Institut de Recherche
IRMAPE
en Management et Pays Emergents
GROUPE ESC PAU

CAHIERdeRECHERCHE N°9

Institut de Recherche en MAnagement et Pays Emergents du Groupe ESC PAU

The Groupe ESCPAU Institute for Research in Management and Emerging Countries

décembre**2007**

Préambule

« Ecrire n'est pas l'élément neutre de la Recherche mais au contraire contribue à la révéler. Alors, considérer dès le départ l'écriture comme ressort essentiel du travail à effectuer permettra d'en déployer toute la potentialité formatrice.

Une telle démarche ne va pas de soi. Aussi, il convient de la mettre largement à l'épreuve, d'en expérimenter les effets dans d'autres contextes, d'accompagner sa recherche avec elle et de mettre à profit différentes rencontres possibles autour d'elle, le retour de la part des autres étant essentiel pour savoir ce que l'on a écrit ».

L'écriture produit ou modifie la pensée et ses formes.

Yves REUTER

Composition et présentation du Comité de Lecture

Monsieur **Gérard HIRIGOYEN**
Président du Pôle Universitaire de Bordeaux

Monsieur **Paul LE CANNU**
Professeur à l'Université de Paris I - Panthéon Sorbonne

Monsieur **Gérard DENIS**
Professeur et Doyen de l'Université de Pau et des Pays de l'Adour

Monsieur **Bertrand SAPORTA**
Professeur Emérite de Sciences de Gestion à l'Université Montesquieu-Bordeaux IV

Monsieur **Christian PRAT DIT HAURET**
Maître de Conférences à l'Université Montesquieu-Bordeaux IV

Monsieur **Thierry POULAIN REHM**
Maître de Conférences à l'Université Montesquieu-Bordeaux IV

Monsieur **ALI YAKHLEF**
Professeur à l'Université de Stockholm

Madame **Anne NELSON**
Professeur en Systèmes d'Informations à High Point University

Sommaire

The Effects of Rewards on the Motivation of Experts to Transfer their Knowledge Page 5

Par Dr. Ali Yakhlef, Laurent Sié, Dr.Philippe Julien

A continuous approximation procedure for determining inventory distribution schemas within supply chains Page 21

Par Nikhil A. Pujari, Trevor S. Hale, Faizul Huq

La critique de la théorie de l'exploitation Page 47

Par Dr. Patrice Cailleba

The Role of Formal and Informal Chinese Institutions in Shaping the Entrepreneurial Landscape Page 63

Par Jean-Michel Quentier

The Effects of Rewards on the Motivation of Experts to Transfer their Knowledge

By Dr. Ali Yakhlef
Director of Research
Stockholm University School of Business / Sweden

Laurent Sié
MBA Manager
Groupe ESC PAU / France

Dr.Philippe JULIEN
Vice President Upstream Coordination
TOTAL S.A. / France

Article accepted by the 8th EURAM (European Academy of Management) Conference, 14th - 17th May, Ljubljana and Bled Slovenia

Abstract:

Theorists have taken a keen interest in the factors that boost or diminish employees' motivation to share knowledge. Of a crucial concern is: what are the effects of rewards on knowledge transfer? Why do rewards sometimes undermine and sometimes enhance employees' disposition to share their knowledge? The present paper suggests that the passion for knowledge is a predictor of an intrinsically-driven behavior, which is, in its turn, a predictor of a knowledge-sharing behavior. When extrinsic rewards boost self-regulation, a feeling of competence and self-relatedness, they may enhance knowledge sharing because they provide a fostering-ground for curiosity and passion.

Keywords:

expertise, passion, knowledge transfer, intrinsic motivation, Self Determination Theory

Introduction

It is conventionally held that firms can choose between two types of knowledge-based strategies: exploration vs. exploitation (March 1991). Both strategies require a great deal of knowledge transfer and sharing among organizational members. Whereas the sharing of explicit knowledge saves time, the sharing and transfer of tacit knowledge improves quality, increases innovation and signals competence to customers (Haas and Hansen 2007). Hence, research into the determinants that affect initiatives to share knowledge has brought to the fore the significance of motivation (Argote, McEvily and Reagans 2003). Motivation is the energy that activates individuals towards reaching an end. While the organizational benefits of knowledge sharing are recognized, it is not clear why organizational members would be motivated to willingly share their knowledge with others if they are not rewarded for that. Hence, invoking the logic of contribution-reward systems, research tends to emphasize the significance of rewards as a device to increase employees' motivation to transfer and share knowledge (Menon and Pfeffer 2003). In general such rewards range from tangible benefits, such as money, to intangible ones, such as social recognition and acknowledgement (Gagné and Deci 2005).

However, most studies tend to treat motivation as a discrete, unitary phenomenon, in the sense that people are motivated or not motivated. As a response to this line of criticism, researchers are suggesting that motivation is a continuous and multi-layered concept, implying that individuals can be more or less motivated, and that there are different forms of motivation (Quigley, Tesluk, Locke and Bartol 2007). In particular, researchers make a distinction between a psychological type referred to as intrinsic, and an economic one, or extrinsic motivation (Deci 1975; Ryan 1998). Whereas motivation is intrinsic when an "activity is undertaken for one's immediate need satisfaction" (Osterloh and Frey 2000 p.539), motivation is extrinsic when it 'serves to satisfy indirect needs, for example money" (Osterloh 2007 p.7). Furthermore, researchers assume that there is a tension between the two forms of motivation. If, with regard to intrinsic motivation, the ideal incentive system would lie in the meaningful work content itself (Osterloh and Frey, 2000), attempts at controlling intrinsically motivated behaviors (through, for instance the use of economic incentives and rewards) may have a negative, crowding-out effect that may reduce performance (Osterloh and Frey 2000). Adam meta-analysis of 128 laboratory experiments confirmed that tangible rewards undermine intrinsic motivation (Deci and al. 1999). Because intrinsic motivation is regarded as crucial in connection with the transfer of tacit knowledge, it is relevant to investigate how external rewards impact upon the transfer process. As noted by Chesbrough and Teece (1996), when the transfer of tacit knowledge is at stake, introducing market mechanisms is bad advice. If the transfer of tacit knowledge requires intrinsic motivation and if intrinsic motivation cannot be compelled, the transfer of tacit knowledge is then more difficult than previously recognized. How should companies motivate their employees to share knowledge is still an open issue. What impact do rewards for sharing tacit knowledge have on employees? Why do incentives have a positive effect on employees in some contexts and not in others? In what cases do empowering or controlling rewards boost or diminish the motivation to share knowledge? Why can rewards sometimes undermine the self-confidence and self-efficacy of employees or enhance their sense of competence and social esteem which negatively or positively may affect the motivation to transfer knowledge?

The present paper is an attempt to address these questions, emphasizing the interactive interplay between the type of knowledge to be transferred, the source's motivation to transfer knowledge and the form of rewards that animate them. Our assumption is that 1) tacit knowledge is more likely to have affinity with intrinsic motivation; for that reason; 2) it is more readily sharable and diffusible among individuals; and 3) that rewards are more associated with intrinsic satisfaction of the work itself. To understand this dynamics, we have interviewed 28 departing experts within the field of geo-sciences who work with newly recruited-engineers. These experts are supposed to transfer their knowledge which they have acquired through many years of experience. Given that 30 per cent of the company's experts will retire in the next five years, and given that their expertise knowledge involves a form of tacit and sticky (hard to transit to other) knowledge (Szulanski 1996), finding ways of promoting and facilitating the transfer process is of a strategic importance for the company. What would be the effects if the company introduces extrinsic rewards?

The argument in this paper will unfold in the following way. In the next section, we review previous accounts on motivation in the context of knowledge transfer, pointing out the relative neglect of intrinsic motivation by researchers and the effects of managerial attempts to control it (Section 2). In section 3, we present the method and the empirical data elicited from our respondents. Our analysis (section 4) shows how the experts under consideration are driven not by economic rewards, but rather by intrinsic motives that draw on the degree of perceived autonomy in defining and deciding on their tasks, the degree to which they perceive themselves (and are perceived by their peers) as competent, and the degree to which they feel part of a social network. Finally, in Section 5, implications of the present study for the theory and practice of knowledge transfer are drawn.

1. On experts, tacit knowledge and motivation

From a psychological perspective, experts are defined in terms of intrinsic characteristics because they invest time and interest in their expertise. This is so, because as Ericsson and Lehman (1996) point out, experts acquire their expertise through extensive deliberate practice. Deliberate practice is defined "as the individualized training activities especially designed by a coach or teacher to improve specific aspects of an individual's performance through repetition and successive refinement" (Ericsson and Lehman 1996: p.278). It is not the capabilities of persons that enable people to become experts but it is the extended intense practice which causes physiological, anatomical, and even neurological adaptations in the body (Ericsson and Lehman 1996). The term "expert" is rooted in the Latin adjective "expertus", meaning experienced in and with something (Smith, 1991). In their study of the difference between expert and novice, Tanaka et al (2005, p.145) consider that: "an obvious difference between experts and novices is that the former has greater exposure to objects from their domain of expertise than do novices". Hinds et al (2001, p.1236) add that "we consider experts to be those with both knowledge and experience in applying this knowledge to a variety of problems within the domain".

Expertise involves using one's abilities to acquire, store, and utilize at least two kinds of knowledge: explicit knowledge of a domain and implicit or tacit knowledge of a field (Sternberg et al., 1995). Thus, the mode of acquiring expertise alsobecomes a crucial feature of what expertise is: in fact, experts learn by doing (Greeno and Simon 1998). For Mieg (2001, p.2) "Experts in the literal sense are experimentalists: they know from active, reflexive experience". Foley and Hart (1999: p.234) summarize what they understand by expert in the following way: "someone who has attained a high level of performance in the domain as a result of years of experience". Since expertise comes with years of experience and since it is based on learning-by-doing, it is related to the form of knowledge that Polanyi (1966) refers to as tacit knowledge. Furthermore, because it requires a great deal of investment and interest in an area, we can assume that it is animated by a strong drive of intrinsic motivation.

1.1 - Intrinsic-Extrinsic Motivation

The term 'motivation' in the knowledge management literature is generally used to refer to an individual's motives that are designed to benefit the self (e.g. extrinsic vs. intrinsic, Osterloh and Frey, 2000) and to achieve a certain degree of satisfaction. Social and monetary rewards have been found as motives that facilitate and encourage knowledge transfer within organizational settings (Argote et al 2003; Menon and Pfeffer, 2003). For instance, individuals may be primarily motivated to share information and knowledge because they want to accrue recognition and rewards or because of a desire to help their colleagues (Burgess, 2005). The implication is that if the paid rewards and the perceived rewards are not in balance, the intended plans to transfer knowledge will most likely be doomed to failure (Osterloh et al., 2002).

Correlatively, absence of motivation has negative effects on knowledge transfer (Burgess, 2005; Osterloh et al., 2002; Osterloh and Frey, 2000). Researchers have stated many reasons to explain why a person may not be willing to transfer and share their knowledge with others: "A source of knowledge may be reluctant to share crucial knowledge for fear of losing ownership, a position of privilege, superiority; it may resent not being adequately rewarded for hard-won success; or it may be unwilling to devote time and resources to support the transfer" (Szulanski 1996, p.31). This view presumes that motivation is an explicit psychological state which consciously guides people's behavior. This may not always be the case since we do not sometimes know what really motivates different individuals (Burgess, 2005). Furthermore, research into knowledge-sharing behaviors is limited, for instance, we cannot explain the relationship between incentives to encourage knowledge transfer and outcomes (Brown and Duguid 1991; Quigley et 2007).

