A relationship-based theory of the firm: a Chinese perspective

Xin Li
Asia Research Centre
Copenhagen Business School
xl.int@cbs.dk

Abstract
In this paper, I argue that all the existing theories of the firm, i.e., the transaction cost, knowledge-based, and entrepreneurial theories, are insightful yet partial since they only deal with one or another possible situation in which the first ever firm came to exist. In addition, all of them have a common problem of buyer/entrepreneur/firm-centrism because they all ignore the role of employees in the formation of the firm. I argue, since they are complementary to each other, a new, comprehensive and integrative theory of the firm must be able to unify them with a consideration of employees. I propose a relationship-based theory of the firm (R’BT) as such a candidate theory. The R’BT places the employer-employee (or broader, firm-stakeholder) relationships at the centre of its whole theoretical framework and argues that the notion of relationship harmony is fundamental in explaining the nature of the firm.

Keywords: theory of the firm, relationship, harmony, transaction cost, entrepreneurship, knowledge

1. Introduction

Entering into the new century, the world has witnessed a string of global crises, ranging from the ‘moral crisis’ caused by scandals such as Enron in 2001 and Madoff in 2008, to the environmental crisis such as pollution and climate change, and to the global financial crisis. What can be seen in them are the direct and indirect linkages between those macro-level crises and the micro-level firm behaviors. The simultaneous failure of the business communities to achieve economic, environmental and social sustainability urges us to rethink the ineffectiveness of our management knowledge and to reexamine the basic tenets of our field (Freeman and Zollo, 2009). In this paper, I focus on one foundational issue: the theory of the firm.

In economics, the theory of the firm is the most rapidly expanding subfield (Foss, 1996a: 470). In a narrow sense, i.e., in accordance with Coase (1937), a theory of the firm should address three issues regarding the nature of the firm: the existence, the boundaries, and the internal organization of the firm. According to Conner and Prahalad (1996: 479), the value of developing a theory of the firm for management and organizational researchers is to identify a necessary building-block for addressing the broader issue of performance difference between firms. Thus, Foss (1999) suggests that a strategic theory of the firm should address one extra issue: the competitive advantage of the firm.
Broadly speaking, there are three schools of thought within the existing literature on the theory of the firm: transaction cost economics (TCE), knowledge-based theory (KBT), and entrepreneurial theory. The TCE is based on Coase’s (1937) transaction cost theory and currently dominated by Williamson’s (1975, 1985) opportunism-based theory of asset specificity. The KBT is pioneered by Kogut and Zander (1992), Conner (1991), Conner and Prahalad (1996), Grant (1996) and Pitelis and Teece (2009). The entrepreneurial theory is proposed by Casson (1998, 2005), Foss and Klein (2005) and Langlois (2005). Each of the three schools offers different answers to the three Coasian themes.

Foss (1999) points out that in this research field, there are two approaches toward theory advancement. Some people prefer developing the theory of the firm along each line of thinking while some others see an integrative theory combining different schools necessary and desirable. And in fact the entrepreneurial theory is such a kind of integrative effort to synthesize the theory of the firm and the theory of entrepreneurship long existed separately in economics. In this paper, I want to go a step further to integrate all the three schools of thought by proposing a relationship-based theory of the firm (R’BT).

2. Literature review

The TCE argues that, according to Coase, the firm exists because it can reduce transaction costs incurred when using price mechanism (i.e., market). However, when one market-based transaction is incorporated into the firm, it will cause an increase of coordination cost within the firm while it reduces transaction cost. So, the firm will tend to expand until the increase of coordination cost is equal to or larger than the decrease of transaction cost. With the assumption that human nature is opportunistic and boundedly rational, Williamson operationalized Coase’s verbal theory by developing his theory of asset specificity, according to which, the firm exists to avoid opportunism-related transaction costs; for a transaction involving specific assets, the more specific the assets to the transaction, the higher the potential for post-contract opportunism therefore higher transaction costs; and in such a situation, the natural solution to avoid such opportunism is to incorporate the assets into the firm. Different from Alchain and Demsetz (1972), both Coase and Williamson hold that within the firm the management has an organizational authority over the employees to direct them by fiat what to do and how to work; and Williamson further argues that the firm acts as an internal court for resolving disputes within the firm.

The KBT offers a different perspective which sidesteps the issue of opportunism. Kogut and Zander (1992) argue that the firm exists because it can provide a social community for sharing, creating and transfer of knowledge more effectively than through market. The boundary of the firm (or, the make-or-buy decision) is then dependent on three

1 Although Conner (1991) initially aims to argue for resource-based theory as a new theory of the firm in comparison to the transaction cost theory, Conner and Prahalad (1996) argue that ‘a knowledge-based view is the essence of the resource-based perspective’.
factors: the current knowledge, learning capacity and combinative capabilities of the firm. With regard to the internal organization question, the KBT emphasizes on ‘organizing principles’ which frame how individual and functional expertise is structured, coordinated, and communicated (Zander and Kogut, 1995). So, if the TCE is about ‘avoiding negatives’, i.e., opportunism and transaction cost, the KBT is about ‘creating positives’ (Conner, 1991: 140), i.e., gains from specialization in knowledge acquisition and storage (Grant, 1996:113).

