mortgages (ARMs) have the shortest expected term and lower interest rates, making them the least risky, as industry-wide default rates suggest. Our data enables us to measure a firm’s portfolio of these different types of loans, as well as the default and delinquency rates, to validate these conjectures.

To confirm that this categorization was indeed indicative of transaction risks we conducted three sets of tests. First, using robust regression, we examined the relationship between integration and overall default risk. Our results suggest that default risk is indeed correlated (significant at p<.05) with market procurement. From the regression estimates, we calculate that an entirely integrated bank faces a default rate of 1.72%, while the rate for a bank that buys all loans on the market is 2.06%. Second, we compare default and delinquency rates of each loan category for each firm over time using the non-parametric Wilcoxon test. Our results suggest that FHA/VA loans have default rates substantially higher than ARM or FRM loans (p<.001). FRMs also have higher default rates than ARMs but the contrast is less substantial (p<.1). We obtain similar results using an alternative data source for overall defaults and delinquencies by loan type from the Department of Housing and Urban Development (HUD) over our sample period suggesting the patterns in our data are consistent with the market. However, the fact that some loans are riskier is not enough to establish the existence of differential transaction costs.

To use the loan categories as surrogate measures for the extent of transaction risks, we also need to demonstrate that riskier loans are comparatively riskier to procure through the market. This leads to our final test, in which we use robust regression to compute correlations between the use of retail production (integration) and the percentage of delinquent loans by loan type. We find that the more a bank is integrated, the lower the risks of default losses, but that the correlation patterns vary depending on the loan type. The correlation is highly significant for the more dangerous FHA/VA loans – regression estimates suggest that a fully integrated firm faces losses
on these loans of about 1.43% in contrast to a 2.77% loss rate for a specialized firm (significant at p<.001). This difference drops to .16% for FRMs (p<.05) and is not significant for ARMs. So for the “risky” loans, the use of the market leads to significant costs of misrepresentation; whereas for the standard loans, there is little, if any, transactional hazard induced by using the market. This means that loan types do provide a reasonable direct measure of the transactional hazards a firm faces, at least with regard to misrepresentation and measurement (Akerlof, 1970).

Our measure of transaction costs, then, is the composition of the loan portfolio in terms of percentages of these three loan types – firms with a larger proportion of “dangerous” loans should be more integrated in order to mitigate such risks. Measuring TC in this manner has the advantage that it represents a direct measure of transaction risks relevant this setting; the shortcoming is that because “hold-up” is not a major risk in this sector, we cannot make any generalizations about the relative explanatory power for these other TC risks versus our hypothesized productive capability factors.

Measures of Productive Capability. A key advantage of the MBFRF data is that the productivity, as conventionally measured in the industry, can be captured at the level of the individual value chain segment. For our purposes, we focus on the distinction between loan retailing (producing the loan) and loan warehousing and marketing (funding and selling to securitizers). All the mortgage banks in our sample are active in warehousing, and most are also active in retail production. As discussed earlier, we define productive capability as the efficiency with which firms turn their inputs into outputs. Given that output (mortgage loans) are fairly homogeneous and standardized, we can approximate this efficiency by measuring labor productivity and operational margins for each segment. We thus use two types of measures for the productive capability of loan origination and loan warehousing. The first set of measures regard efficiency per employee (a) in loan origination and (b) the warehousing segment. The next set of measures consist of the operating margin in (a) loan origination and (b) warehousing, consistent with standard industry practice (Posner and Nambar, 2002).
Specifically, our first set of measures of productivity is loans per employee for each segment and income per the loan number in each of the two segments, as we are fortunate to have both loan numbers and employees per segment. Although the labor productivity metric omits a measure of capital as is conventional in productivity analysis, it is unlikely that this introduces significant biases as physical capital for mortgage banks (office space, computers) is highly correlated with employment. Thus, our labor productivity measure is likely to be closely related to the more comprehensive “multifactor productivity” metric in this industry. These productivity measures can be computed separately for the retail segment (measuring the number of loans originated by the firm by each FTE employee in origination); and the warehousing segment (measuring the number of loans, whether produced in-house or bought through the market, warehoused by the bank by each FTE employee in warehousing).

The second set of measures we use are the operating margins obtained in each segment though the MBFRF income breakdown. While measures of per-segment margin can be noisy, especially in financial institutions, they do provide useful information that can be used to corroborate the FTE-based efficiency metrics. The existence of both margin measures and labor productivity measures provide us with an unusually direct measure of productive capabilities. Finally, note that given the distributional properties of the margin and of the FTE per loan metrics, we used the natural logarithms of these measures in our regression.