In their study of the dynamics of motivation, Osterloh et al. (2002) emphasize two types of motivation: intrinsic and extrinsic motivation and the importance of balancing between them. Individuals are extrinsically motivated if they are able to satisfy their needs indirectly, especially through monetary compensation (Osterloh and Frey, 2000). "Money as such does not provide direct utility but serves to acquire desired goods and services, an assumption which is basic to all economic analysis" (Osterloh et al. 2002, p. 64). Extrinsically motivated coordination in firms is achieved by linking employees' monetary motives to the goals of the firm (Osterloh and Frey, 2000). Here the ideal incentive system is strictly pay-for-performance. Extrinsic rewards can take the form of recognition, increased responsibility, advancement, better supervisory relations, better peer relations, increased pay, or job security (Dermer, 1975). Extrinsic rewards are also a good signal that conveys the idea that the time experts spend sharing knowledge with new comers is highly valorized by the organization (Kogut and Zander, 1992, 1993; Pan and Scarbrough, 1999), and that experts' efforts are appreciated.

In contrast to extrinsic rewards, intrinsic motives involve an activity that is undertaken for the immediate satisfaction of one's needs (Osterloh and Frey, 2000). Intrinsic motivation 'is valued for its own sake and appears to be self-sustained' (Deci 1975, p.105, in Osterloh et al. 2002, p. 64). Intrinsic motivation can be directed 1) to the activity's flow - for example reading a book, 2) to a self-defined goal - for example climbing a mountain, or 3) to the obligations of personal and social norms for their own sake - e.g. benevolence, identity, norms of distributive fairness and procedural fairness (Osterloh et al. 2002). Intrinsic motivation is fostered by commitment to the work itself, which must be both satisfactory and fulfilling to the employees (Dermer, 1975). Therefore, the ideal incentive system is in the meaningful work content itself (Osterloh and Frey, 2000).

1.2 - Self-Determination Theory

Given the importance of intrinsic motivation, it is instructive to inquire into the conditions that enhance, sustain or subdue and forestall the innate propensity that catalyzes it? Theorists have developed what is called Self-Determination Theory (SDT) whose aim it is to specify the factors that impact intrinsically motivated individuals (Ryan and Deci 2000). Crucial among these are the needs for competence, relatedness and autonomy. The theory argues that social, contextual events that conduce to feelings of competence during action could enhance intrinsic motivation for that action. Positive performance feedback, for instance, was found to enhance intrinsic motivation, whereas negative feedback diminished it (Deci 1975), this is so, because these are mediated by perceived competence. Ryan (1982) added that feelings of competence will enhance intrinsically-driven behavior only if it is accompanied by a sense of autonomy, that is, by an internal perceived locus of causality (deCharms 1968).

People must not only experience feelings of competence, but also feel that their behavior is self-determined and caused by themselves, and not dictated from the outside. In this connection, some researchers hold that extrinsic rewards can undermine intrinsic motivation (Deci 1975), because people feel that external influence is a loss of internal control and reduction in autonomy. Additionally, research

has shown that, not tangible rewards, but also threats, deadlines, directives, pressured evaluation and imposed goals diminish intrinsic motivation, because they conduce toward an external perceived locus of causality. In contrast, choice, acknowledgement of feelings, and opportunities for self-direction were found to enhance intrinsic behavior, confirming a sense of autonomy (Ryan and Deci 2000). On this count, an autonomy-supporting context, in contrast to a controlling one, would incite more intrinsic motivation, encourage initiative, curiosity and a desire for challenge.

Over and above autonomy and competence, intrinsic motivation requires a third factor, namely feelings of social relatedness. Satisfaction of the need for relatedness implies an increase in intrinsic motivation (Ryan and Deci 2000). Because our sense of self grows in part out of our relations and interactions with others, a sense of belonging and a feeling of being part of, and being appreciated by other members of the community supports intrinsically motivated action. People will be more inclined to pursue a certain behavior if they feel that their social network supports and approves of that. For example, students are more intrinsically motivated when they experience their teachers as warm and caring. In the main, a secure, relational base appears to provide some support for intrinsic motivation.

In sum, most research in this area tends to suggest that the social context plays an important role in facilitating or thwarting intrinsic motivation. Of crucial importance is that the activity itself is inherently interesting for its own sake, involving creativity and exploration. As noted above, intrinsic motivation is vulnerable to the negative impact of external control and management through the use of extrinsic rewards. Research suggests that if people are rewarded for doing what they would have done out of interest in the activity itself, they will be less likely to do it in the future without being paid (Deci and al. 1999). This transfer of power or control from the inside of the individual to the person who externally controls and manages the rewards will diminish intrinsic motivation. When extrinsic rewards are used to harness for some purpose that is not in harmony with the interest of the individual, it will be experienced as a loss of control and of freedom and a sense of not being the origins of their action. In contradistinction, when the rewards are intrinsic to the person's self understanding and purposeful fulfillment, when perceived as providing supporting evidence relevant to the individual's sense of freedom, of responsibility and of competence, such rewards would have positive effects.

1.3 - Crowding-out Effects

As noted above, research suggests that introducing extrinsic rewards on an intrinsically-driven activity will reduce motivation such that the total effort towards a task diminishes (Osterloh and Frey 2000). What this implies is that if rewards come to be perceived as the cause of engaging in activity, the motivation to continue that activity will be reduced because it will be viewed less interesting and enjoyable in its own right. Crowding-out effects have gained wide recognition in the literature (Osterloh and Frey 2000). Amabile (1996) asserts that rewards have a diminishing effect primarily on creative tasks, including higher-level problem-solving: the more complex the activity, the more it's hurt by extrinsic reward. For her, creative work cannot be controlled, or operantly conditioned for the following reasons. First, extrinsic rewards encourage people to concentrate on a task, to execute it as fast and as efficiently as possible, thereby avoiding exploration and risk-taking. Second, people will feel that rewards are exercising control over their activities and reducing their autonomy. Such sense of loss of control and autonomy will affect performance and creativity negatively (Amabile and al. 1994). Finally, rewards can crowd out intrinsic motivation by converting it into less desirable, and thus they may not do them with great interest. As emphasized by Ryan (1975), there's a difference between saying, "I'm giving you this reward because I recognize the value of your work" and "You're getting this reward because you've lived up to my standards."

Hence, based on the above, and assuming that the transfer of tacit knowledge requires intrinsic motivation, the aim of remainder of the paper is to explore possible crowding or amplifying effects on individuals' intrinsic motivation to share their tacit knowledge with their colleagues. In particular, when extrinsic rewards are perceived as leading to a sense of loss of self-determination and impaired self-esteem intrinsic motivation is likely to subside. This is so because extrinsic motives shift the locus of control from the inside to the outside, leading to a reduction in self-determination (Rotter 1966). Furthermore, if the introduction of extrinsic reward is perceived as undermining a person's sense of competence and social esteem, it will negatively affect their intrinsic motivation to share their tacit knowledge. By contrast, external rewards will be likely to re-enforce intrinsic motivation if individuals perceive them as supportive and self-esteem-fostering. Rewards, in the form of recognition and meaningful, positive feedback for work, increase intrinsic motivation. Likewise, team-based structures that foster emotional ties and bonds of loyalties that enable personal relationships, that are most likely strongly raise the intrinsic motivation to cooperate (Dawes et al. 1988, Frey and Bohnet 1995) and to be willing to share tacit knowledge with others.

In sum, external intervention may have a positive or negative effect on the efforts individuals put into their work, depending on whether it is dominantly perceived as a promoter of one's sense of self-control and self-esteem or whether it is perceived as constraining, degrading and controlling. Hence our general framework, captured in the following representation (figure 1) would suggest that, in the context of transfer of tacit knowledge, which is a largely intrinsic activity, rewards that are perceived as promoting self-control, self-esteem and the social status of individuals would have a positive, re-enforcing effect on the motivation to transfer knowledge. By contrast, if rewards are perceived as impairing these conditions, would have discouraging and inhibiting effects on individuals' motivation to transfer their knowledge. This framework will guide us in generating and interpreting the information underlying this study.

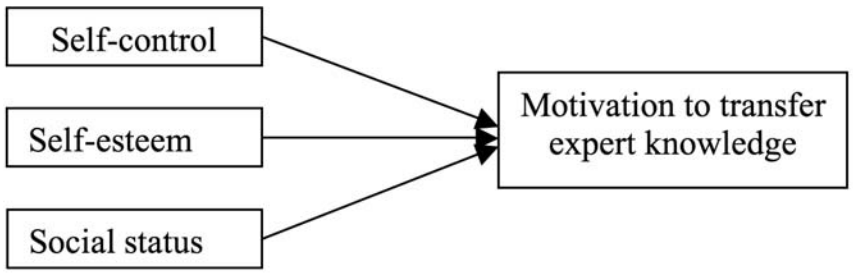

Figure 1: The effects of rewards on the motivation to transfer tacit knowledge

2. TOTAL: a Case Illustration

As noted above, our empirical site involves an international oil company, Total. Total is one of the four largest international oil companies in the world. With operations in more than 130 countries, Total engages in all aspects of the petroleum industry, including upstream operations (oil and gas exploration, development and production, LNG) and downstream operations (refining, marketing and the trading and shipping of crude oil and petroleum products). The company also produces base chemicals (petrochemicals and fertilizers) and specialty chemicals for the industrial and consumer markets (rubber processing, adhesives, resins and electroplating). In addition, Total has interests in the coal mining and power generation sectors. Total is also developing renewable energies, such as wind and solar power, and alternative fuels. The company employs 96 000 persons; in 2006, its sales were about 154 billion Euros. The net income was about 12.6 billion Euros for the same year; the larger part of it coming from its "upstream" activities (Total's Annual Report 2006).

> Total is determined to remain one of the most dynamic and successful energy groups in the world today by:
>
> - paying careful attention to safety and environmental protection, which allows the group to improve the reliability of its industrial plant;
> - pursuing an investment strategy aimed at profitable organic growth;
> - leveraging its top expertise in high-growth sectors (deep offshore, LNG, heavy crudes) and implementing a wide-ranging human-resources program with the emphasis on diversity at the managerial level; and by
> - maintaining good geographic diversity of both reserves and production.

One of its competitive advantages is based on the ability of its geoscientists to discover new reserves in an as efficient a manner as possible (Total's Annual Report 2006). Consequently, the company is very dependent on the long-term knowledge of its experts; i.e. the tacit knowledge that is internally accumulated through years of experience. The life-cycle of this expertise is fairly stable over a time span of about 30 years or even more. Within the strategic geosciences division, which is the focus of the present study, geophysicists, geologists and reservoir engineers are regarded as knowledge workers. Their work can be described in the following way: geophysicists derive and interpret underground images of the earth. Geologists analyze how the petroleum system is formed and the features of the underground. The reservoir-engineers determine the dynamic properties of the reservoir to predict the behavior of a petroleum field, to optimize its production and to compute its reserves. If the wrong decision is made - such as drilling a well that proves unproductive, the company will incur a very large financial loss, about several millions of Euros.

Given the high stakes involved in these decisions- decisions which can only be processed by expertise knowledge - transfer of expertise is a central issue in the company. Add to this the fact that most of those scientific experts will retire in the next 10 years. Their average age is more than 50 years old. In this field, expertise is acquired through decades of practice and learning by doing. In the face of imminent mass retirement, the company is undertaking a number of training programs. Training does not solve the whole problem given that the knowledge in question is tacit and thus cannot be easily transferred to newly recruited novices overnight. To these ends, the geo-science division organizes work process in project-based and cross-age teams. In their concerted efforts to transfer knowledge to newly-minted engineers and scientists, experts are taking a leading role, assuming the responsibility for teaching, transferring their expertise and guiding the novices. Our focus is thus on a number of experts in order to determine what encourages or dis-encourages them to share their knowledge with newly-recruited engineers. We assume that the experts under consideration possess a form of knowledge that is tacit and driven by intrinsic motivation.

The empirical material is derived from 28 geo-scientist interviewees at Total's geosciences division. We interviewed:

-5 experts with more than 20 years of experience;

-2 experts retired but still working on research projects;

-8 specialists (senior researchers, future experts, more than 15 years of experience);

-3 technical managers who yet managed technical teams with experts and specialists

-10 less experienced scientists with less than 5 years of experiences, who interact with experts

The informants have 12 different nationalities: Algerian, Angolan, Dutch, English, French, Iranian, Israeli, Lebanese, Nigerian, Norwegian, Russian, and Spanish. Interview time ranged from 50 to 120 minutes: average time was 70 minutes. During these interactions with the experts and specialists, we focused our questions on how they define and relate to their field of expertise and what motivates them to readily and willingly transfer their expertise to younger generation. In order to know about the perspective of the novices and technical managers, our questions to them revolved a round how they experience the transfer process and what, according to them, are the requirements for an effective knowledge transfer process. In order to protect the identity of the interviewees, we have used fictive names.