What are their responses to each other? The knowledge-based theorists accept the validity of the TCE and argue that the KBT is a complement to the opportunism-based perspective (Conner and Prahalad, 1996), nevertheless, Conner and Prahalad (1996) argue that the knowledge-based considerations can not only explain the existence of the firm but also outweigh opportunism-related ones (i.e., TCE) as Conner (1991) tries to make a case that firms exist for reasons primarily related to “creating positives” rather than “avoid negatives” and argues that ‘unlike transaction cost theory, resource-based theory as theory of the firm does not depend on the presence of opportunistic potential’ (Conner, 1991: 144).

In contrast, the TCE defenders such as Foss (1996a, 1996b) dismiss the knowledge-based perspective being able to independently explain the key question of firm existence. Foss (1996a) made a ‘very powerful’ critique of KBT (Barney, 1996) by arguing that ‘it is erroneous to think that higher order organizing principles are a qualitative differential of the firm relative to the market, although they are probably in reality a quantitative one. That is, the qualitative presence of higher order organizing principles does not necessarily distinguish the market relative to the firm; rather, there may be “more of it” in the firm than in the market, as Arrow (1974) forcefully argued. The argument that firms better cultivate higher order organizing principles demands, however, precisely an argument from opportunism’.

However, in the literature, the TCE is very controversial and a target of heavy criticisms. For instance, Masten, Meehan and Snyder (1991) make it crystal clear that Coase’s original transaction cost argument is easy to make and impossible to refute, which implies the Coase’s transaction cost theory is not scientific as Popper (1934, 1959) has argued that scientific theories must be falsifiable. Barney (2001) echoes this charge and label it as ‘Coasian tautology’. Coase (1988) himself and Demsetz (1988: 147) argue that Williamson’s hypothetical linkage between asset specificity and transaction cost is weak, as Masten et al.’s (1991: 4) further point out that ‘a finding that higher asset specificity leads to a larger probability of integration could, in principle, obtain even if the hypothesis that asset specificity raises contracting costs were invalid’. But the most serious attacks come from Pfeffer (1994) and Ghoshal and Moran (1996). They argue Williamson’s version of transaction cost theory is ‘bad for practice’ because of the false or at least imprecise assumption about human nature, i.e., being opportunistic. According to psychology theories, people’s attitudes and behaviours are likely to be modified in response to their perceptions how they are treated by other people surrounding them. Negative perception derived from mistrust due to the opportunism assumption may reinforce people’s opportunistic attitude and behaviors through a positive feedback loop.
Therefore, the opportunism becomes a self-fulfilling prophecy. Due to the self-fulfilling propensity of the opportunism-based theory, it is very dangerous in management practice to design organizational control and incentive mechanisms according to this theory. The irony of the opportunism-based TCE theory is likely to cause more opportunistic behavior they theory is designed to reduce.

Even so, Foss (1996b: 519) still holds firmly the ‘conventional’ view that ‘we cannot do without concepts such as opportunism if we wish to explain the existence of the firm’ (Foss, 1996a: 474). In the meantime, despite Williamson’s (1999: 1093) critique of the KBT relying on ‘ex post rationality’, Foss (1996: 470) commends KBT that ‘properly interpreted, knowledge-based theories may help shed light on issues relating to the boundaries and internal organization of the firm’. For example, Grant (1996) is a good piece. What is interesting is that both Foss (1996a, 1996b, 1999) and Conner (1991: 143) see TCE and KBT complementary and share the view that an integration of the two perspectives may be fruitful for further theory development. One such integration effort results in the entrepreneurial theory of the firm.

The entrepreneurial theory tries to synthesize the theory of entrepreneurship and the theory of the firm. They follow Knight (1921) to view entrepreneurship as judgment, which they argue provides a natural answer to the question of the existence of the firm, namely, because judgment cannot be purchased on the market, the entrepreneur needs a firm - a set of alienable assets he controls - to carry out his function (Foss and Klein, 2005). The changes of firm boundary (i.e., via mergers, acquisitions, divestitures, and other reorganizations) ‘are best viewed as responses to a valuation discrepancy’ (Foss and Klein, 2005: 17) as the entrepreneurs expect efficiency improvement ‘by replacing poorly performing managers, creating operating synergies, or establishing internal capital markets’ (Foss et al., 2007: 1176). Regarding internal organization, they discuss how entrepreneurs process and synthesize information to identify new opportunities for coordination (Casson, 1998) or how entrepreneurs delegate judgmental decision-making rights to subordinates and how to manage the ‘derived’ judgment and entrepreneurship (Foss, Foss and Klein, 2007).

They see entrepreneurship as ‘a crucial but neglected element’ (Langlois, 2005) in explaining the nature of the firm because ‘the establishment of a new business venture is the quintessential manifestation of entrepreneurship’ (Foss et al., 2007: 1167), viz. ‘the firm exists because of entrepreneurship’ (Langlois, 2005). Therefore, entrepreneurship is ‘a necessary element in any comprehensive synthesis’ of otherwise ‘partial theories’ of the firm (Casson, 2005: 327-9). Casson (2005: 345) argues that since both the TCE and KBT ‘are compatible with the theory of entrepreneurship and can, therefore, be synthesized within this framework’, thus, the addition of entrepreneurship ‘is not minor refinement of the theory of the firm, but represents a radical change’ (Cass, 2005: 327).