Hypothesis 2 suggests that vertical integration in retail will be positively associated with the margin per loan produced in retail production: the more attractive downstream production is, the more a firm does. Therefore, we expect a negative correlation between the number of FTE per loan produced and integration, as this measure is the inverse of efficiency (the higher the measure, the less efficient the firm). We also expect a positive correlation between the retail production margin and the degree of integration into retail.
Hypothesis 3 suggests that integration into retail production will be negatively correlated with the productive capabilities in warehousing. Therefore, we expect that the margin per loan in the upstream parts of production (warehousing) will correlate negatively with integration: The better a firm is in warehousing, the more it will attempt to use the market (loans purchased from brokers) to leverage and profit from its capability. For the same reason, we expect a positive correlation between the number of employees in warehousing and retail integration (the weaker a firm is in warehousing, the less it will be inclined to use the market to buy loans so it can warehouse them).4

Control Variables: Time. In addition to the variables noted above, we include a year dummy variable. First, each year the pool of respondents changes, and thus we need to correct any such compositional bias. Second, in years with high demand, mortgage banks are known to use more outside agents, since their own branches are not enough to cover the necessary demand (Nishiguchi, 1994). For that reason, integration may be related to the economic cycle, and year dummies should absorb this effect.

Scale. Another important control variable is scale. Scale could potentially be important for two distinct reasons; one relating to technology of production, and the other one relating to market power. A traditional argument is that integration depends on the absolute scale, inasmuch as a minimum scale (i.e., ratio of fixed to variable costs) might be required to make integration viable. The example of pin manufacturing (Smith, 1776) illustrates that the pin-makers in a small village would not have the requisite scale to integrate into metal molding, whereas a large, urban pin factory might have the requisite scale to do so. Also, larger scale helps by smoothing demand from many different sources, which leads to economies of agglomeration and economies of throughput maximization (Chandler, 1962). Similarly, integration may require higher capital expenditures, which, in an imperfect capital market, may only be economical to larger firms (Galbraith, 1967).

Another economic argument considers the role of market power in prompting integration. In imperfectly competitive industries, the pursuit of monopolistic or oligopolistic rents motivate
integration: Firms may want to integrate in order to raise rivals’ costs (Salop and Scheffman, 1983), control scarce resources (Galbraith, 1967; Porter, 1980), eliminate multiple marginalization (Salop, 1979; Dixit, 1983), improve the ability to price discriminate, or to obtain a strategic upstream supply (Stigler, 1951; Arrow, 1975; Riordan and Sappington, 1987).

For these reasons, we control for scale, and we can also get some prima facie evidence on the relative validity of these economics-related arguments by observing if integration is a linear function of scale. More specifically, we capture scale through a measure of assets, to which we added the asset square term to consider the non-linear effects. Other measures of scale such as equity, total number or dollar volume of loans produced, and number of employees were also employed, and yielded similar results.

**Methods**

Our main regressions are on the percentage of loans produced in-house, measured in dollar terms. We report four types of results. First, we report ordinary least squares (OLS) with all firms and years pooled, estimated with Huber-White robust standard errors accounting for repeated observations of the same firm. Failure to correct for this error structure tends to inflate t-values (White, 1980).

The second type of analysis consists of robust regressions. We utilized the Stata 8.1 rreg procedure which operates by calculating an OLS estimate, calculating the Cook’s D influence statistic and performing a generalized least squares using the influence statistics as weights. This procedure reduces the effect of outliers (Cook, 1979; Berk, 1990; Hamilton, 1991), which is particularly important in financial industry data where financial assets and returns can vary more widely than in “real” corporations. The alternative approach of individually investigating the outlier points is not feasible due to our confidentiality agreement.

To control for unobserved heterogeneity, and to ensure the robustness of our findings, we also use a panel data fixed effects model (cf. Baltagi, 1995), which effectively controls for any time-
invariant firm specific factors in our unbalanced panel. However, our confidence in these models is guarded, as during the observation period there has been substantial merger and acquisition activity which can influence the estimates of fixed effects models and number of repeated observations of firms in the sample is not large which tends to make these estimates imprecise.

Finally, we use logistic (Logit) regressions, again with Huber-White standard errors, to explore the impact of TC and capabilities on the choice of firms that chose to be fully integrated, or fully specialized. These Logit regressions are run only on the sub-sample of our segment which is nearly entirely integrated (more than 90% of loans produced in house) as opposed to those that are almost entirely reliant on the market (less than 10% of loans produced in house). Other cut-off points, including 100% integration or specialization, yielded qualitatively similar results, but we focus on the 90% cutoff to maintain a reasonable sample size.