Findings

a. The intrinsic nature of expertise

The majority of the experts we have interviewed consider their work as inherently interesting. They emphasize the importance of curiosity and exploration as prerequisites for conducting their work. Nicholas summarizes his understanding of what he is doing: "when you are an expert, in general it is because you invested many years into your work, and for that you need to be curious about learning new things the whole time. To become an expert, curiosity is necessary because it enables you to develop an inquisitive and investigative mind". Adam, one of the informants we interviewed, describes why he has chosen to be a geo-scientist in the following way: "I had this 'sweet' taste for discovery and exploration: I always wanted to find the solutions that others couldn't... I just like discovering". One is motivated in his work as long as they are curious about his field of study. Frederich supports this view point: "The willingness to exert one's self and the curiosity towards one's job are hallmarks of what it is to be an expert. Of course, we should not forget social recognition and job satisfaction as additional factors that determined my decision". Vincent defines job satisfaction in terms of satisfaction of his curiosity to learn: "I take pleasure in learning in order to satisfy my curiosity; if I cannot satisfy my curiosity to learn I am not happy with doing my job". John who is a geo-statistician, what motivates him in his work is "curiosity to learn, to understand and to think, it is an inner drive that sustains [his] thirst for learning". Kadder adds that "when I was a kid I liked to disassemble my toys because I was so curious about what was inside them and how they worked ...this is why I became an engineer". John said more or less the same thing: "my wife is often complaining that I cannot take my mind off my work, because even when we are on holiday, such as when climbing a mountain, I would stop and examine the stones and rocks that look unfamiliar to me". For Patrick, a former geophysical research manager, "the ideal expert has more than curiosity. S/he also needs to be passionate about the objects of their knowledge. Because this form of knowledge can only be learned by doing and through many years, one needs curiosity, enthusiasm, passion and humor". John adds that: "a true geologist has a certain relation, a physical, but not a sexual one, with stones: s/he would touch it, feel it, smell it and sometimes lick it in order to come closer to it".

In this sense, expertise is related to the emotional commitment experts invest in their

knowledge. They regard enthusiasm, curiosity and passion as intrinsically-bounded elements of expertise. Becoming an expert is a lengthy and painstaking process which cannot be achieved without emotional investment and passion.

b. The relational nature of expertise

Experts have emphasized the nature of expertise, involving an intellectual challenge, requiring an insatiable thirst to learn and explore new things. In this section, we discuss the social context in which expertise thrives and develops. Next to the experts' predisposition to learn, the interviewees have underscored the relational nature of expertise, that is, it arises in social interactions with others.

David, a specialist in uncertainty assessment says that: "to have expertise without transferring is nonsense." Because expertise is an attribute bestowed on experts by others, it can only manifest itself in social acts of sharing and helping others solve concrete problems. Absence sharing and transferring, an expert would not be recognized as such. When we asked the experts about what motivates them to share their expertise and how they view the process of transferring it to younger colleagues we were surprised to learn that for them transferring knowledge is a process of learning, or rather, of co-learning. For Nicholas there is no distinction between knowledge transfer and knowledge creation: "You are neither a source only, nor just a recipient, you are both, that is, a knowledge creator. (...) it is when you are discussing with somebody you are engaged more in a process of knowledge creation process rather just a transfer process. (...) You cannot learn alone, it makes no sense". Maria seconds that view: "when you work with someone else, you always learn something". John maintains that "when learners are curious and highly motivated we feel that our efforts are meaningful and thus more motivated to work with them".

For this reason, they claim that knowledge transfer requires curiosity, enthusiasm and passion. As contended by Maria, curiosity has an impact not only on the acquisition of knowledge, but also on the transfer of knowledge: "What motivates an expert is this permanent thirst for learning, this willingness to learn but also to share". Adam adds that "If you are curious, if you like to discover, to share your knowledge with others you are able to discover new knowledge". Even though experts see transferring knowledge as part of their daily tasks, they do recognize that oftentimes, the novices beat experts as learners.

In other words, they described the process of transferring their knowledge along the same lines as the processes that defines learning it. As noted by Björn, being an expert is "a constant search for the unknown, you are compelled to learn new things every day, and since being an expert means that you spend most of your time transferring and sharing your knowledge, you are bound to learn from these exchanges with others". The practice of being an expert implies working with others and it is from those practices that experts learn. To the extent that experts learn while interacting and exchanging ideas with others, their passion to learn is their passion to transfer their knowledge. For Justin you need "enthusiasm to touch people where they are sensitive, (...) it goes beyond rationality, it is not related to cognition, although it deals sometimes with how certain technologies function".

Over and above the social, interactional elements that define expertise, our informants see their expertise as an institution in its own right: Adam says that: "I consider transferring knowledge fundamental to the sustenance and survival of our profession; and if we fail to transfer it, we would not measure up to our profession". In a sense, by transferring their knowledge, experts believe that they expand the field of their expertise to other individuals. Adam says that "the fact that I am not going to stay here for long, I want to see my expertise transferred back to my company, I do not want it to go to waste. Since it is my profession and since it will not have much utility for me personally, I want to pass it over to other generations who can sustain it and develop it. The worst scenario I can imagine is to see it wither away

after I leave this company". Nicholas elaborates on this point by saying that "when you are passionate about something you want to make it concrete and objectified, so it will lead its own life". When a departing "expert transfers his knowledge, he takes part, at a small or big scale, in the process of enriching his field of expertise" (Nicholas). Frederich sees this feeling of contributing to this stock of collective knowledge as "a very strong motivator for transferring his knowledge" (Frederich). Adam even goes as far as saying that "because my expertise doesn't belong to me, I am lucky to posses it, because I have stayed in this profession long enough, gathering everything I wanted to know during these years, and therefore it would be a big loss if this know-how would disappear after I retire".

What seems to emerge from our interviews with these departing experts regarding knowledge transfer is that they are intent on seeing their expertise live on, and be transmitted to younger generations. The majority of the interviewees emphasized the significance of emotional drives (curiosity, enthusiasm and passion) as motives for learning/transferring their expertise. Of course enthusiasm and passion are not enough, expertise has to be proven, has to be manifested through practicing it, through sharing it with others. Sharing is the social aspect of expertise.

c. The effects of rewards

The extent that the experts under consideration do not make a distinction between transferring their knowledge and learning, we were interested in finding out what actually would motivate them to be more willing to share or transfer their knowledge. More specifically, we asked questions about whether extrinsic rewards such as bonuses, social recognition, etc, would play any role in the degree to which they are motivated. Would such extrinsic rewards enhance or reduce their motivation to share more of their expertise? If extrinsic rewards do not have a strong effect, what would?

As to straightforward question whether monetary rewards would enhance their motivation to transfer knowledge, predominantly the answer was in the negative: “Money is no longer a reward for expertise. If it were the case, we [experts] would not be where we are today” (Adam). To that, Nicholas adds: “monetary rewards will neither affect experts’ curiosity to explore and search, nor their interest in their jobs”. Frederich expressed himself in more explicit terms: “money is a reward for job performance but not for transferring one’s expertise”. It seems as though “…most experts would still transfer their knowledge if they were not paid for it, just because they are interested in their field of competence. When one expert transfers his/her knowledge, s/he is enriching his/her field of expertise, in however small way that may be” (Nicholas).

However, factors, such as the feeling of being competent, emerged as pertinent. Frederich says: “I personally feel that public acknowledgement is a strong motivator for an expert like me … I get a feeling of satisfaction when I see that people recognize my competence in these conferences and that they really listen to what I got to say”. Another relevant statement was made by Frederich: “Doing your job well is a fundamental issue of expertise. This links expertise to the feeling of belonging to a community I previously talked about. The feeling of contributing to the collective performance is a very strong motivator to transfer knowledge” (Frederich). Finally, John who is about to retire, put it in the following way: “The fact that I am not going to stay here for long, and because I am not going to need my expertise anymore, it is crucial for me to see it live on after I am gone. It is of no use to me anymore, and for this reason I ‘give’ it [back] to my company … It is because my expertise doesn’t belong to me, I have been lucky to have had the opportunity to acquire it, through years of experience at this company, and it would be painful to see it disappear after I retire”.

3. Analysis and discussion

It is conventionally held that rewards can be used to increase the motivation to transfer knowledge (Alavi and Leidner 2001; Argote and Ingram 2000; Argote and al. 2003; Cabrera and Cabrera 2002; Cabrera and al. 2006; Goodman and Darr 1998, Hansen 1999, Hansen and al. 2006; Quigley and al. 2007; Menon and Pfeffer 2003; Spender and Grant 1996, Szulanski 2000). Our case provides a more complicated picture. Focusing on the transfer of tacit knowledge which is assumed to be sustained by intrinsic motivation, our findings tend to lend support to Osterloh and Frey's (2000) assumption that, for intrinsically motivated individuals, the ideal incentive system would be a meaningful content of work. Our study suggests that these intrinsically—driven experts work mainly in various projects, where sharing knowledge is the basis of teamwork. For instance, when an operational geoscientist has a measurement problem that he discusses with an expert, this problem is one further case for the expert to think about, and which enriches his experience. Transferring and sharing expertise knowledge is part of the whole team's daily job.

As noted above, the dominant logic underlying intrinsic motivation is not mainly understood in terms of exchange value; in terms of 'prices' for their work, but it is more in terms of 'prizes'. Their attachment to their field of expertise is not only a means to an end, but it is largely an end in itself - performing an activity for the pleasure of performing it. This is so because, in our case, extrinsic rewards, such as money and promotion, were found adequate only by few of the interviewed experts. In terms of social rewards, peer recognition of an individual expert's expertise are regarded significant, as these raise his/her perceived status inside the organizations, and thus, affords him/her more power and credibility (Thomas-hunt et al., 2003). One of the experts highlighted the importance of the recognition by their peers as more important than the public acknowledgement in front of other employees. Peers validate experts' knowledge and afford them a status as experts. This social recognition is seen highly motivating by most of the experts under consideration. It was argued that the function of monetary rewards is merely a sign of recognition of experts' intellectual superiority. In this case, money is itself not a means to end but it a symbol of high esteem. It stands for (or is a symbol of) something that is more coveted by the experts – namely that their bosses hold them in high esteem. This is what explains why monetary rewards were seen the least motivating factor in the sharing of expertise. Experts are expecting recognition for their status as experts, not for sharing their expertise. When experts talked about monetary rewards, they emphasized the importance of linking them to organizational goals but not to individuals. For them, rewards should be paid only if these overall goals are achieved. Informants warned against the risk that rewards may imply: changing a culture of sharing into a culture of bargaining and exchange.

From our study it seems to emerge that the distinction between knowledge and knowledge creation is hard to maintain in the context of expertise knowledge. This is so because expertise is relational, that is, its acquisition requires interactions with others. Interactions are occasions for experts to confront new problems and new challenges, which enable them to augment their stock of expertise. Furthermore, expertise presupposes a strong attachment to an object, or a field of interest, of the curiosity of the individuals. Hence, experts are first and foremost so passionate about their object of interest (object-orientation) that they are readily disposed to share it and leverage it in interactions with others (relational): "you are so excited and passionate about your area of expertise that you cannot help talking about with others!" (Björn). This strong connection between object-orientation and the social, relational aspects of expertise does not necessarily require extrinsic rewards.

Of course, as we have noted, extrinsic rewards can modify positively or negatively this relation. When extrinsic rewards amplify the sense of competence, control and relational esteem, experts' motivation to share their knowledge is augmented. The reason is that such extrinsic rewards are not perceived by the experts as attempts to displace the locus of control from their inside to some other external agents. In this context, even monetary rewards would be themselves as a symbol of acknowledgement of their competence, and their social status – which are two factors that reinforce and encourage them to share their expertise and thus learn more.