In my view, all of the three perspectives are insightful yet partial. Both opportunism and entrepreneurship are necessary and complementary for explaining the existence of the firm because obviously we cannot exclude the two possibilities for which the first ever firm came to exist, i.e., the first ever firm may emerge out of an entrepreneur’s desire to
establish a firm or out of a buyer’s solution to avoid opportunism the seller might act. I agree with Foss (1996a) that KBT does not offer an ultimate answer to the existence question because in both situations in which the first ever firm may come to exist, knowledge consideration is not the fundamental factor. However, I argue the KBT may shed light on the entrepreneurship perspective, i.e., one possible rationale behind entrepreneurship is that the entrepreneur might believe in the potential synergy by effective utilizing others’ knowledge. So, it is clear there is a linkage between the three perspectives.

Nevertheless, I have a common criticism of the three perspectives, namely, their answers to the firm existence question clearly show their buyer/entrepreneur/firm-centrism, or, in Zajac and Olsen’s (1993) term, ‘a single-party’ emphasis that ‘neglects the interdependence between exchange partners’. For instance, the transaction cost theory is buyer-centric because it says the buyer decided to form a firm to avoid the costs of transacting with the sellers. It ignores the role of the sellers: why should the sellers obediently accept to become the buyer’s employees just because the buyer wanted to avoid some costs? The entrepreneurial theory is entrepreneur-centric because it says the entrepreneur decided to form a firm in order to have some people to make derived judgments for him. But, why should those people working freely on market obediently accept to become the entrepreneur’s employees to take on the responsibilities of judgmental decision making? Similarly, why should those people give up their freedom to become someone’s employees just because someone else wanted to take advantage of their own knowledge? All of them (TCE, KBT, entrepreneurial theory) are silent on such a question or might have never considered it an issue.

What all the existing theories miss is a theoretical consideration of the employees, i.e., those who accept to work for other people, which is an indispensible party of any economic transactual relationship. Obviously, those people who work with freedom on market will not obediently accept to become other people’s employees without benefiting from such a change of status. Some scholars including Coase’s (1988) and Ghoshal and Moran’s (1996) have called for an alternative theory. I argue, a new, comprehensive and integrative theory of the firm must explicitly address this ‘why becoming employee’ question. In this paper, I propose a relationship-based theory as such a candidate theory. As a hint, the concept of relationship indicates a consideration of both parties of a relationship. In the following sections, I will use the concept of relationship and the notion of relationship harmony describing the quality of a relationship to frame my whole argument, which I intend to integrate all the existing theories of the firm as well as some elements of the stakeholder theory, the latter being not primarily a theory of the firm in the strict sense, i.e., it does not primarily aim at addressing the key question of firm

---

2 Pitelis and Teece (2009: 10) clearly understand the problem of ignoring the role of employees as they critique the ‘efficiency gains from transaction costs…may not be an adequate explanation for employees voluntary accepting to work for employers’. However, their account is still entrepreneur-centric as they argue ‘firm exist because of the intentions of their principals-to-be’ (p. 11), i.e., ‘for their principals-to-be to capture value (profit from) their approvable value creating advantage’ (p. 10). They clearly refer firm principals-to-be as the entrepreneurs (p. 10) or owner-shareholders (p. 8).
existence, yet offering insights for boundaries, internal organization and competitive advantage issues of the firm.

3. Foundations of the R’BT

3.1 Assumptions about Human nature

The relationship-based theory of the firm (R’BT) assumes human nature is multifaceted and consists of two parts: a negative part and a positive one. The negative part is further divided into self-interest and opportunism. The positive part also has two subparts: cooperation and altruism, which some economists seem to recognize, e.g., Fehr and Gächter (2000). Therefore, Williamson’s assumption about human nature is very biased and ‘bad for practice’ because he solely focuses on the worst part of human nature – opportunism.

R’BT also assumes that human nature is not static but environment-sensitive. According to dissonance theory (Aronson, 1980; Festinger, 1957), there is a positive feedback relationship between human behavior and the modification of human attitude as a means of reducing dissonance (Ghoshal and Moran, 1996: 21). Since human behavior is often constrained by surrounding environment, human attitude is therefore modified by the external signal from the surrounding environment. These signals may include: how others surrounding you behave, how they treat you, and how they talk about human nature, etc.

To relax a bit our assumption about human nature being environment-sensitive, even if we cannot be sure if the two parts (and their subparts) can/will change in relative size, we are still confident in arguing that human attitudes derived from the respective parts of human nature can/will change. In social psychology, there is a crowding theory which argues human’s intrinsic motivation may be crowded out by extrinsic motivation, such as monetary incentive, and under some circumstances, it is difficult to restore the crowded-out intrinsic motivation even the extrinsic motivation is removed (Deci and Ryan, 1985; Lane, 1991; Lepper and Greene, 1978). This crowding-out effect is acknowledged by other organization scholars like Osterloh and Frey (2000) and some economists such as Frey (1994) and Frey and Oberholzer-Gee (1997).