In all four methods, we sequentially introduce explanatory variables for scale, for transaction cost related factors and finally capability related factors. The availability of multiple methods (as well as more specifications we tried in addition to those reported) indicate significant robustness in our results. Finally, as we note in the research design section, we are not only interested in the statistical significance of the coefficients associated with each theory; we are also interested in the extent to which the measures can explain the variation in our dependent variable.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
Include Tables 2, 3 and 4 about here & \\
\hline
\end{tabular}
\end{table}

\textbf{RESULTS}

Model 1 in Table 3 presents the correlation between the degree of integration, control variables, and scale. We can see that our control measures do not drive the results: Rather than the predicted positive linear relationship between scale and integration, there is a “u-shaped” relationship, with a strong and undoubtedly negative linear term, and a positive squared term. Integration is neither linear, nor monotonically positive function of scale. Similar results obtain in the Logit regressions of Table 4, although they are less strongly significant.
Hypothesis 1 (on the role of TC) received mixed support, as we can see through Model 2, that includes controls and the TC related measures. In the logistic regression, which considers what explains the choice of firms to be entirely integrated or entirely specialized, the sign and magnitude of the coefficients are what the theory would predict (see Table 4). Relatively safe loans (FRM) are related negatively to the decision to be fully integrated, and intermediate-risk loans such as ARM are not statistically correlated to integration. Finally, we find that the greater the proportion of risky loans a firm produces (FHA/VA), the greater the probability that the firm would be integrated, although this finding becomes significant only in the full model.

The picture changes once we consider not only the binary choice of integrating versus not integrating, but also the precise level of integration for the large number of firms that rely on both the market and their own retail branches. Model 2 in Table 3 contains the results for the OLS, robust and fixed-effects regression on the continuous measures of integration. In OLS, the composition of the loan portfolio is not statistically significant in explaining integration. However, in our most restrictive specifications, robust and fixed effects regressions yield results that occasionally contradict the transactional logic.5

What is more interesting to us is the amount of variance explained. Even in the Logit regression, where the results go in the expected direction, the increase in fit (pseudo-R²) is 2.6% (from 1.6% to 4.2%). To put this in context, the addition of capability metrics increases the pseudo-R² to 72.6% (see below). The ability of transactional factors to explain integration is even weaker in the continuous measure of integration, where the fit increases from 5.1% to 5.7% in OLS, and the within-R² in fixed effects increases from 4.2% to 5.7%.

Hypotheses 2 and 3 (relating to productive capability) are very strongly supported as we can see by looking at the results of Model 3, which includes the productive capability measures. Table 3 provides the results for the continuous measure of integration. As we can see, the capability-related coefficients in terms of productivity (FTE per loan) are significant in the expected
direction. Thus, the greater the efficiency downstream (the smaller FTE in origination is), the
greater the extent of integration into that segment. Firms that are better than others in origination
are integrated in that vertical segment (H2). Furthermore, the coefficient of the efficiency in the
upstream part of production (loan warehousing) is strongly negatively associated with integration
(H3). The better a firm is in upstream production (the smaller FTE in loan origination), the more it
tries to leverage its advantage by relying on the market (the smaller our dependent variable).

Similar findings obtain for the margin measures we use. The higher the origination margin in a
segment, the more that segment is used (as opposed to the market) and the greater the degree of
integration (H2). Also, the higher the margin in warehousing, the greater the reliance on the
market (H3). The most important issue, however, is to note the dramatic increase in explanatory
power: In OLS, the $R^2$ increased from 5.6% to 50.5%. Given the nature of the data, this is a
remarkable fit.

These findings are strongly corroborated in the robust regressions, where the coefficients are
even stronger on the predicted direction. Indeed, margin in retail production, the only capability
metric which was not above the significance threshold, becomes clearly significant. Also, scale
effects almost vanish in the robust specification, once we include the capability metrics.

The Fixed-effects analysis is consistent with the OLS and robust regression in supporting H2
and H3 --most of the coefficients are highly significant in the same direction. The reduction of
significance was expected in a fixed-effects specification; indeed, if fixed-effects estimates were to
be fully relied upon, we might have seen an even greater reduction of the significance in our
coefficients, since using that method, fixed long-run components of firm capability are removed,
leaving only the transient component. Either way, the fixed component alone cannot explain
integration, since capability differences not only vary between firms, but also over time. This
happens as some firms become better and some relatively worse, consistent with the path-
dependent process of capability development described in our theory section, so that our capability measures are important even after controlling for time invariant firm characteristics.\(^6\)

The Logit analysis is also consistent with OLS, Robust, and Fixed Effects. The coefficients for both H3a and H3b are significant and in the hypothesized direction, except for the warehousing margin measure (H3b). Perhaps more importantly, once we add the capability-based metrics, pseudo-R\(^2\) in the Logit regression increases from 5.6% to 72.6%.

**DISCUSSION**

This analysis suggests that, at least in our particular setting, differences in productive capabilities are key drivers of the decision to integrate. While a transactional logic may be correct (at least in determining the binary institutional choice), capability considerations dominate. This finding, of course, could be contested as the result of the specificities of our setting. Yet such a criticism would be misinterpreting the major premises of our study, which are, first, to indicate that productive capabilities can and often do play a major role in the determination of vertical scope; and second, to suggest that, rather than empirically examining whether particular perspectives are correct, we have to focus on the relative role of different factors in explaining vertical scope, with the acknowledgment that their relative role will depend on the specificities of each context. That, however, does not mean that we should give up on measuring relative impact; rather, we should design multi-setting, multi-industry studies to explore the relative ability of different factors (including both TC and capabilities) to explain variation in scope.