Concluding remarks

Theorists in knowledge management usually make a distinction between knowledge acquisition and knowledge transfer. Our study tends to suggest that expertise is relational, that is, it is a quality that is bestowed upon someone who discloses it during problem-solving interactions with peers. Such interactions are not only opportunities for experts to transfer their knowledge, but also occasions for them to learn from each particular case they encounter. It is nurtured and fostered in dialogues among groups, rather than within the innermost of an individual. Assuming its dialogical nature, expertise is a form of knowledge where learning and transferring knowledge are two processes that co-configure each other.

Expertise knowledge transfer highlights an instance of transfer where knowledge processes (transfer and acquisition), the transmitter (expert) and the recipient are intrinsically-bounded up. To the extent that here is a strong link between expertise knowledge, the source and the target, the role of external incentives and motivation recede to the background. These play an active role only if they undermine this triadic relationship, that is, if the expert's competence in the object of their passion is questioned, if the status of the source is not recognized by its peers as an expert, and if the expert does not feel as a sovereign source of their transfer initiatives. However, extrinsic rewards will have a positive effect if they contribute to the cohesion of these three aspects of expertise knowledge transfer: (passion-driven) knowledge, (self-regulated) sources and (encouraging and supporting) recipients.

Further empirical research may investigate more explicitly the role of extrinsic motivation on this triadic relationship of this expertise knowledge transfer process. This study has put forward passion and curiosity as a strong predictor of an intrinsically-driven knowledge-sharing behavior. It has also thrown some light on the interplay among passion, feeling of competence, of self-determination, and of self-esteem.

Further empirical research may investigate more explicitly the role of extrinsic motivation on this triadic relationship of this expertise knowledge transfer process. This study has put forward passion and curiosity as a strong predictor of an intrinsically-driven knowledge-sharing behavior. It has also thrown some light on the interplay among passion, feeling of competence, of self-determination, and of self-esteem.

References

Agnew, N. M., Ford, K. M. and Hayes J. P. (1997) '*Expertise in context: Personally constructed, socially selected, and reality-relevant?*' In P. J. Feltovich, K. M. Ford, and R R Hoffman (Eds.), Expertise in Context, pp 219-244. Menlo Park, CA:AAAI Press.

Alavi, M. and Leander, D. (2001), "Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues", *MIS Quarterly*, Vol. 25, n° 1, p. 107-176, March

Amabile, T.M., Hill, K., Hennessey, B.A., & Tighe, E. (1994), The Work Preference Inventory: Assessing intrinsic and extrinsic motivational orientations, *Journal of Personality and Social Psychology*, 66, 950-967.

Amabile T. M. (1996), C*reativity in Context*. Boulder, Colorado: Westview Press.

Argote L., Ingram P. (2000). Knowledge transfer: A basis for competitive advantage of firms. Or*gan. Behavior Human Decision Processes*, 82, 150-169.

Argote L., McEvily B. and Reagans R. (2003), Managing Knowledge in Organizations: An Integrative Framework and Review of Emerging Themes. Man*agement Science*, Vol. 49, No. 4, pp. 571-582.

Brown J. S., and Duguid P. (1991), Organizational learning and communities of practice, *Organization Science*, 2 40-57.

Burgess D. (2005). What Motivates Employees to Transfer Knowledge Outside Their Work Unit? *Journal of Business Communication*, Vol. 42, No. 4, pp. 324-348.

Cabrera, A. and Cabrera, E.F. (2002), Knowledge Sharing Dilemmas, *Organization Studies*, 23(5): 687–710.

Cabrera A., Collins W. C. and Salgado J. F. (2006), Determinants of individual engagement in knowledge sharing, *International Journal of Human Resource Management*, 17:2, February 245–264

Chesbrough H. W. Teece D. (1996), "When is Virtual Virtuous: Organizing for Innovation", with, Ha*rvard Business Review*, Jan.-Feb. 65-73

de Charms, R. (1968), *Personal causation*, New York: Academic Press

Deci, E.L. (1975). *Intrinsic Motivation*, Plenum Press, New York, US.

Deci E. L., Koestner R., & Ryan R. M. (1999), A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation, *Psychological Bulletin*, 125, 627-668.

Dermer J. (1975), The Interrelationship of Intrinsic and Extrinsic Motivation, *Academy of Management Journal*, Vol. 18, No. 1, pp.125-129.

Ericsson K. A., and Lehmann A. C. (1996), Expert and exceptional performance: evidence on maximal adaptations on task constraints. *Annual Review of Psychology*, 47: 273-305

Foley, M. and Hart, A. (1992), Expert Novice Differences and Knowledge Elicitation, In R. R. Hoffman (Ed.), Th*e Psychology of Expertise: Cognitive Research and Empirical* AI: 233-269. Mahwah NJ: Springer-Verlag.

Gagné M., & Deci E. L. (2005), Self-determination theory and work motivation, Journal of Organizational Behavior, 26, 331-362

Goodman, P. S., & Darr, E. D. (1998), "Computer-aided systems and communities: mechanisms for organizational learning in distributed environments", *MIS Quarterly*, 22(4), December, 417- 441

Greeno, J. S., and Simon, H. A. (1988) Problem solving and reasoning, In R. C. Atkinson (Ed.), *Steven's handbook of experimental psychology*, 2nd ed.: 589-639. New York: Wiley

Haas M. and Hansen M.T. (2007), "Different Knowledge, Different Benefits: Toward a Productivity Perspective on Knowledge Sharing in Organizations.", *Strategic Management Journal* 28, 1133-1153

Hansen, M.T. (1999), The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits, *Administrative Science Quarterly*, Vol. 44, No. 1, pp. 82-111.

Hansen M. T., L. Mors, and B. Lovas (2005), "Knowledge Sharing in Organizations: Multiple Networks, Multiple Phases." *Academy of Management Journal*, Oct, Vol. 48 Issue 5, p776-793

Hinds P. J.., Patterson M., and Pfeffer J. (2001) "Bothered by Abstraction: The Effect of Expertise on Knowledge Transfer and Subsequent Novice Performance", *Journal of Applied Psychology*, 2001, Vol. 86, No. 6, 1232-1243

Kogut B., Zander U. (1992). Knowledge of the firm combinative capabilities, and the replication of technology, *Organization Science*, 3(3): 383–397.

March J. J. (1991), Exploration and exploitation in organizational learning. *Organization Science*, 2:71–87.

Menon, T. and Pfeffer, J. (2003), Valuing Internal vs. External knowledge: Explaining the Preference for Outsiders. *Management Science*, Vol.49, No.4, pp. 497-513.

Mieg H. A. (2001), *The Social Psychology of Expertise: Case Studies in Research, Professional Domains, and Expert Roles*, ed Lawrence Earlbaum Associates

Osterloh, M. (2007), "Human Resources Management and Knowledge Creation", in: Ichijo, Kazuo, and Ikujiro Nonaka, eds. *Knowledge Creation and Management : New Challenges for Managers*. Oxford University Press,

Osterloh, M. and Frey, B. (2000), Motivation, Knowledge Transfer, and Organizational Forms, *Organization Science*, Vol. 11, No. 5, pp. 538-550.

Osterloh M., Frost J. and Frey B. (2002). The Dynamics of Motivation in Organizational Forms, *International Journal of the Economics of Business*, Vol. 9, No. 1, pp. 61-77.

Pan SL and Scarbrough H (1999), Knowledge Management in Practice: An Exploratory Case Study of Buckman Labs, *Technology Analysis and Strategic Management* 11(3) 359-74

Polanyi, M. (1966), *The tacit dimension*. New York: Doubleday.

Quigley N: R., P: E. Tesluk, E: A. Locke, K: M. Bartol (2007), A Multilevel Investigation of the Motivational Mechanisms Underlying Knowledge Sharing and Performance, *Organization Science*, Volume 18 , Issue 1 (January), Pages 71-88

Rotter J. B. (1966), Generalized expectancies for internal versus external control of reinforcement, *Psychological Monographs*, Vol.80, no.1, pp.609-633

Ryan, R. M. (1982), Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory, *Journal of Personality and Social Psychology*, 43, 450-461

Ryan, R. M. (1998), Human psychological needs and the issues of volition, control and regulatory focus. In J. Heckhausen & C. Dweck (Eds.), *Motivation and self-regulation across the life span*, New York: Cambridge University Press.

Ryan, R. M. and Deci, E. L. (2000), Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25, 54-67.

Smith, J. A. 1991. *The idea brokers. Think tanks and the rise of the new policy elite*, New York: Free press.

Spender, J-C. and Grant, R.M. (1996) 'Knowledge and the Firm: Overview', *Strategic Management Journal*, 17: 5–9.

Sternberg, R. J. and Horvath, J. A. (1999). *Tacit Knowledge in Professional Practice researcher and practitioner perspectives*, ed Lawrence Earlbaum Associates

Szulanski, G. (1996). Exploring Internal Stickiness: Impediments to the Transfer of Best Practice Within the Firm. *Strategic Management Journal*, Vol. 17, Winter 96 special issue, pp. 27-43.\

Tanaka J. W., Curran Tim, and Sheinberg David L. (2005), "The Training and Transfer of Real-World Perceptual Expertise", *Psychological Science*, Vol. 16-Number 2, p. 145-151

Thomas-Hunt M., Ogden T. and Neale M. (2003), Who's really sharing? Effects of social and expert status on knowledge exchange within groups, *Management Science*, 49 (4), 464-477.

A continuous approximation procedure for determining inventory distribution schemas within supply chains

By Nikhil A. Pujari
Ohio University, Athens, USA

Trevor S. Hale
University of Houston - Downtown, Texas, USA

Faizul Huq
Groupe ESC Pau, France

Abstract

This paper presents an integrated inventory distribution optimization model that simultaneously incorporates the issues of location, production, inventory, and transportation within a supply chain. The objective is to determine the optimal number and size of shipments under varying but commonly practiced production and shipping scenarios. A continuous approximation procedure is proposed to determine the optimal number and size of shipments. Three production and shipping scenarios are investigated and closed form expressions for the optimal number of shipments for each scenario are obtained. A numerical example is presented to demonstrate the usefulness of the model.

Keywords:

Distribution system; Inventory management; Supply chain; Continuous approximation

Introduction

Supply chain management (SCM) is the systematic analysis and educated decision-making within the different business functions of an organization resulting in smooth and cost-effective flows of resources - material, information, and money. In other words, it is the coordination and synchronization of the flow of resources within a network of suppliers, manufacturing facilities, distribution centers and customers.

The supply chain of an organization consists of different functions at each planning stage. According to Ganeshan and Harrison [10], these functions can be broadly classified into four categories - location, production, inventory, and transportation. Each of these functions plays a major role in the overall performance of the supply chain. So it becomes essential to execute each one of them in an optimal manner to ensure an efficient supply chain performance.

These functions have been mostly studied and optimized individually. However, in the recent years, there have been attempts to consider and analyze these functions collectively to study their interactions and provide a better solution for overall optimization. The goal of this paper is to develop an integrated supply chain optimization model that brings together the four functional categories of location, production, inventory, and transportation. There are numerous benefits of integration that can help companies improve their performance by way of collaborative planning and execution. According to the National Research Council, one of the most visible and obvious areas for improvement for manufacturing companies is in the area of inventory control. With the *combined* optimization approach presented herein, inventory can be controlled at more globally optimum levels across the entire supply chain leading to lower holding costs, reduction of required warehouse space, reduced material handling activities, and timely deliveries. In the next section, we present an extensive literature review followed by a section on model development.

Subsequent to the model development section, we present three general cases of production and shipping control patterns. A continuous approximation approach is then used to determine optimal number of shipments under multiple scenarios arising from common production and shipping patterns. In the process of approximating the optimal number of shipments, tactically optimal inventory levels are also simultaneously derived. To demonstrate the efficacy of this integrated approach, a hypothetical case is presented and solved for one of the more complex scenarios. The final section of the paper presents the conclusions followed by a brief discussion on future research and extensions.