There is also a crowding-in effect which allows external intervention, a reward or regulation, to strengthen or ‘crowd in’ intrinsic motivation (Frey, 1994: 335). Fehr and Gächter (2000: 159) point out that ‘many people deviate from purely self-interested behavior in a reciprocal manner’, namely, ‘in response to friendly actions, people are frequently much nicer and much more cooperative than predicted by the self-interest model’. Because of the propensity of reciprocal behavior, many social psychologists led by Martin Seligman, the then president of the American Psychological Association, have called for a positive psychology movement which is ‘concerned with what is right with people and building on that instead of trying to just fix what is wrong with people’ (Luthans and Youssef, 2004:151). Positive psychology has spurred two related movements that emphasize the positivity and strength of human being. The first is the positive organizational scholarship (POS), and the second the positive organizational...
behavior (POB) (e.g., Luthans, 2002a, 2002b). I argue organization and management studies in general urgency need a positive theory of the firm and management in order to escape from the ‘pathological spiraling relationship’ (cf. Enzle and Anderson, 1993; cited in Ghoshal and Moran, 1996: 25) or the ‘self-fulfilling prophecy’ of opportunism argument (Ghoshal and Moran, 1996: 21). The proposed R’BT is intended to be such a positive theory of the firm by purposefully emphasizing the positive side of human nature while downplaying the negative side.

3.2 The concept of relationship

According to Merriam-Webster online dictionary, relationship is defined as ‘the relation connecting or binding participants in a relationship’, which indicates ‘the state of being related or interrelated’. In this sense, I make clear from the outset that, the concept of relationship I use in the R’BT is very different from the one used in resource-based view (RBV) because the R’BT sees relationship as the dynamic relation between two parties rather than what the RBV talks about relationship as a particular type of resource the firm owns. This is why I term my relationship-based theory as R’BT, to distinguish from the RBV.

Apparently, there are two types of economic relationships, market-based and firm-based. Market-based relationships include arm-length exchange relationships and longer-term cooperative relationships between strategic alliance partners. Firm-based relationships can be divided into task-oriented and people-oriented relationships, the former concerns how people are related in their joint accomplishment of tasks while the latter concerns how people are related socially and psychologically.

Here, I summarize four different views of commercial relationship: contractual, capital, game, and option views.

The contractual view sees relationship as a kind of contract because commercial relationships are based on contracts, explicit or implicit. Kay (1993) identifies three types of commercial contract: spot, classical and relational contracts. Many scholars from different disciplines have discussed the concept of relational contract, such as law scholar Macneil (1974, 1978, 1980), economists like Olson (1965) and Grossman and Hart (1983, 1986) and sociologists like Bradach and Eccles (1989), Fox (1974), Granovetter (1985), McAuley (1963), Thompson et al., (1991) and Zucker (1986). In the R’BT, any relationship is underpinned simultaneously by two or three types of contract in Kay’s typology. Take the example of a firm’s relationships with its customers. When a customer buys a firm’s products from the market, the buying-selling is a spot contract. Meanwhile, regulations of consumer rights protection serve as the legal contract between the firm and its customers, although this is often not discussed during the buying-selling process. Moreover, there might be a potential relational contract between the firm and its customers in that, if the firm takes care of its customers, they may become loyal to the firm. Another example is the relationships between a firm and its employees. The employer-employee relationship involves the legal employment contract as well as a psychological contract. If the firm looks after its employees more than the employment
contracts require, the employees will be more satisfied psychologically, hence they would be willing to put more effort in their jobs as reciprocity, the commitment, productivity and creativity of the work force may thus be increased.

The capital view pays attention to the potential capital embedded in relationships. The modern concept of capital encompasses four types: economic capital, human capital such as knowledge and skills, social capital such as networks and trust, and positive psychological capital such as confidence, hope and optimism (Luthans and Youssef, 2004). While economic capital and human capital can be generated independently by individuals, social capital can only be captured from embedded resources in a social network (Lin, 2001) of relationships. Lazega and Pattison (2001) propose that social capital, as resources embedded in social network, includes information, colleagues’ goodwill, advice and sometime emotional support, as well as many other means that serve individual and collective ends. Therefore, the social capital can enhance mutual trust and cooperation and therefore strengthen the interdependence between two parties, leading to the increase in commitment, productivity and creativity of individuals. The concept of relationship in R’BT includes social capital as well as the positive psychological capital because the development of a person’s confidence, hope, optimism and resiliency (i.e., positive psychological capital) is unlikely if he or she works in a very discouraging and conflicting environment. So a harmonious relationship is a precondition for developing the positive psychological capital.

Economists and management theorists frequently use game theory to analyze economic or business relationships (Casson, 2000; Kay, 1993; McMillan, 1992; North, 1990). IBM’s founder Thomas J. Watson Sr. once spelled out that ‘business is a game, the greatest game in the world if you know how to play it’ (cited in McMillan, 1992). Kay (1993) uses game theory to explain the cooperation and coordination of business relationships. Many scholars point out that mutual trust is the cornerstone of long-term relationship (Fuller and Lewis, 2002; Jarillo, 1988; Johannisson, 1987; Kay, 1996; Lorenzoni and Ornati, 1988). From a game theoretical perspective, building relationship is actually a repeated game in which trust, reciprocity, and mutual gains are necessary for an ongoing cooperation. The Japanese corporate system has had a harmonious capital-labour relationship owing to the cultural emphasis on trust (Matsumoto, 1991). A stable long-term relationship is hardly possible if the benefit of engaging in a relationship is only partial rather than mutual. In a sense, people or organizations establish or join in relationships for the sake of gaining benefits from the relationships. Chen (2004) explains why a guanxi network is vital to Chinese people and points out that the favour exchange is the norm of building a guanxi or a ‘guanxi web’. Su and Littlefield (2001) analyze the favour-seeking and rent-seeking dimensions of guanxi and claim reciprocity is the core of a guanxi. Larson (1992) also suggests that reciprocity norm is one of important factors underpinning the duration and stability of relationship.