But let us return to our main finding, which is the predominance of productive capabilities and hence gains from trade as motivators of vertical scope. While this finding, and the corresponding theory, had not been explicitly articulated and indicated to date, it is still consistent with all earlier empirical work that acknowledged the role of capability heterogeneity between firms. Walker and Weber (1984), for instance, found that cost benefits were the most important predictors of firms’ make-vs.-buy choice. More recently, Leiblein and Miller (2003) found that the inclusion of firm-
specific variables dramatically improved the fit of their model. We would re-interpret their finding that firms tend to integrate when they were integrated in similar products in the past, as a surrogate of these firms’ ability to be efficient in the processes in which they are integrated. Hoetker (2004) also found the differences in capabilities of suppliers (as measured by past patents) to be a strong predictor of supplier choice. We would generalize this to suggest that inter-firm differences in capabilities are a key driver of vertical specialization.

More broadly, the surprisingly limited empirical evidence (e.g., Lieberman and Dhawan, 2000) indicates that variation in capabilities is persistent and important within an industry, and, in all likelihood, within firms in the value chain (cf. Duncan, 1998). As such, we should expect that many settings would be much like ours, where comparative advantage happens to play a major role in predicting vertical scope. Clearly, the question of variation in productive capabilities merits dedicated analysis.

Our findings show that in examining vertical scope, we have to consider what is “behind” the market over and beyond the capabilities of governance interfaces to coordinate economic activities. In addition to firm-specific capabilities in using the market or being integrated, it also matters how good each firm is, compared with other potential industry participants, in each stage of the value chain. Simple logic suggests, yet much research neglects, that if a firm has a comparative advantage in a particular part of the value chain, another firm has a corresponding weakness. For gains from trade to exist, there must exist a capability imbalance, so that both parties gain from specialization. This can happen because there is an asymmetric distribution of capabilities, with one firm being good upstream and another downstream, and hence specialization occurring.

Yet while we consider that differences in productive capabilities are very significant and should be studied further, we also believe that this analysis should not be done to the exclusion of the analysis of transactional factors. On the empirical level, we find that transactional factors are
important. But, more importantly, on the theoretical level, we consider that transaction costs and capability differences not only work as independent motivators of vertical scope, but, crucially, they interact to shape the Institutional Structure of Production (Coase, 1991). TC determine the extent to which these latent gains from trade, driven by the distribution of productive capabilities in an industry, are or are not sufficient to justify vertical specialization. Much as global taxation can curb international trade, costs of using the market may prevent specialization and reliance on the market. So TC are such that they catalyze the differences in capabilities; their reduction enables the underlying differences in productive capabilities to be brought to the fore, and as a result, allows specialization to occur. However, it is not the reduction of transaction costs itself that matters but the relative reduction in these costs, when compared to the capability differences that lead to the latent gains from trade.

Finally, another connection between productive capabilities and TC is that the very reduction of TC may be motivated by the existence of latent gains from trade. Interested parties, which stand to gain from specialization, may push and invest for TC to be reduced, expecting their investments in reducing TC to be repaid through their share of the gains from trade (Jacobides, 2004). The joint analysis of differences in productive capabilities and transaction costs / vertical scope holds much promise for the future and opens up new venues of research.

Our approach is rooted in the analysis of firm-level heterogeneity, and is hence commensurate with recent applications of the RBV to the analysis of vertical scope. Yet on the other hand, there are some key differences in our approach. Rather than focusing on the objective of sustainable competitive advantage (Combs and Ketchen, 1999; Schilling and Steensma, 2002), we focus on the emergence of comparative advantage. While theorists may be interested in the objective of sustainable advantage, and on the rarity, non-imitability, and non-substitutability of a resource (Barney, 1986), we argue that firms often base their decisions of scope on their comparative advantage.
Furthermore, our results suggest that even modest advantages may drive scope; and to understand these advantages we must shift our analytical focus from the specificities of each firm to the macro-patterns of variance and capability heterogeneity in an industry, along a value chain. This focus on the systematic patterns of productive capability distribution opens up exciting new venues for research. The question becomes, what drives the extent of co-variation of capabilities in an industry? For instance, is the lack of such co-variance a function of knowledge bases (Winter, 1984, 1987) where the divers of upstream and downstream performance are independent? Or are there, in particular settings, real managerial diseconomies of scope, so that being in one part of the value chain adversely affects performance in another? That is, might vertical specialization be related to inappropriate use of common HR and organizational practices or to the establishment of firm-wide “rules of thumb,” or simple rules, common across the firm (Eisenhardt and Sull, 2001)? Is it due to the establishment of “dominant logics” (Prahalad and Bettis, 1986), simplifying devices (Gavetti, Levinthal and Rivkin, 2003) and mental models (Markides, 2000), which help one part of the value chain and hinder another? Asking such questions and understanding the distribution of comparative advantage will lead us to re-integrate the theory of production and the theory of exchange (Demsetz, 1988; Langlois and Foss, 1999), and build a more robust set of explanations of the institutional structure of production (Coase, 1992; Madhok, 2002; Jacobides and Winter, 2004). The central question becomes: When and why do productive capabilities not co-vary? This is what we should consider in order to gain a deeper understanding of what drives vertical scope.