1. Background

One of the first steps to run an effective supply chain is the strategic positioning of the manufacturing facilities, warehouses and distribution centers. A number of quantitative models use mixed-integer programming (MIP) to solve such supply chain optimization problems. Others have looked to continuous approximations to study such systems. The landscape of these two research areas are reviewed below.

1.1 - Mathematical programming models

Canel and Khumawala [2], in addition to providing a thorough review of the literature for the uncapacitated, multi-period international facilities location problem (IFLP), they formulated a mixed-integer programming (MIP) model and solve it using a branch and bound technique. The decision variables include the countries to locate manufacturing facilities, and their production and shipping levels. Canel and Khumawala [3] later developed heuristic procedures for solving the MIP model of a similar IFLP.

One of the first attempts of using MIP was done by Geoffrion and Graves [11], where a MIP model was formulated for the multicommodity location problem. This seminal research involved the determination of distribution center (DC) locations, their capacities, customer zones and transportation flow patterns for all commodities. More recently, Geoffrion and Powers [12] provided a comprehensive review of the development of distribution systems research.

Cohen and Lee [5] developed an analytical model to establish a material requirements policy based on stochastic demand. They develop four different sub-models with a minimum-cost objective. A mathematical algorithm at the end decides the optimal ordering policies to minimize the costs. Cohen and Lee [6] also developed a deterministic analytical MIP model to maximize the global after-tax profits through optimal policies for facility network design and material flows. The decision variables for the network design issues include location and capacities of all production facilities whereas those for material management issues include sourcing decisions, production and distribution planning. The model decides the optimal resource deployment for a particular policy option.

Ganeshan [9] presents an integrated model that synchronizes the inventory at retailers and warehouse, and the demand at the warehouse. One of the important contributions of this paper is the development of an inventory-logistics framework that includes inventory and transportation components in the total cost function. It also identifies the need for a model that caters to more than two echelons in the supply chain and a better way to model random delay at the warehouse.

Chandra and Fisher [4] compare the computational aspects of solving the production and distribution problems separately and in a combined model. A number of cases with different values for the various parameters for these models are analyzed to compare the performances. Final analysis shows an improvement of up to 20% in some cases for the integrated approach. The authors propose that a segregated approach is suitable if there is a sufficient reserve of finished products between the production and shipping phases. On the other hand, an integrated approach is suitable for cases where the distribution costs are significantly higher and the related networks are dense.

Lodree et al. [16] considered the integration of customer waiting time with the production–distribution functions in a supply chain. Optimization policies for models with combinations of production, inventory and transportation costs, along with customer waiting times were proposed to determine the produc-

tion rate and the sequence of vehicle shipments. Benjamin [1] analyzed the inventory and transportation cost minimization problem addressing the decisions such as production lot size at each facility, total amount shipped from each facility and the amount shipped every few weeks after receiving the order. This problem is treated as a combination of the economic production lot size problem, the transportation problem and the economic order quantity problem. Three basic costs have been identified – production cost, transportation cost and inventory cost. The author concludes that integrated formulation can be applied to a wide variety of problems such as multi-stage trans-shipment problems with setup costs. Also, setup and fixed transportation costs can be analyzed simultaneously.

1.2 - Continuous approximation models

Masel and Pujari [17] considered the integration of customer waiting costs in integrated supply chain analysis using a continuous approximation approach. Expressions for customer waiting cost, inventory cost, and the transportation cost were obtained in three separate sub-models. These sub-models are then analyzed collectively in a single comprehensive minimization model for simultaneous optimization. Classical optimization was then used to solve this model to obtain a closed form solution to determine the optimal number of shipments that minimizes the total costs. This is one of the few attempts to include qualitative aspects of the supply chain, such as customer satisfaction in terms of waiting time reduction, in the overall analysis. Another important aspect of this paper is the use of continuous approximation to model the inventory distribution patterns, which has been adopted in this research as well. This approach helps in obtaining a clear understanding of and accurate modeling of the inventory patterns given their tendency to follow periodic cycles. The optimal value for number of shipments obtained in the end results in a total minimum cost that includes customer waiting cost, inventory cost, and transportation cost.

Dasci and Verter [7] considered the integration of production and distribution functions by proposing an alternative approach based on the use of continuous approximation of costs and demands as opposed to discrete MIP models. Simultaneous, instead of sequential optimization of configuration decisions in production–distribution networks were proposed to avoid sub-optimality. Decision variables were the number and locations of facilities and their service regions. Closed form solutions were obtained to minimize the fixed costs of facility location and operation, and the transportation costs.

Li and O'Brien [14] developed an ingenious dynamic programming model for supply chain integration development between the strategic and tactical levels. Although not a continuous approximation approach, the parallel to the *area of research* herein is worth noting.

Langevin et al. [15] presented an overview of continuous approximation models developed for freight distribution problems. The authors categorized these models into six classes depending on one or more origins, one or more destinations, and with or without transshipments. Their review went on to show that the use of continuous approximation models in conjunction with optimization methods proves that it can be a powerful tool for problem solving. Also, these models provided significant analysis capabilities for decisions on strategic as well as operational levels due to their use of concise summaries of data rather than detailed data points.

The authors point to the relatively sparse research and interaction in the past between mathematical programming and continuous approximation methods but also predicted an increased development of such combined models due to their complementary nature.

In the more general literature, others (e.g., [13,8]) have also advocated the use of continuous approximation models to supplement mathematical programming instead of replacing it. This is due to the advantages of continuous approximation such as ease of interpretation, use of minimal information, and reliance on distribution functions rather than exact data points.

Following this line of reasoning with respect to supply chain modeling, it is postulated that there exists an unexplored avenue of research that incorporates the use of MIP models as input to and/or in conjunction with continuous approximation approaches. This research aims at using continuous approximation models on tactical levels to fine-tune results obtained from previously employed MIP models (e.g., location–allocation models, routing-location models, customer–plant–product assignment models, etc.) on the strategic levels. The goal here is to help develop an integrated model aimed at the cost minimization of the inventory and distribution functions in the supply chain of an organization. A multi-product, multi-period production–distribution system is considered here, with deterministic demand and linear transportation costs.

2. Model development

The supply chain considered here consists of a set of manufacturing facilities with limited production capacities situated within a geographical area. Each of the manufacturing facilities can produce one or all of the products in the company's portfolio. The customer demands for multiple products are to be satisfied from this set of manufacturing facilities. It is assumed here that the location–allocation decisions have already been made keeping in mind the location, cost, and capacity constraints. As a result, the plant–product–customer relationships, too, have already been established thus assigning product quantities to be produced at designated plants and delivered to respective customers. These product quantities, then, act as inputs to the integrated model proposed here.

The customer locations and their distances from each of the existing facilities are known. So, on the strategic level, the transportation cost required to ship the products from the manufacturing facilities to the customers is an important factor. Moreover, sometimes urgent orders need that the products be shipped to the customers at the earliest which means transportation times are also important for customer satisfaction.

Once the strategic level location–allocation decisions have been determined, the continuous approximation model developed herein uses these decisions as an input and fine-tunes them further to provide a detailed distribution plan that optimizes the transportation and inventory costs.

The model optimizes the material flow in the system to determine the optimal distribution plan based on the different inventory distribution patterns considered. These inventory distribution patterns may be unique not only for each customer but also for each of its products. The model determines the number of shipments, and the shipment sizes for each pattern thus providing a customized solution for each company–customer relationship that minimizes the total inventory holding costs, penalty costs, and transportation costs.

The inventory distribution patterns are governed by the following four factors:

a. *Unit production time (y)*: This is the time taken to produce one unit of a product at any manufacturing facility. It is assumed that for a given product, the unit production time does not vary for each facility.

b. *Penalty cost (w)*: It is described as the cost associated with the time the customer has to wait to receive the product after placing the order. It is assumed that the customer expects the delivery of the products as soon as possible. However, since the products are manufactured after an order being placed, it takes some time for the products to be delivered to the customer. The penalty cost is incurred for each product in the order for each time unit until the product is delivered. The penalty cost is negotiated with the customer beforehand at the time of placing the order. It may also be formulated internally in the organization as a tool to measu-

re the losses due to tardy customer deliveries. These losses could be explained as short-term losses such as order cancellations or long-term losses such as losing business from customers due to reduced satisfaction levels. Moreover, this may also result in a deterioration of the company's image in the market. Although some of these may not be purely tangible losses, they still need to be included in the analysis in order to increase the accuracy of the model with respect to the costs incurred.

c. *Inventory holding cost (h):* It is the cost associated with storing the product inventory until it is shipped to the customer. This cost is incurred between when the product is manufactured at the facility, and its delivery to the customer.

d. *Time between shipments (L):* The products may not be shipped to the customer as soon as they are manufactured due to factors such as non-availability of trucks or insufficient workers. There is also some time required for actual loading of the products in the trucks. Sometime the products may first travel to the distribution center (DC) from where they may be finally shipped to the customer. The time taken due to these factors is collectively considered here as the lead time between shipments.

There are other factors also that indirectly affect the inventory distribution pattern such as number of machines available for a given time period, number of trucks available at any instant, shipping rate, travel time, unit travel costs, and shipment size. A list of parameters and decision variables used in the model is given below in Table 1.

The integrated model represents a production/distribution supply chain in which the products are manufactured at the plants and shipped to customers in multiple shipments at regular intervals, until the demand is satisfied. After the commencement of production, the products would be stored in the inventory if shipment is not possible immediately. Otherwise, the products are loaded on the trucks to be shipped to the customer.

Products continue to be produced during the time the shipments are being sent to the customer. As explained earlier, there may be some time required between successive shipments which is considered as the time between shipments. The plant continues manufacturing the product until the total quantity of customer demand is produced. It is possible to store the whole order and ship it at the end of production. However, this option would incur higher inventory cost for storing a large number of products for a long time. It will also incur penalty costs because the customer would have to wait till the end of production to receive the products.

A detailed framework for the integrated model is shown in Fig. 1.

Assumptions:

a. The production quantity (Q) of each product to be produced at each manufacturing facility for each customer has been determined already and acts as an input to the integrated model along with other data.

b. The customer is ready to accept the products as and when the shipment takes place.

c. Penalty costs are incurred for the time the customer has to wait to receive the product.

There can be different cases of inventory distribution patterns based on the difference between production rate (PR) and shipping rate (SR) (i.e., PR = i*(SR) where $i = 1, 2, 3$), continuous or intermittent production and/or shipping, and instantaneous or gradual shipping. Here 'i' is used as a multiplier to denote the ratio between PR and SR. The inventory distribution patterns will in turn influence the inventory costs, penalty costs and transportation costs. Five different cases for continuous production and instantaneous shipping are described below along with the expressions for each of the costs. These cases have been identified to model as many scenarios as possible for continuous production and instantaneous shipping in a generalized framework. Although there can be almost infinite number of cases depending on variations in each of the governing conditions, the cases described here represent the most generic

situations that may occur due to these variations. In order to categorize the problem domain it was necessary to restrict the classification, albeit, in such a way that any of the possible scenarios can be categorized in at least one of these cases.