The option concept was borrowed by strategy scholars to explain strategic management as ‘a process of organizational resource-investment choices, or option’ (Bowman and Hurry, 1993: 760) (e.g., Bowman and Hurry, 1987; Hurry, 1993; Hurry et al., 1992; Kester, 1984; Kogut, 1991; Myers, 1984; Sharp, 1991). In financial market, the
development and use of option contracts were motivated by risk-averse people’s intuition of ‘keeping options open’ against the uncertainty of future (Bowman and Hurry, 1993: 761). It can be said that strategies emerge from investment in resources which generate options for future investments. In the R’BT, relationships can also be understood from the option perspective. When two parties establish a relationship, each side does not know how the other side will behave in the future. Due to risk-aversion consideration, both parties want to strengthen the relationship through investing in the relationship, resulting in a gradual accumulation of each party’s sunk cost of engagement. This incremental investment in the established relationship serves as an effective mechanism to reduce the opportunism. Another advantage of building relationship is to retain the option for future opportunity deriving from the current relationship. This is evidenced in the Chinese practice of building ‘guanxi web’ – a network of personal relationships based on reciprocity (Luo, 1997). By investing in different relationships, they essentially keep their future options open. While the western approach of relationship building is intended as an option to avoid risk and contain opportunism, the Chinese guanxi network is used as an option for access to future opportunities. This demonstrates the complexity and casual ambiguity (Einhorn and Hogarth, 1986) of relationship options compared to contractual options (i.e., option contracts). In this sense, we can say that the relationship engagement is more sophisticated than other strategic options (Bowman and Hurry, 1987).

3.3 The concept of harmony

When it comes to describing the state of a relationship, two opposite words naturally spring to our minds: conflict and harmony (Cheng, 2006). Conflict is normally associated with competition while harmony with cooperation. Scholars have often called for cooperation (Child et al., 2005; Doz and Hamel, 1998), especially, in the firm-stakeholder relationships (Gulati and Kletter, 2005; Jones, 1995; Jones and Wicks, 1999; Rindova and Fombrun, 1999; Ring and Van de Ven, 1992). Relationship harmony is viewed as crucial in the Confucian societies. Obviously, harmonious relationships can effectively reduce (albeit not to eliminate) some contracting costs associated with agency, shirking, underinvestment and opportunism due to the enhanced trust and cooperation (Jones, 1995). Therefore, harmony rather than conflict in relationships is desirable and an enduring goal to pursue for the majority of human beings.

But, what is harmony? Although there might be various definitions of the concept, I don’t want to dig into the philosophic meaning of harmony, instead, in this paper, I offer a practical perspective on harmony in economic relationship. To avoid the difficulty of defining a public concept, I choose to describe the state of harmonious relationship, i.e., the quality of relationship (Ghoshal and Moran, 1996: 37), by dimensionalizing the concept with different aspects. Jones (1995) and Jones and Wicks (1999) point out trust and cooperation are fundamentally important in managing stakeholder relationships. In addition, how both parties of the relationship perceive whether their respective goals fit each other, viz. whether they have shared interests, is also very important for the ongoing relationship (Ghoshal and Moran, 1996: 36). Accordingly, I dimensionize relationship harmony with three variables: shared interests, mutual trust, and obligation fulfillment.
This means, if the two parties of a relationship have shared interests, mutual trust and can fulfill their own obligations, their relationship is harmonious. Obviously, there is a degree issue, namely, the more shared interests, the more mutual trust and the better they fulfill their obligations, the more harmonious their relationship is. I argue, only when all three variables reach a certain degree, the minimalist harmony state will come into being. This is to say, if any one variable scores too low, it will result in disharmony even if the other two variables score relatively high. For instance, if there is a conflict of interests, i.e., lack of shared interests, then there can hardly be mutual trust and obligation fulfillment. If one party does not well trust the other, i.e., low mutual trust, no matter how well the other party fulfils its obligation, their relationship can hardly be harmonious. Similarly, if one party often does not fulfill its obligation, no matter how good their relationship initially may be, the other party will gradually lose trust in it, i.e., mutual trust damaged. To put it another way, if any of the three variables scores lower than a minimal level, the relationship is not harmonious, viz. there is relationship disharmony.

I also distinguish two types of harmony: one being a realistic harmony in which no body is unhappy about the relationship, the other an idealistic harmony in which everybody is happy about the relationship. To maintain a stable relationship, a realistic harmony is necessary. To build a health relationship, an idealistic harmony should be ceaselessly pursued.