**Limitations**

While this study does indicate that a factor we had not explored so far, namely, productive capabilities, matter, it is subject to a number of limitations. On the empirical level, our study does not directly test all aspects of Williamsonian TC. While we do believe that in our setting the appropriate TC are those related to misrepresentation and risk. So whereas we do think that we have unusually direct measures of these costs, confirmed by our empirical analyses, we still do not directly test for the relative ability of Williamsonian TC to explain variance. Thus, our findings
cannot rule out that further unobserved TC are empirically as important as capability differences in explaining scope, especially in other settings. That being said, the measures we used for capability differences cannot be reasonably construed as indirectly associated with Williamsonian TC, so that our basic finding, that is, that productive capabilities matter, still stands.

On the theoretical level, there are additional limitations. First, our paper presumes the existence of a clear, separable governance interface; that is, presumes that there is a potential market, with some level of TC “tax”, linking two discrete steps in the production process (say, A and B) so that we can identify the productive capabilities associated with step A, the capabilities associated with step B, and the governance capabilities linking steps A and B. While this could be done in our setting, it is often the case that the structure of the value chain is not quite as clearly demarcated, and that we cannot easily distinguish between the different types of productive capabilities and dis-entangle them from governance capabilities. Indeed, a key analytical challenge which we did not tackle is to examine why some segments do not even have the option of being vertically distinct. While we would argue that there is something in the nature of capabilities that is associated with what is often called “technological inseparabilities”, our approach both takes them for granted, and cannot explain them. We consider recent research on this topic (Baldwin and Clark, 2003; Jacobides, 2004) to be an important complement to our approach. In particular, we think that there is much promise in examining how productive capabilities and governance capabilities affect and determine each other over time (Argyres and Zenger, 2004); and how capabilities and transaction costs co-evolve, shaping the structure of the production system (Jacobides and Winter, 2004). Thus, while we do not offer a comprehensive theory of vertical scope, we hope that pointing out an important empirical regularity (that productive capability differences, at any point in time, drive scope), as well as articulating the comparative advantage view will help support future research in scope, capabilities, and their co-evolution.

Second, while our objective was to provide a more inclusive theory of vertical scope, we also did not directly consider any of the “dynamic” drivers of integration. In the theory section, we
noted that a motivation for vertical integration is the protection or access to a stream of revenues from upstream innovation. This might not be particularly relevant in a mature industry like mortgage banking, yet it would be a significant omission in a more dynamic setting (White, 2000; Afuah, 2001; Schilling and Steensma, 2002; Leiblein and Miller, 2003). Likewise, we have not considered another important dynamic element: The role of technological change and how integration relates to it. Afuah (2001), for instance, demonstrated that integration in old technologies is detrimental to overall performance, possibly because it impedes the systemic improvement of capabilities along the value chain.

Another limitation of our study is that we have only confirmed that capability differences drive scope. Yet while we speculated on what drives these capability differences, and hypothesized that there will be settings in which the differences of capabilities along the value chain may be high, and others in which they will be low, we have neither empirically tested that, nor have we studied what makes some settings and not others prone to gains from specialization and trade. Similarly, we have not examined how capability differences and capability distributions evolve over time, affecting both vertical scope and the division of labor between particular firms in the industry.

Finally, our paper takes the vantage point of the entire industry, looking at the logic that drives specialization; it tries to shift from the “ceteris paribus”, partial equilibrium analysis inherent in TCE as well as the RBV, to a systemic analysis of what drives specialization in a given population of firms with a particular distribution of transaction costs and capabilities. However, while such an approach yields some fresh theoretical insights, it does not enable us to speak directly to the problem a manager faces in determining the scope of their firm. In particular, our approach is silent on the cost of developing capabilities; and it does not examine whether either organic growth or M&A can change the capability distribution over time, so that firms engage in M&A or selective nurturing of capabilities so as to avoid vertical specialization and compensate for their capability gaps (cf. Karim and Mitchell, 2000). We consider these topics as important extensions of our research. On the basis of the knowledge that capability differences do affect vertical scope,
we hope that future research will examine why in some settings firms manage to balance their capability gaps, precluding specialization; and whether this happens organically or not. Such an approach, building on our findings, could potentially expand on recent research by Helfat and Raubitschek (2000), Karim and Mitchell (2000) and Mitchell and Shaver (2003), applying these author’s insights to the question of vertical scope as it evolves.