Table 1
Parameters and decision variables for the integrated model

Parameters	
y	Unit production time
w	Unit penalty cost per unit time
h	Unit inventory holding cost per unit time
L	Time between shipments
Q	Total quantity produced
r	Fixed cost of shipment
t	Travel time per shipment
g	Transportation cost per unit time per shipment
Decision variables	
N	Optimal number of shipments
V	Optimal shipment size

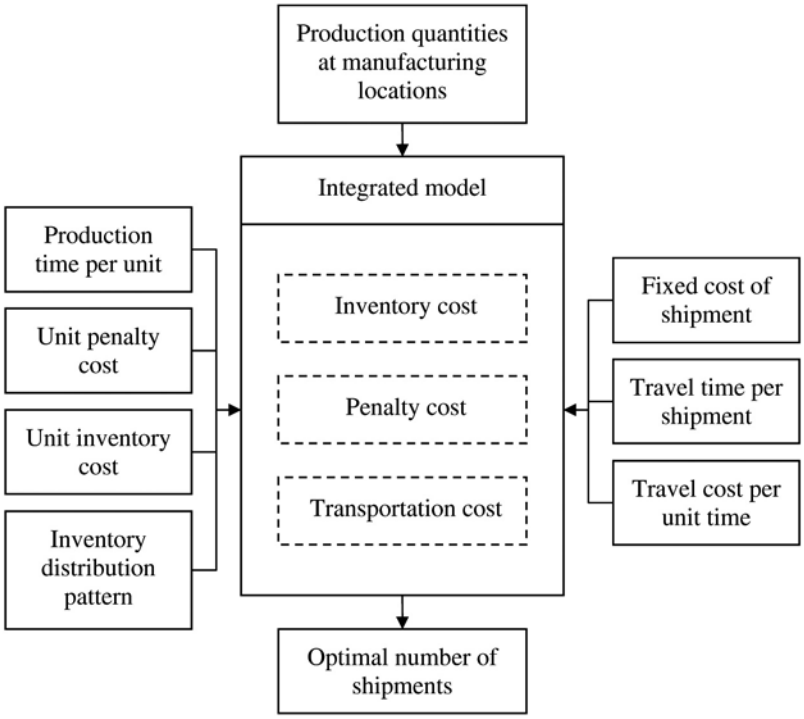

Fig. 1. Detailed framework of the integrated model.

2.1 - Case 1: Continuous production, instantaneous shipping; *PR = SR (i.e., i = 1)*

In this case the production and shipping functions are synchronized and occur at the same rate *(i.e., i = 1).*

This case has been considered in Masel and Pujari [17], where closed form expressions have been obtained for inventory costs, customer waiting costs, and transportation costs. This can be considered as the simplest and one of the most commonly occurring cases. According to the authors, for this case, products are produced for the first shipment and continue to be produced while the first shipment is being loaded and shipped to the customer which requires time l. These additional products produced in time l are then added to the inventory.

This lead time (l) may include time required for activities such as product being shipped from the manufacturing facility to DC, loading time, etc. Thus the first shipment of products is shipped to the customer after time $l + iVy$ where l is the lead time for the first shipment and iVy*(i.e.,*Vy since i*=1)* is the time required to produce the first V units. In other words, the time between shipments is large enough to permit production of at least V units between shipments. The subsequent shipments occur at regular intervals of time iVy until the full order is shipped. It is assumed here that production will always stop at the end of any time period iVy and not during that time period. This helps in maintaining uniform periodicity in the inventory distribution patterns. Fig. 2 shows the inventory distribution pattern for Case 1.

The total inventory held over the entire time horizon, obtained by the summation of area under the graph, determines the total inventory costs. For ease of computation, the total area is divided into distinct and unique sub-areas denoted by I, II, III, IV, and V to be calculated separately. The calculation of these sub-areas for other cases is described later in detail. Further, the penalty and transportation costs are obtained. Finally the sum of all these costs is minimized to determine the optimal number of shipments.

The expression for optimal number of shipments obtained in [17] is shown below.

$$N = \sqrt{\frac{Q^2 y(h+w)}{2(r+tg)}}.$$

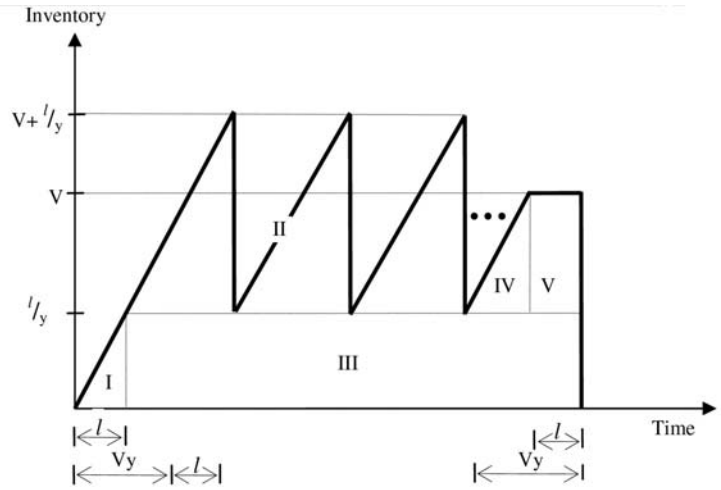

Fig. 2. Case 1: Inventory distribution pattern.

2.2 - Case 2: Continuous production, instantaneous shipping; Two-tiered production rate

In this case production occurs at two different rates while the shipping rate remains constant. As soon as units equivalent to the first shipment size are produced, the shipping activity is commenced resulting in units being shipped to the customer at the end of time (Vy). It represents the time taken to produce V units. Unlike Masel and Pujari [17], however, the lead time for the first shipment l, as described in Case 1, is ignored in this case as well as the subsequent cases, since l is assumed to be insignificant as compared to the total production time for the entire production quantity Q. The subsequent shipments occur at regular intervals until the full order is shipped. However, due to the two-tiered production rate the duration of these intervals change as the production rate changes.

The change in production rate may occur due to various constraints such as availability of machines, workers, machine sharing due to multi-product lines, and so on. Alternately, there could be constraints due to inventory storage limitations on the customer side due to which it may not be possible for the customer to take delivery of the products. So the production rate could be decreased to avoid inventory build-up at the manufacturing facility.

Given that the initial (PR1) and latter (PR2) production rates are different we can express the relationship between the two production rates as follows:

$$\mathrm{PR1} = d^{*}\,(\mathrm{PR2})$$

where,

$$d \in R, \quad d > 1.$$

For the case where $d < l$, the overall inventory level will be almost similar to that for $d > 1$, the magnitude of *'d'* on one side being inverse of that on the other side. The inventory costs will depend on the actual value of *'d'*. However, the penalty costs will be higher for $d < 1$ because higher number of products will be shipped late. The case with $d > 1$ is considered here for analysis as shown in Fig. 3.

The shipments occur at regular intervals of time (Vy) during one half of the production cycle and time ($\mathrm{d}Vy$) during the other half until the full order is shipped. The production rate could be higher for the first half and slower for the next half of the production cycle, as shown in Fig. 3, or vice versa.

For ease of computation and exposition, following assumptions are considered:

a. The change in production rate occurs after half the shipments are completed. There can be potentially infinite number of cases depending on when and how many times the production rate changes throughout the cycle. So to model this system behavior to some extent, it becomes necessary to consider some degree of specification in the scenario.

b. To support the first assumption it also becomes important to assume that the number of shipments is even. This helps in maintaining equal number of shipments on either side of the change in production rate thus helping the modeling effort.

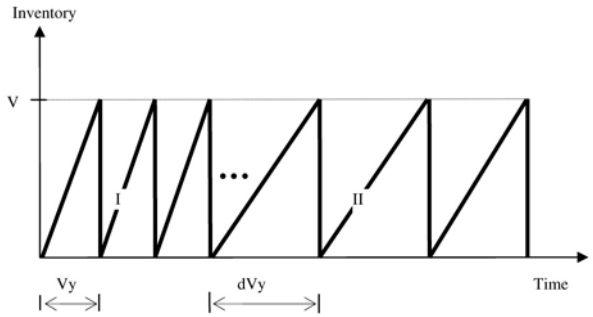

Fig. 3. Case 2: Inventory distribution pattern.

c. The change in production rate does not occur while a shipment is being produced, but at the end of production of a shipment. In other words, this change occurs at the end of any time period Vy (or $\mathrm{d}Vy$) and not during that time period.

d. It is also assumed that production will always stop at the end of any time period Vy (or $\mathrm{d}Vy$) and not during that time period.

2.2.1. Inventory cost

The total inventory held during the full shipping cycle is equal to the total area covered by the graph. It is divided into two unique sub-areas which are calculated as shown in the following expressions. For example, Area I is calculated by multiplying the area of the repetitive triangle with the times it is repeated (N/2).

Area I : $$\left[\left(\frac{1}{2}\right)(Vy)(V)\right]\frac{N}{2}=\frac{NV^2y}{4},$$

Area II : $$\left[\left(\frac{1}{2}\right)(\mathrm{d}Vy)(V)\right]\frac{N}{2}=\frac{\mathrm{d}NV^2y}{4},$$

Total inventory $$=\frac{(d+1)NV^2y}{4}.$$

This expression actually represents the total inventory held over the entire time horizon. It can thus be expressed as the total inventory-time units which, when multiplied by the inventory holding cost, determines the total inventory cost.

So the total inventory cost is given by,

$$C_I=\left(\frac{(d+1)NV^2y}{4}\right)h=\frac{(d+1)Q^2yh}{4N}.$$

The shipment size (V) is obtained by dividing the total production quantity (Q) by the number of shipments (N). This same relationship is used to determine the expression for inventory cost for all the subsequent cases.

2.2.2 - Penalty cost

Time required for first shipment = $(Vy + t)$

Penalty cost for first shipment = $(Vy + t)\, wV$

Similarly; penalty cost for second shipment = $(2Vy + t)\, wV$

So, penalty cost for $[N/2]$th shipment,

$$= \left(\sum_{n=1}^{\frac{N}{2}} (nVy) + t \right) wV = \frac{N^2 V^2 yw}{8} + \frac{NV^2 yw}{4} + \frac{NVtw}{2}.$$

Similarly,

Penalty cost for $[N/2 + 1]$th shipment $\left(\left(\frac{N}{2} + 1 \right) \mathrm{d}Vy + t \right) wV,$

Penality cost for [(N - 1)/3 + 2]th shipment $\left(\left(\frac{N}{2} + 2 \right) \mathrm{d}Vy + t \right) wV.$

So, penalty cost for Nth shipment,

$$= \left(\sum_{n=1}^{\frac{N}{2}} \left(\left(\frac{N}{2} + n \right) \mathrm{d}Vy + t \right) \right) wV = \frac{3\mathrm{d}N^2 V^2 yw}{8} + \frac{\mathrm{d}NV^2 yw}{4} + \frac{NVtw}{2}.$$

So total penalty cost for 1 to N shipments is given by,

$$C_P = \frac{(3d+1)N^2 V^2 yw}{8} + \frac{(d+1)NV^2 yw}{4} + NVtw = \frac{(3d+1)Q^2 yw}{8} + \frac{(d+1)Q^2 yw}{4N} + Qtw.$$

2.2.3 - Transportation cost

Transportation cost is given by,

$$C_T = (r + tg)N.$$

The optimal number of shipments N that minimizes the total distribution cost function, obtained by summing the total inventory cost, penalty cost, and the transportation cost, is found from the first order conditions as follows:

Minimize $C_{IPT} = \frac{(d+1)Q^2 yh}{4N} + \frac{(3d+1)Q^2 yw}{8} + \frac{(d+1)Q^2 yw}{4N} + Qtw + (r + tg)N$

$$\frac{\partial}{\partial N} C_{IPT} = -\frac{(d+1)Q^2 yh}{4N^2} - \frac{(d+1)Q^2 yw}{4N^2} + (r + tg) = 0.$$

Optimal number of shipments,

$$N = \sqrt{\frac{(d+1)Q^2y(h+w)}{4(r+tg)}}.$$

2.3 - Case 3: Continuous production, instantaneous shipping; *PR>SR*

Here the production function occurs at a faster rate as compared to the shipping function. This may occur due to factors such as limited availability of transport services, limited availability of workers for shipping functions, or other similar constraints. Time (Vy) represents the time taken to produce V products. The first shipment of products is shipped to the customer after time iVy, which is the time between two shipments. In other words, the time between shipments is large enough to permit production of at least iV units. As noted previously, l, the lead time for the first shipment is ignored. The expression for the time between shipments is determined by the difference between production rate and shipping rate. Let us consider that their relation is given by,

$$\mathrm{PR} = i^* (\mathrm{SR}),$$

where,

$$i \in R^+, \quad i \geqslant 1.$$

Then time between shipments is given by,

$$L = i^* V^* y.$$

For the case where $i = 1$, and ignoring l, Case 3 reduces to Case 1.