4. The nature of the firm

4.1 Existence of the firm

Market-based relationships are normally not harmonious, due to the lack of shared interests (i.e., conflict of interests) and mutual trust between the buyers and the sellers, and opportunistic potential on the part of the sellers. Due to such relationship disharmony embedded generally in market-based relationships, neither side of a market-based relationship can gain maximized self-interest out of their relationship. The R’BT argues the firm came to exist to remedy the problem of relationship disharmony and to provide a vehicle for both parties to gain more self-interests than what available on market.

The transaction cost theory implicitly assumes that there had existed market transactions before one party of the transaction decided to employ the other parties to form a firm. In so doing, the TCE naturally precludes the role of entrepreneurship in explaining the emergence of the firm. The R’BT accepts this assumption (i.e., market transaction preceding the establishment of the firm) as just one possible situation in which the first firm came to exist. The other possible situation the R’BT acknowledges is that the first firm came to emerge out of the first ever entrepreneur’s personal desire. Therefore, the R’BT is to offer a new solution encompassing both opportunism and entrepreneurship perspectives. But, what is new here in comparison to the TCE and entrepreneurial theory is that the R’BT does not believe that it is either the high transaction costs or the needs of delegating judgmental decision rights that characterizes the situation out of which the first firm was born. As we can imagine, at the time when the first ever firm came to emerge, the business should be not too complex as in modern time, and what the
entrepreneur needed might be simply finding some laborers in local neighborhood to do
some not so complex tasks under the entrepreneur’s intention and direction. Therefore,
the costs of searching for labor providers should be fairly low and there might not need to
be any formal contracts therefore costs of contracting could be ignored. The monitoring
cost should also be a trivial issue too. Similarly, it is highly unlikely that the first ever
entrepreneur wanted and needed other people to make judgmental decisions for him/her.

Then, what is the relationship-based theory of firm existence. Let us start from the first
situation where prior market transactions existed. Imagine a person A who needed some
one to do some things for him. He came to the market and found person B. A told B
about his purpose and asked if B was willing to do so. While A was willing to pay $0.8,
B asked for $1.2 for that task. After negotiation, they agreed the deal at $1. Then B did it
and got $1 from A. At the end of this deal, A wanted to continue this transaction and B
accepted. In the extended deal, A was not willing to pay more for the same service while
B was of course not going to do more service for A for the same amount of pay. So, their
market-based transactions continued as usual for a while until one day B started to shirk.
The possible reason for such shirking behavior might be either due to B’s opportunistic
nature or B’s perception that his labor was underpaid. Why such perception? It is because
B might believe his service well deserved $1.2 as he initially asked for but rejected by A,
or at least was worth of more than $1 otherwise why A wanted to continue this deal. Due
to B’s shirk, the service he provided in this task then was less than before, which was
then detected by A. A was unhappy with the outcome and rejected to pay the same fee as
before, and then A cut the payment to $0.9. In the following process of transactions, the
similar behaviors happened again, and a vicious circle occurred. So, the service B
provided to A became less and less while the fee A paid to B reduced accordingly until
one day the transaction was discontinued due to a certain degree of disharmony
accumulated in their relationship.

In the above story, we see two stages of the market-based transaction between buyer A
and seller B. The first stage is the $1 deal where no opportunism is involved. The second
stage is the period where the vicious circle of $1 → $0.9 → $0.8 took place due to the
opportunism on the part of the seller. In each stage, the buyer A might feel dissatisfied
with the deal because what he could gain from the market-based transaction was no more
than $1. In the meantime, neither can the seller get more than $1 payment. Imagine, if A
was at the same time in deals with more than one seller, say B and C. In this situation,
what A could get out of the market-based transactions with B and C was no more than $2
while he paid out no more than $2 as well. In the first no opportunism involved stage, the
buyer A might suddenly have had an idea that if employed by him the sellers B and C
might be able to generate more than $2, say $3 valued service, due to buyer A’s effective
management and coordination between B and C. Then there will be $1 extra generated by
such employment relationships. Buyer A thought if he was to offer $1.2 each to sellers B
and C for the same tasks, then B and C would be willing to become A’s employees with
an agreement to accept A’s authority what to do and how to do. Then, in the end, buyer A
could still gain extra $0.6 (= $1 - $0.2 x 2) while sellers B and C each gain $0.2 extra.
With such an idea, buyer A approached sellers B and C with such idea. Since all of them
could benefit from the new employment relationships, they all accepted. So, the first ever
firm emerged because it solved the problem of relationship disharmony inherent in 
market-based relationships and provide a mechanism for all the parties involved to gain 
more self-interest benefit than whatever available on market.

If the firm emerged in the first stage of the market-based transaction, i.e., $1 deal without 
opportunism involved, then buyer A’s entrepreneurship is the sole determinant of the 
emergence of the firm. In contrast, if the firm emerged in the second stage, i.e., 
$1 \rightarrow $0.9 \rightarrow $0.8 vicious circle, then it is either or both of opportunism avoidance and 
trepreneurship of buyer A that made the firm emergence possible. In the case of 
trepreneurship, the KBT can shed light on the mechanism of synergy generation, i.e., 
$1 + $1 = $3 > $2$, because working together and under direction, the personal knowledge 
resided in different people can be better shared and combined to improve economic 
efficiency, as first argued by Adam Smith (1776) with the idea of specialization and 
coordination. This is how the R’BT integrates the opportunism, entrepreneurship and 
knowledge perspectives as well as taking into consideration of the role of employees.