Concluding Remarks

This study proposes that in the determination of vertical scope, we have to consider separately the capabilities of governance of different governance interfaces (markets or integration) as well as the productive capabilities of the firms in an industry. It argues that behind “the market” lies another firm, willing to supply a good or service at a particular price; and that the ability of any such firm to produce in a cost-effective manner will be an important part of the calculus of the make-vs.-buy decision. In a world of significant variation in these productive capabilities between firms, capability differences may be responsible for a significant part of the decisions to integrate or not -- exactly as it happened in our setting. Transaction costs can thus be seen as a tax that seriously impacts, but does not in and of itself generate, the need to specialize.

Our deceptively intuitive finding, that the choice of scope is largely determined by comparative capabilities, opens up some fascinating questions, once we shift the focus from the individual firm to the population of firms and the related distribution of capabilities. The analysis of the co-variance of productive capabilities of different firms along the value chain should itself become the object of study, which can help us understand the deeper causes of vertical scope. By expanding TC analysis and combining it with a more careful study of the managerial structure of production, we will be able to obtain a more robust and more representative theory of vertical scope. Building such a theory, and focusing on the ability to explain variance, as opposed to the correctness of a particular theoretical prediction can thus help us not lose sight of the forest for the trees.
References


Endnotes

1 The logic is as follows. Suppose the buyer knows the value with certainty, while the seller believes the value to be equally likely across a range of possible values. If the buyer proposes a price \( p \), the seller concludes that the value must be less than that price \( p \), which leads them to estimate the average value to be \( \frac{1}{2} p \) (assuming all values below \( p \) are equally likely). Thus, trade is not possible at any price greater than zero. It is the inability of the seller to credibly communicate the value, and the buyer’s recognition that the seller has incentives to misrepresent the value, that inhibits trade.

2 A related rationale for mixed governance rests on “differential scalability” between vertical segments – that is, on the possibility that it is costlier to expand, say, the downstream operations than to expand the upstream operations. Consider a firm with two equally strong vertical segments, which wants to grow and leverage its absolute competitive advantage: As it can grow upstream faster and more economically than it can grow downstream, it will eventually both produce downstream, and use the market, so as to benefit from its superior and more scalable upstream capability.

3 Loan defaults are costly for mortgage banks in several ways. First, if a loan defaults and there were underwriting errors, mortgage banks are often obligated to buy back the loan from secondary investors and bear the credit loss directly. Second, even if the mortgage bank is not responsible for credit risk, managing default creates significant operational costs of collections, foreclosure and subsequent asset management. Thus, the costs of procuring a “lemon” from the market is large.

4 Note that all the firms in our sample have at least some presence in the warehousing segment (as this is the criterion for their inclusion in the database), so that we do not have any pure vertical specialists in retailing. We do, however, have a few vertical specialists in warehousing (118 firm-years, i.e. 7% of the sample). For these firms, we do not include any measures of capability in the retail segment, as they do not have any retail production. This omission, as well as the omission of the pure retail specialists, is likely to introduce a conservative bias as for the role of capabilities: We have good reason to expect (and we know through other industry studies) that vertical specialists have superior capabilities in the segments they specialize in, and were or would be poor in the segments they are not present in. This means that we should expect that the results in our sample systematically understate the role of H3 as it potentially applies to the universe of mortgage banking, including all specialists with imputed capability measures in areas they are not currently active in.

5 The reader might observe that the coefficient signs on the composition portfolio of the loan portfolio do change between different models and between the logit and OLS / robust / fixed effects regression. This, however, is an expected statistical artifact since all loan categories add up to 1: A negative coefficient in, say, FRM and a positive coefficient in FHA/VA is equivalent to having a positive coefficient in both FRM and FHA/VA inasmuch as the beta for the FRM coefficient is smaller than the one in the FVA, as in both of these cases a shift from FHA/VA to FRM will lead to reduction in the level of the dependent variable (i.e., an increased reliance on the market).