The subsequent shipments occur at regular intervals of time *iVy* until the full order is shipped. For ease of computation, it is assumed for this case as well as for all the subsequent cases that production will always stop at the end of any time period iVy and not during that time period. In this type of scenario, optimization of number of shipments, and in turn, shipment size would be needed mostly when transportation costs are sufficiently larger than inventory costs. If they are not, it would still be cost effective to ship products as and when shipping is possible, without any optimization of the shipment size.

Depending on the value of i, the total number of shipments (N), and the inventory distribution pattern, the number of shipments sent during and after production follow a fixed ratio. After detailed analysis of the distribution patterns, the following Table 2 is prepared which accurately determines the ratio for a particular case.

The table shows ratios *x : y* where, *x* and *y* are the number of shipments during and after production, respectively. In most of the cases, the production and shipment activities are carried out simultaneously until the production stops. After that, shipments are continued from the existing inventory levels. For $i = 2$, N will always be an even number. For $i \geqslant 3$, the values taken by N will be either "multiple of i" or "multiple of i" + 1. When N takes a value equal to "multiple of i"+1 the production may need to be extended to produce an additional shipment quantity to obtain that value of N. These cases are marked with an asterisk (*). These modifications are evident from the inventory patterns in the cases, for example, where $i = 3$ and $N = 4, 7, 10. \ . \ .$ etc. The table shows the ratios for limited values of i and N, the basic purpose being to show the underlying patterns of periodicities in the ratios. This table could be further extended for more values of i and N for future research purposes.

Also, based on the value of i, there can be two sub-cases as described below. For this research, it is assumed that $i = \{1,2,3\}$. This assumption takes into consideration real-life situations where it would be feasible to expect that production rate will not exceed the shipping rate by a factor of more than 3. Since the case for $i = 1$ has already been considered, we now consider two sub-cases for $i = 2$ and $i = 3$, respectively.

2.3.1 - Case 3a: Continuous production, instantaneous shipping; *PR = i* (SR) ; i = 2*

In this sub-case, the production rate is twice the shipping rate as shown in Fig. 4. This implies that products equivalent to two shipment sizes are produced during the time between shipments. Furthermore, production

Table 2
Ratio of shipments during and after production

No. of shipments	Value of i for PR = i*(SR)					
i = 2	i = 3	i = 4	i = 5	i = 6	i = 7. . .	
N = 2	1:1					
N = 3		1:2				
N = 4	2:2	1:3*	1:3			
N = 5			1:4*	1:4		
N = 6	3:3	2:4		1:5*	1:5	
N = 7		2:5*			1:6*	1:6
N = 8	4:4		2:6			1:7*
N = 9		3:6	2:7*			
N = 10	5:5	3:7*		2:8		
N = 11				2:9*		
N = 12	6:6	4:8	3:9		2:10	
N = 13		4:9*	3:10*		2:11*	
N = 14	7:7					2:12
N = 15		5:10		3:12		2:13*
N = 16	8:8	5:11*	4:12	3:13*		
N = 17			4:13*			
N = 18	9:9	6:12			3:15	
N = 19		6:13*			3:16*	
N = 20	10:10		5:15	4:16		
N = 21		7:14	5:16*	4:17*		3:18
N = 22. . .	11:11	7:15*				3:19*

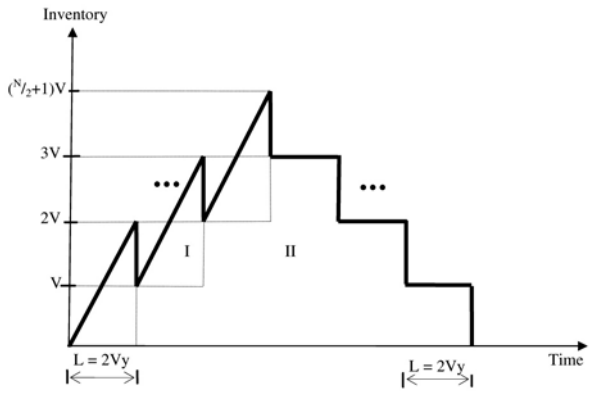

Fig. 4. Case 3a: Inventory distribution pattern.

of the entire order coincides with the completion of half of the necessary shipments. This is depicted as the "saw-tooth" region in Fig. 4. Then the remaining products are stored in the inventory while shipments are being carried out at regular intervals until the entire order has been shipped and inventory level reaches zero. This is shown as the "stair-steps" in Fig. 4. It is important to note that the number of shipments will always be even in this case.

The optimal number of shipments is then determined by,

$$N = \sqrt{\frac{Q^2 y(h+w)}{2(r+tg)}}.$$

Please refer to the Appendix for a detailed account of how the inventory, penalty, and transportation costs are calculated for all cases.

2.3.2 - Case 3b(i): Continuous production, instantaneous shipping; *PR = i (SR); i = 3; N* not a multiple of 3

In this sub-case, the production rate is three times the shipping rate. This implies that products equivalent to three shipment sizes are produced during the time between shipments. Furthermore, (similar to the previous sub-case) the production of entire order coincides with the completion of a third of the necessary shipments. This sub-case has been further classified into two sub-cases of its own depending on whether the total number of shipments, N, is a multiple of three or not. In other words, since $i = 3$, N takes a value which is either a "multiple of 3" or "multiple of 3" + 1. This further classification helps in accurately modeling the inventory and penalty costs which are different due to the slight change in the inventory distribution patterns. When *N* is a multiple of three (case 3b(i)), an additional batch of V products has to be manufactured at the end of the production run. When N is a multiple of three (case 3b(ii)), there is no need to produce this small batch, and shipment can commence right after the last batch of $3V$ units has been produced. These two variations of case 3b are shown in Figs. 5 and 6, specifically at the "top" of their respective distribution graphs. Thus, the optimal number of shipments is as follows,

$$N = \sqrt{\frac{3Q^2 y(h+w)}{2(r+tg)}}.$$

2.3.3 - Case 3b(ii): Continuous production, instantaneous shipping; *PR = i (SR); i = 3; N* is a multiple of 3

Fig. 6 shows the inventory distribution pattern for this case.

Likewise, the optimal number of shipments is,

$$N = \sqrt{\frac{3Q^2 y(h+w)}{2(r+tg)}}.$$

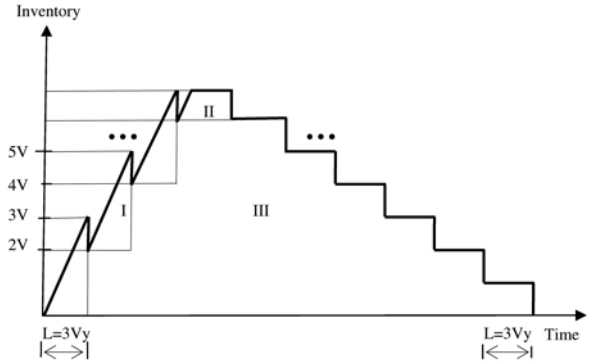

Fig. 5. Case 3b(i): Inventory distribution pattern.

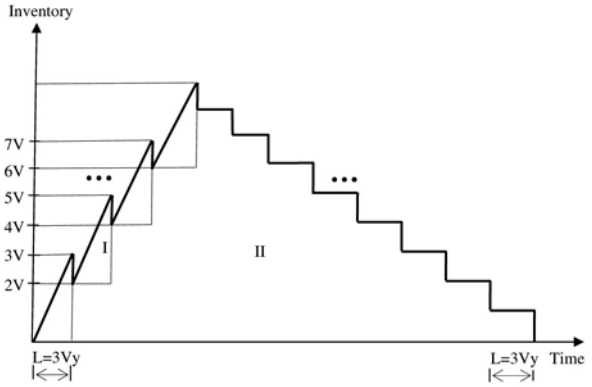

Fig. 6. Case 3b(ii): Inventory distribution pattern.

3. Implementation

The integrated model considers each plant–customer–product relationship individually and minimizes the inventory, penalty, and transportation costs by determining the optimal number of shipments. Let us consider an example problem with 5 locations, 2 customers and 3 products. The production quantities for each product at each plant for each customer are assumed to be determined from previous analysis. Let us consider one of the allocations for our model implementation, say, $Q = 2000$ units. It means 2000 units of a particular product are produced for a specific customer at an allocated plant. Table 3 shows additional data for the problem. If we consider the relation being modeled by case 3a, the optimal number of shipments is obtained by,

$$N = \sqrt{\frac{Q^2 y(h+w)}{(r+tg)}}.$$

Table 3
Data for example problem

Parameters	Notation	Value
Unit production time	y	0.5 hours/unit
Unit penalty cost	w	0.001 $/unit/hour
Unit inventory holding cost	h	0.05 $/unit/hour
Total quantity produced	Q	2000 units
Fixed cost of shipment	r	500 $
Travel cost per unit time per shipment	g	50 $/hour/ship
Travel time for a shipment	t	5 hours

Substituting the values we obtain the following results:

Number of shipments = 12
Shipment size = 167 units
Inventory cost = $29166.67
Penalty cost = $2176.67
Transportation cost = $9000.00
Total cost = $40343.33

We can continue the same calculations for each value of Q and determine the optimal number of shipments depending on the case being modeled for the relation represented by that Q.

Conclusions and future research

The literature pointed to an unexplored area of research in the use of continuous approximation approaches in developing inventory distribution schemas for supply chains. This research provided a continuous approximation procedure is proposed for tactical level optimization of inventory and distribution functions. In doing so, closed-form expressions were obtained for the optimal number of shipments for the different cases presented here. Table 4 presents a comprehensive list of the closed form expressions for all the cases analyzed herein.

Table 4
List of cases and corresponding closed-form expressions for N

S. no.	Case description	Case no.	i	Constraint on the value of N	Expression for optimal value of N
1	Continuous production, instantaneous shipping	1	1	–	$N = \sqrt{\frac{Q^2 y(h+w)}{2(r+tg)}}.$ *(Source: Masel and Pujari [17])*
					$N = \sqrt{\frac{(d+1)Q^2 y(h+w)}{4(r+tg)}}.$ [PR1=d^*(PR2)]
3	Continuous production, instantaneous shipping	3a	2	–	$N = \sqrt{\frac{Q^2 y(h+w)}{(r+tg)}}.$
4		3b(i)	3	N not a multiple of 3	$N = \sqrt{\frac{3Q^2 y(h+w)}{2(r+tg)}}.$
5		3b(ii)		N is a multiple of 3	$N = \sqrt{\frac{3Q^2 y(h+w)}{2(r+tg)}}.$

These closed form expressions for the optimal number of shipments and the continuous approximation procedure developed show a high degree of robustness with respect to different input parameters. Indeed, this is one of the main reasons for originally opting for this approach in order to solve the integrated model for minimizing total costs. All of the cases presented in this research covered real-life situations that may arise from varying production and shipping rates for continuous production and instantaneous shipping scenarios. This provides a comprehensive coverage of inventory distribution patterns to minimize inventory, penalty, and transportation costs.

Analysis of expressions from Table 4 shows that the final expression for optimal value of N does not change depending on the condition that N is or is not a multiple of 3. This condition though, definitely affects the inventory distribution pattern. Specifically, when N is not a multiple of 3, it necessitates the slight adjustment in production (discussed previously). However, it does not affect the inventory, penalty, and transportation costs thus maintaining the same expression for optimal value of N.

In terms of areas for future research, the cases and sub-cases introduced herein considered production rate up to three times the shipping rate. Although this presents a reasonably realistic situation to analyze, it still poses some limits in terms of a generic analysis. Future research could be directed towards finding a generic expression for the optimal number of shipments for a given value of i, where i represents the ratio of production rate to shipping rate.

Also, in terms of the ratios presented in Table 2, this research presents a limited set of ratios for N and i values. Future research could be further expanded to study the ratio patterns and propose expressions to predict ratios for a given combination of N and i values. Although it may not be possible to obtain these ratios for any combination of N and i, it may still be possible to develop expressions for certain combinations. There may also be governing factors such as upper boundaries to these values which could be considered in determining the ratios.