4.2 Boundaries of the firm

Since the R’BT argues that the firm is a better governance mode to increase self-interest 
gains of all parties of transactual relationships, then why not organize all transactions in a 
big firm? This question can be divided into two interrelated sub-questions: first, why not 
just one big firm without any markets? second, why there are other firms competing with 
each other? For the first, we can follow Coase’s (1937) logic, managing many 
transactions within the firm will incur coordination costs, if the increase of coordination 
cost outweigh the increase of self-interest gain on the entrepreneur side, then the 
trepreneur will stop to incorporate extra transaction into the firm. For the second sub-
question, the answer is that when many transactions (and the people involved) are 
incorporated into the firm, the employees may find the relationships in the firm not 
harmonious or less harmonious than before or in other firms, they may choose to leave, 
either going back to market or going to other firms. Why?

R’BT argues the firm tends to expand as long as it can attract more people and maintain a 
certain degree of harmony in their relationships as well as provide better chances for 
newcomers’ personal goal realization. But the firm won’t expand infinitely because 
existing relationships may be disharmonized with the gradual expansion of the firm for 
different reasons. For example, many long-time employees may feel they are less cared 
about by the central management if they used to enjoy much more close relationships 
with the management. Another reason might be more people means less benefit can be 
shared if the profit growth is slower than the payroll expansion. So, the firm size expands 
until the number of insiders quitting the firm due to perceived decreasing harmony by 
expansion is equal to or greater than the number of newly recruited with a perception that 
this firm may offer them a better chance for realizing their personal goals than anywhere 
else or working alone.

Above logic can also be applied to groups of peoples (i.e., subunit within the firm). If the 
top management fails to maintain harmonious relationships with its subunits, the subunits

Department of International Economics and Management, CBS, Working Paper wp1-2010-xl 12
may quit via management buy-out (MBO) or with other solutions. On the other hand, if
the central management perceives there is no synergy or even disharmony between a
subunit and the rest of the firm, the firm may divest it. Understandably, the divesture of
subunits can be done for those units purchased from market, which implies the firm
might have misjudged the potential for enhancing harmony in the relationships by
shifting the inter-firm relationships into intra-firm relationships, or even if the initial
judgment was not a mistake by bringing them into the firm, the switched relationships
turned out to be less harmonious than prior arm-length relationships.

With regard to vertical integration, it is still the relationship between the firm and its
business partner (another firm) that dictates the acceptable governance structure. Say, the
firm has so far had a harmonious relationship with one supplier. If there is no any
disharmonious contingency emerging in their relationship, they will carry out their
businesses as usual. Imagine, one day, if the supplier suddenly requests the firm to pay an
extra fee for the same transaction carried out before. Whether this request will bring
disharmony or not into their relationship will then depend on how the firm perceives this
request. There are several possibilities: firstly, if the firm perceives it as fair because the
supplier might have had an increased cost for providing the same product or service due
to the increased costs of raw materials and labors, then in order to maintain a harmonious
relationship, the firm will fully accept this extra fee request; secondly, if the firm
perceives the request is not fair because the extra fee requested is higher than the
increased cost to the supplier, then there are two options for the firm to choose, for one, if
the extra fee is acceptable, the firm may tolerate and accept this request for the sake of
maintaining the relationship though a bit disharmony will be built into their relationship;
for another, if the extra fee is not acceptable, then the decision of the firm depends on
how important this buyer-supplier relationship to the firm: if it is not important and if it is
possible to find another supplier for the same product and service, the firm will choose
switch to another supplier; if the relationship is very important or there is no other
suppliers for the same product and service, the firm may choose to purchase this supplier.

4.3 Internal organization

R’BT offers another viewpoint on authority in the firm. Rather than lean towards either
Coasian fiat-type authority of the employer over employees or Alchian and Demsetz’s
(1972) no authority at all argument, the R’BT argues the owners or their representatives
(i.e., managers) do have authority over the employees or subordinates because by joining
the firm the employees voluntarily accept some kind of organization authority. But the
authority has its limit. According to Simon (1951, 1995), the employer’s authority over
its employees is restricted by the employees’ ‘zone of acceptance’ under the assumption
of utility maximization rather than the neoclassical assumption of profit maximization
(Simon, 1995: 279) where ‘profit [consideration] is only one among a number of [other
considerations] and enters into most subsystems [of the firm] only in indirect ways’
(Simon, 1964: 21). And the real authority lies in the capacity for the employer, under the
precondition of maintaining a harmonious relationship, to prioritize its organizational
goals over its employees’ individual goals.
This prioritization-under-harmonization capacity based authority indicates that if the employer abuses its bargaining power, if its fiat falls outside the employees’ ‘zones of acceptance’, the employees may choose to leave the firm, and then there will be no authority over them anymore. But, if the employer firstly builds and harmonizes its relationships with its employees, then the employer can more easily prioritize its organizational goals over employees’ personal goals, and the employees tend to accept such authority of prioritization.