6 On the one hand, if capabilities were time-invariant, then the fixed effects (within) regressions of capability measures on integration should teeter on the edge of significance, whereas we see that in our model they do not; they are strongly significant in the expected direction. On the other hand, by comparing the R-squared in our fixed effects / within model (35.4%) with the R-squared of the overall OLS regression with no firm dummies (50.5%) we observe that some of the variance is explained by time-invariant firm-level differences. However, this result may underestimate the inter-temporally constant element of capabilities, for statistical / sample reasons alone: The problem is, in particular, the contaminating effect of M&A activities, which we know to operate in this industry. As a result of M&A, a firm with a set of capabilities \( C_1 \), leading to a choice of integration \( I_1 \) at some initial time, will become a different firm, with set of capabilities \( C_2 \) (and resulting integration \( I_2 \)) after a merger. As identifying M&A in the sample was not feasible due to our data privacy restrictions, we acknowledge the limits of a fixed-effects model to capture consistent inter-temporal and firm-specific factors, and the limited ability of time-invariant fixed-effects to capture more variation may be due to this.
<table>
<thead>
<tr>
<th><strong>Focus of the theory</strong></th>
<th>TCE – “Measurement Branch” (<em>ex ante</em> problems)</th>
<th>TCE- Williamsonian Opportunism (<em>ex post</em> problems)</th>
<th>Capabilities and RBV Approach</th>
<th>Comparative Advantage Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Attributes of the transaction</strong></td>
<td>Information asymmetry and impactedness; inability to articulate what is needed</td>
<td>Potential for hold-up due to dedicated assets leads to TC</td>
<td><em>No focus / Prediction</em></td>
<td>TC is a “tax”, driven by information asymmetry, impactedness or transactional risks</td>
</tr>
<tr>
<td><strong>Attributes of the transactors</strong></td>
<td><em>No focus / Prediction</em></td>
<td><em>No focus / Prediction</em></td>
<td>(a) Focus on path dependent <em>governance capabilities</em> (b) Focus on absolute advantage or existence of unique, rare resource to be exploited</td>
<td>Focus on <em>productive capabilities</em> and their distribution within segment and along the value chain (which determine total gains from trade)</td>
</tr>
<tr>
<td><strong>When do firms integrate?</strong></td>
<td>Information problems (due to misrepresentation or info asymmetry or inability to coordinate through prices) make integration desirable</td>
<td>When the threat of opportunistic renegotiation (due to asset specificity and uncertainty) is greater than the cost of internal governance</td>
<td>(a) The better the governance capabilities, the more the market can be used (b) Firms will integrate where they have absolute advantage or to protect a resource</td>
<td>When the costs of using the market (the TC “tax”) is greater than the potential gains from trade from specialization between any two firms</td>
</tr>
</tbody>
</table>

**TABLE 1**

What Drives Vertical Scope? Comparing the Different Theoretical Perspectives
<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integration in Origination</td>
<td></td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Assets</td>
<td></td>
<td>-0.1736</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asset Square</td>
<td></td>
<td>-0.1010</td>
<td>0.8809</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% FVA loans</td>
<td></td>
<td>0.0596</td>
<td>-0.0748</td>
<td>-0.0349</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% FRM loans</td>
<td></td>
<td>-0.0664</td>
<td>-0.0173</td>
<td>-0.0172</td>
<td>-0.4304</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% ARM loans</td>
<td></td>
<td>0.0126</td>
<td>0.0328</td>
<td>0.0096</td>
<td>-0.2690</td>
<td>-0.1083</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTE in Production per loan</td>
<td></td>
<td>-0.0600</td>
<td>0.0320</td>
<td>0.0270</td>
<td>0.0283</td>
<td>0.0521</td>
<td>-0.1129</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTE in Origination per loan</td>
<td></td>
<td>0.2166</td>
<td>0.0491</td>
<td>0.0420</td>
<td>0.0419</td>
<td>0.0270</td>
<td>-0.1081</td>
<td>0.8903</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>Warehousing Margin</td>
<td></td>
<td>-0.1629</td>
<td>0.3136</td>
<td>0.1484</td>
<td>-0.0418</td>
<td>-0.0274</td>
<td>0.1055</td>
<td>-0.5915</td>
<td>-0.5699</td>
<td>1.0000</td>
</tr>
<tr>
<td>Origination Margin</td>
<td></td>
<td>0.2660</td>
<td>-0.1172</td>
<td>-0.0530</td>
<td>0.0229</td>
<td>-0.1486</td>
<td>-0.0865</td>
<td>0.0952</td>
<td>0.1666</td>
<td>-0.2446</td>
</tr>
</tbody>
</table>
TABLE 3
Integration (continuous measure) in Loan Origination as a Function of Size, Transaction Costs and Capabilities
The Dependent Variable is Retail Integration (percentage of total loans made in-house) across all models and all specifications