R’BT also does not prefer the formal control argued for by TCE scholars, especially Williamson (1975, 1985). In China where relationships and harmony are highly appreciated and cultivated the management philosophy goes it is not impossible to achieve a perfect type of control: governing without intervening. Governance without interference can be understood like this: don’t interfere your subordinates, if you selected them, trust them, empower them, and encourage their creativity, and then they might get tasks done effectively, efficiently and creatively. If leaders at each level follow this ‘governing without interference’ principle to manage their subordinates who in turn manage their own tasks, then the whole company can be governed without interference.

But this is not to deny the necessity and value of formal mechanism to exert a certain degree of control on the employees. As mentioned before, there are two types of relationship in the firm: task-oriented and people-oriented. The R’BT argues, to harmonize the task-oriented relationships, the firm needs to establish certain formal regulations such as award-and-penalty rules and to implement certain management tools such as enterprise resource planning (ERP) system. For the people-oriented relationships, it is better to rely on more informal means to harmonize them, for example, to facilitate socialization among the employees, to take care of employees and fellow colleagues, to enhance mutual trust, and to nurture a friendly and flexible corporate culture, etc.

4.4 Competitive advantage

Economics is concerned with efficiency (i.e., best, optimal) use of limited resources (Marschak and Radner, 1972: 3). However, many economists have paid more attention on economic resources such as physical and financial resources than on human resources. R’BT argues human resources of an individual such as time, effort, knowledge and creativity are scarce resources to himself. Where to work, what to work with, and how to work are very important decisions which will determine the efficiency of the use of his human resources. Obviously, working in a disharmonious environment will hamper the allocation efficiency of human resources. So in order to improve the allocative efficiency, the human resources must find a harmonious work place. This explains why we say the firm has an advantage over the market in helping the individuals to gain more self-interests out of their commercial relationships.

An economic organization in general strives to make economic rent, i.e., above average profit. There are four major types of economic rent: first, Porterian rent derived from strong bargaining power in contractual relationships. According to Porter’s (1980) model of five forces of industry, we understand that a firm can squeeze or appropriate the value-
added of its suppliers and customers, and its employees as well. Obvious, this rent is generated at some other’s expenses; second, Ricardian rent derived from superior resources the firm controls, which is argued for by the resource-based view; third, Schumpeterian rent derived from entrepreneurial innovations or creative destructions (Schumpeter, 1934) which generate more value than existing economic system; the last one, I term as Confucian rent. Under harmonious relationship people work in a more committed, more productive or/and more creative way than otherwise if there is no harmony in the relationship. The extra output above normal generated by working more committedly, productively and creatively is the Confucian rent.

But above analysis does not explain the competitive advantage of one firm compared to another firm. The R’BT explains the competitive advantage of a particular firm by arguing that the more harmony in the firm, the more potential to have a competitive advantage over others for two reasons. First, harmonious relationships within the firm allow the firm to better utilize the current human resources available to the firm. Second, more harmonious relationships within the firm compared with others help the firm to attract talented people from other places (market or other firms) to join the firm due to its good reputations and those people’s expectations of better utilizing their own human resources. The R’BT asserts the fundamental and ultimate competitive advantage of the firm rests in its ability to effectively attract, retain and utilize the talented people. Relationship harmony is then a determining factor underpinning that ability.

R’BT argues harmonious relationships provide the firm a relationship advantage, i.e., the relative high potential for the employees’ reciprocal efforts to help the firm to realize its organizational goals, which can be argued as the foundational level advantage of the firm, while the low cost and differentiation advantage are the end level competitive advantage. Any other competitive advantages in between these two levels are categorized as intermediate level advantages, such as innovative, entrepreneurial and quality advantages. As efficiency of allocation is not equal to efficiency of utilization, so the R’BT further argues there must be a transformational mechanism for the firm to encourage the commitment, productivity and creativity of the employees and by doing so the firm can then transform the foundational relationship advantage through intermediate level advantages into the ultimately end level advantage, i.e., low cost, or differentiation, or combination of the two.

5. Conclusion

In this paper, I argue that all the existing theories of the firm, i.e., the TCE, KBT, and entrepreneurial theory, are insightful yet partial since they only deal with one or another possible situation in which the first ever firm came to exist. In addition, all of them have a common problem of buyer/entrepreneur/firm-centrism because they all ignore the role of employees in the formation of the firm (cf. Zajac and Olsen, 1993). I argue, since they are complementary to each other a new comprehensive and integrative theory of the firm must be able to unify them with a consideration of the employees. The relationship-based theory of the firm is such a candidate theory because the concept of relationship naturally links the entrepreneurs with the employees and it is compatible with transaction and
knowledge (i.e., capital) perspectives. The R’BT also affirms the significance of entrepreneurship in explaining the existence of the firm because without entrepreneurship there would be no such a thing called the firm.

Since the R’BT places the firm-stakeholder relationships at the centre of its whole theoretical framework and prescribes the value of relationship harmony, this new theory of the firm may have important implications for management. If we follow the logic of the R’BT, then the essence of management is a total relationship management (TRM) and what we need and should strive for is a never-ending harmonization of the firm-stakeholder relationships. Therefore, this paper also makes a contribution in a way to the stakeholder argument.

Reference


Department of International Economics and Management, CBS, Working Paper wp1-2010-xl


Popper, Karl (1959) *The Logic of Scientific Discovery*. New York, Basic Books (originally published in 1934 in German titled as *Logik der Forschung*)