<table>
<thead>
<tr>
<th></th>
<th>OLS (Huber-White SE)</th>
<th>Robust Regression</th>
<th>Fixed-Effects Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MODEL 1</td>
<td>MODEL 2</td>
<td>MODEL 3</td>
</tr>
<tr>
<td>Total Assets (control)</td>
<td>-1.25e-10*** (2.39e-11)</td>
<td>-1.2e-10*** (2.38e-11)</td>
<td>-3.93e-11 (2.46e-11)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asset Square (control)</td>
<td>8.38e-21*** (2.21e-21)</td>
<td>8.4e-21*** (2.21e-21)</td>
<td>1.88e-21 (2.61e-21)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>%FHA-VA loans (TC measure)</td>
<td>-0.0029 (0.0774)</td>
<td>-0.0103 (0.0813)</td>
<td>-0.1601*** (0.0362)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>%FRM loans (TC measure)</td>
<td>-0.1059 (0.0671)</td>
<td>0.1073 (0.0779)</td>
<td>-0.1360*** (0.0334)</td>
</tr>
<tr>
<td>%ARM loans (TC measure)</td>
<td>0.0176 (0.0993)</td>
<td>0.1300 (0.0974)</td>
<td>-0.041 (0.0539)</td>
</tr>
<tr>
<td>FTE per loan, origination (productive capability)</td>
<td>-0.1629*** (0.0148)</td>
<td>-0.1671*** (0.0088)</td>
<td>-0.1107*** (0.0108)</td>
</tr>
<tr>
<td>FTE in warehousing (productive capability)</td>
<td>0.2349*** (0.0132)</td>
<td>0.2551*** (0.0078)</td>
<td>0.1453*** (0.0095)</td>
</tr>
<tr>
<td>Warehousing margin (productive capability)</td>
<td>-0.0251*** (0.0048)</td>
<td>-0.0239*** (0.0035)</td>
<td>-0.0030 (0.0062)</td>
</tr>
<tr>
<td>Loan origination margin</td>
<td>(0.0071) (0.0062)</td>
<td>0.0136*** (0.0042)</td>
<td>0.0163** (0.0053)</td>
</tr>
<tr>
<td>CONTROLS</td>
<td>YEAR DUMMIES</td>
<td>YEAR DUMMIES</td>
<td>YEAR DUMMIES</td>
</tr>
<tr>
<td>cons</td>
<td>0.747*** (0.028)</td>
<td>0.8132*** (0.0592)</td>
<td>1.4514*** (0.1686)</td>
</tr>
<tr>
<td></td>
<td>0.7271*** (0.0191)</td>
<td>0.6114*** (0.0392)</td>
<td>0.8534 (0.1209)</td>
</tr>
<tr>
<td>R² (%)</td>
<td>5.13</td>
<td>5.69</td>
<td>50.54</td>
</tr>
<tr>
<td>N</td>
<td>1792</td>
<td>1790</td>
<td>1166</td>
</tr>
</tbody>
</table>

+ p<0.1, * p<0.05, ** p<0.01, *** p<0.001 (Standard Errors in parentheses)
TABLE 4

Logit Analysis of Integration as a Function of Size, TC and Capabilities

Dependent variable: Retail Integration (0 if <10%, 1 if >90%, dropped otherwise)

<table>
<thead>
<tr>
<th></th>
<th>MODEL 1</th>
<th>MODEL 2</th>
<th>MODEL 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Assets</td>
<td>-8.84E-10*</td>
<td>-9.32E-10*</td>
<td>-2.4E-09*</td>
</tr>
<tr>
<td></td>
<td>(4.577E-10)</td>
<td>(4.819E-10)</td>
<td>(1.04E-09)</td>
</tr>
<tr>
<td>Asset Square</td>
<td>1.026E-19</td>
<td>1.114E-19</td>
<td>3.76E-19*</td>
</tr>
<tr>
<td></td>
<td>(1.027E-19)</td>
<td>(1.062E-19)</td>
<td>(1.83E-19)</td>
</tr>
<tr>
<td>%FHA-VA loans (%TC measure)</td>
<td>0.9255</td>
<td>6.2617***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.8086)</td>
<td>(1.4691)</td>
<td></td>
</tr>
<tr>
<td>%FRM loans in (%TC measure)</td>
<td>-1.1311***</td>
<td>5.0129***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.4838)</td>
<td>(1.3193)</td>
<td></td>
</tr>
<tr>
<td>%ARM loans in (%TC measure)</td>
<td>-0.4323</td>
<td>6.1867*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.9408)</td>
<td>(2.9573)</td>
<td></td>
</tr>
<tr>
<td>FTE per loan in origination (%productive capability)</td>
<td>-3.4460***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.5866)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTE in warehousing (%productive capability)</td>
<td>3.4576***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.5677)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Warehousing margin (%productive capability)</td>
<td>0.0556</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1503)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loan origination margin (%productive capability)</td>
<td>0.0893</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1817)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.1040***</td>
<td>2.5813***</td>
<td>-1.7352</td>
</tr>
<tr>
<td></td>
<td>(0.1672)</td>
<td>(0.3693)</td>
<td>(1.9987)</td>
</tr>
<tr>
<td>Pseudo R² (%)</td>
<td>1.61</td>
<td>4.21</td>
<td>72.62</td>
</tr>
<tr>
<td>LR Chi²</td>
<td>13.31</td>
<td>34.68</td>
<td>276.21</td>
</tr>
<tr>
<td>N</td>
<td>1106</td>
<td>1104</td>
<td>643</td>
</tr>
</tbody>
</table>

+ p<0.1, * p<0.05, ** p<0.01, *** p<0.001 (Standard Errors in parentheses)

Logit Analysis performed with robust standard errors, clustered on Firm ID to account for non-independence of errors. Likelihood Ratio Chi² is reported from the non-robust model.
FIGURE 1
The Value Chain in Mortgage Banking

Figure 1A: The Broader Mortgage Banking Value Chain

Figure 1B: The Activities in the Mortgage Bank’s Value Chain

Note: The activity which mortgage banks can “make” themselves or “buy” through the market is Loan (i.e. retail) origination. This activity, with these five value-adding steps, is our focus.