The penalty function used herein for lateness was assumed to be linear with time. In reality, most customers are willing to wait for a period before a penalty should be imposed. Future research could be aimed at relaxing this assumption by adding a time window for on-time delivery for which no penalty is attached.

Lastly, the research presented herein could be further extended to study the effects of, and implications for managers, gradual shipping with continuous and discontinuous production and shipments. Closed-form expressions can then be obtained for the optimal number of shipments for the variations of each case.

Appendix A. Calculation of inventory, penalty, and transportation costs for each case

A.1 - Case 3a: Continuous production, instantaneous shipping; $PR = i\,(SR);\ i = 2$

A.1.1. Inventory cost

The total inventory held during the full shipping cycle is equal to the total area covered by the graph. It is divided into two sub-areas which are calculated as follows:

$$\text{Area I}: \quad \left[\left(\frac{1}{2}\right)(2Vy)(2V)\right]\frac{N}{2} = NV^2y,$$

$$\text{Area II}: \quad \left[\sum_{n=1}^{\frac{N}{2}-1} nV(2Vy)\right] + \left(\frac{N}{2}\right)V(2Vy) = \frac{N^2V^2y}{2} - NV^2y + NV^2y,$$

$$\text{Total inventory} \quad = NV^2y + \frac{N^2V^2y}{2} = \frac{NQ^2y}{N^2} + \frac{N^2Q^2y}{2N^2}.$$

So the inventory cost is given by,

$$C_I = \left(\frac{NQ^2y}{N^2} + \frac{N^2Q^2y}{2N^2}\right)h = \frac{Q^2yh}{N} + \frac{Q^2yh}{2}.$$

A.1.2 - Penalty cost

Time required for first shipment = $(2Vy + t)$

Penalty cost for first shipment = $(2Vy + t)\ wV$

Penalty cost for second shipment = $(4Vy + t)\ wV$

Total penalty cost from 1 to $N/2$ shipments,

$$= \sum_{n=1}^{\frac{N}{2}} (n \cdot 2Vy + t)wV = \frac{N^2V^2yw}{4} + \frac{NV^2yw}{2} + \frac{NVtw}{2}.$$

Similarly,

$$\text{Penalty cost for } (N/2+1)\text{th shipment} = \left(\left(\frac{N}{2}+1\right)2Vy + t\right)wV,$$

$$\text{Penalty cost for } (N/2+2)\text{th shipment} = \left(\left(\frac{N}{2}+2\right)2Vy + t\right)wV.$$

Total penalty cost from $N/2 + 1$ to N shipments,

$$= \sum_{n=1}^{\frac{N}{2}} \left(\left(\frac{N}{2} + n \right) 2Vy + t \right) wV = \frac{3N^2V^2yw}{4} + \frac{NV^2yw}{2} + \frac{NVtw}{2}.$$

So total penalty cost for 1 to N shipments is given by,

$$C_P = N^2V^2yw + NV^2yw + NVtw = Q^2yw + \frac{Q^2yw}{N} + Qtw.$$

A.1.3 - Transportation cost

Transportation cost is given by,

$$C_T = (r + tg)N.$$

The costs are then added, and differentiated with respect to N and equated to zero, to obtain the expression for optimal number of shipments which minimizes total distribution costs.

Minimize $$C_{IPT} = \frac{Q^2yh}{2} + \frac{Q^2yh}{N} + Q^2yw + \frac{Q^2yw}{N} + Qtw + (r + tg)N$$

$$\frac{\partial}{\partial N} C_{IPT} = -\frac{Q^2yh}{N^2} - \frac{Q^2yw}{N^2} + (r + tg) = 0,$$

$$N = \sqrt{\frac{Q^2y(h + w)}{2(r + tg)}}.$$

A.2 - Case 3b(i): Continuous production, instantaneous shipping; $PR = i\,(SR);\ i = 3;\ N$ not a multiple of 3

A.2.1 - Inventory cost

The total inventory held during the full shipping cycle is equal to the total area covered by the graph. It is divided into three sub-areas which are calculated as follows:

Area I : $$\left[\left(\frac{1}{2} \right) (3Vy)(3V) \right] \frac{N - 1}{3} = \frac{3NV^2y}{2} - \frac{3V^2y}{2},$$

Area II : $$\left(\frac{1}{2} \right) (Vy)(V) + (2Vy)(V) = \frac{5V^2y}{2},$$

Area III : $$2 \left[\sum_{n=1}^{\frac{N-1}{3}} n(2V)(3Vy) \right] + \left[\sum_{n=1}^{\frac{N-1}{3}} (2n - 1)(V)(3Vy) \right] = N^2V^2y - V^2y,$$

Total inventory $$= \frac{3NV^2y}{2} + N^2V^2y.$$

So the inventory cost is given by,

$$C_I = \left(\frac{3NV^2y}{2} + N^2V^2y\right)h = \frac{3Q^2yh}{2N} + Q^2yh.$$

A.2.2 - Penalty cost

Time required for first shipment = $(3Vy + t)$

Penalty cost for first shipment = $(3Vy + t)\ wV$

Similarly; penalty cost for second shipment = $(6Vy + t)\ wV$

So, penalty cost for $[(N - 1) / 3\ ]$th shipment,

$$= \left(\sum_{n=1}^{\frac{N-1}{3}}(n(3Vy)) + \left(\frac{N-1}{3}\right)t\right)wV = \frac{N^2V^2yw}{6} + \frac{NV^2yw}{6} - \frac{V^2yw}{3} + \frac{NVtw}{3} - \frac{Vtw}{3}.$$

Similarly,

Penalty cost for $[(N - 1) / 3 + 1\ ]$th shipment, $\left(\left(\frac{N-1}{3} + 1\right)3Vy + t\right)wV,$

Penalty cost for $[(N - 1) / 3 + 2\ ]$th shipment, $\left(\left(\frac{N-1}{3} + 2\right)3Vy + t\right)wV.$

So, penalty cost for Nth shipment,

$$= \left(\sum_{n=1}^{\frac{2N+1}{3}}\left(\left(\frac{N-1}{3} + n\right)3Vy + t\right)\right)wV = \frac{8N^2V^2yw}{6} + \frac{8NV^2yw}{6} + \frac{V^2yw}{3} + \frac{2NVtw}{3} + \frac{Vtw}{3}.$$

So total penalty cost for 1 to N shipments is given by,

$$C_P = \frac{3N^2V^2yw}{2} + \frac{3NV^2yw}{2} + NVtw = \frac{3Q^2yw}{2} + \frac{3Q^2yw}{2N} + Qtw.$$

A.2.3 - Transportation cost

Transportation cost is given by,

$$C_T = (r + tg)N$$

The costs are then added, and differentiated with respect to N and equated to zero, to obtain the expression for optimal number of shipments which minimizes total distribution costs.

Minimize $\quad C_{IPT} = \frac{3Q^2yh}{2N} + Q^2yh + \frac{3Q^2yw}{2} + \frac{3Q^2yw}{2N} + Qtw + (r + tg)N$

$$\frac{\partial}{\partial N}C_{IPT} = -\frac{3Q^2yh}{2N^2} - \frac{3Q^2yw}{2N^2} + (r + tg) = 0,$$

$$N = \sqrt{\frac{3Q^2y(h + w)}{2(r + tg)}}.$$

A.3 - Case 3b(ii): Continuous production, instantaneous shipping; $PR = i\,(SR)$; $i = 3$; N is a multiple of 3

A.3.1 - Inventory cost

The total inventory held during the full shipping cycle is equal to the total area covered by the graph. It is divided into two sub-areas which are calculated as follows:

Area I : $$\left[\left(\frac{1}{2}\right)(3Vy)(3V)\right]\frac{N}{3} = \frac{3NV^2y}{2},$$

Area II : $$\left[\sum_{n=1}^{\frac{N}{3}-1} n(2V)(3Vy)\right] + \left[\sum_{n=1}^{\frac{N}{3}}(2n-1)(V)(3Vy)\right] + \left[\left(\frac{2N}{3}\right)(V)(3Vy)\right] = N^2V^2y,$$

Total inventory $$= \frac{3NV^2y}{2} + N^2V^2y.$$

So the inventory cost is given by,

$$C_I = \left(\frac{3NV^2y}{2} + N^2V^2y\right)h = \frac{3Q^2yh}{2N} + Q^2yh.$$

A.3.2 - Penalty cost

Time required for first shipment = $(3Vy + t)$

Penalty cost for first shipment = $(3Vy + t)\,wV$

Similarly; penalty cost for second shipment = $(6Vy + t)\,wV$

So, penalty cost for $[N/3]$th shipment,

$$= \left(\sum_{n=1}^{\frac{N}{3}}(n(3Vy)) + t\right)wV = \frac{N^2V^2yw}{6} + \frac{NV^2yw}{2} + \frac{NVtw}{3}.$$

Similarly,

Penalty cost for $[N/3 + 1]$th shipment, $$= \left(\left(\frac{N}{3} + 1\right)3Vy + t\right)wV,$$

Penalty cost for $[(N - 1)/3 + 2]$th shipment, $$= \left(\left(\frac{N}{3} + 2\right)3Vy + t\right)wV.$$

So, penalty cost for Nth shipment,

$$= \left(\sum_{n=1}^{\frac{2N}{3}}\left(\left(\frac{N}{3} + n\right)3Vy + t\right)\right)wV = \frac{4N^2V^2yw}{3} + NV^2yw + \frac{2NVtw}{3}.$$

So total penalty cost for 1 to N shipments is given by,

$$C_P = \frac{3N^2V^2yw}{2} + \frac{3NV^2yw}{2} + NVtw = \frac{3Q^2yw}{2} + \frac{3Q^2yw}{2N} + Qtw.$$

A.3.3 - Transportation cost

Transportation cost is given by,

$$C_T = (r + tg)N$$

The costs are then added, and differentiated with respect to N and equated to zero, to obtain the expression for optimal number of shipments which minimizes total distribution costs.

Minimize $$C_{IPT} = \frac{3Q^2yh}{2N} + Q^2yh + \frac{3Q^2yw}{2} + \frac{3Q^2yw}{2N} + Qtw + (r + tg)N$$

$$\frac{\partial}{\partial N} C_{IPT} = -\frac{3Q^2yh}{2N^2} - \frac{3Q^2yw}{2N^2} + (r + tg) = 0,$$

$$N = \sqrt{\frac{3Q^2y(h + w)}{2(r + tg)}}.$$

References

[1] J. Benjamin, An analysis of inventory and transportation costs in a constrained network, Transportation Science 23 (3) (1989) 177– 183.

[2] C. Canel, B.M. Khumawala, Multi-period international facilities location: An algorithm and application, International Journal of Production Research 35 (7) (1997) 1891–1910.

[3] C. Canel, B.M. Khumawala, International facilities location: A heuristic procedure for the dynamic uncapacitated problem, International Journal of Production Research 39 (17) (2001) 3975–4000.

[4] P. Chandra, M.L. Fisher, Coordination of production and distribution planning, European Journal of Operational Research 72 (3) (1994) 503–517.

[5] M.A. Cohen, H.L. Lee, Strategic analysis of integrated production–distribution systems: Models and methods, Operations Research 36 (2) (1988) 216–228.

[6] M.A. Cohen, H.L. Lee, Resource deployment analysis of global manufacturing and distribution networks, Journal of Manufacturing and Operations Management 2 (1989) 81–104.

[7] A. Dasci, V. Verter, A continuous model for production–distribution system design, European Journal of Operational Research 129 (2) (2001) 287–298.

[8] M.S. Daskin, Logistics: An overview of the state of the art and perspectives on future research, Transportation Research – A 19A (1985) 383–398.

[9] R. Ganeshan, Managing supply chain inventories: A multiple retailer, one warehouse, multiple supplier model, International Journal of Production Economics 59 (1999) 341–354. [10] R. Ganeshan, T.P. Harrison, An introduction to supply chain management, Department of Management Science and Information Systems, Penn State University, Retrieved November 2006 from: <http://lcm.csa.iisc.ernet.in/scm/supply_chain_intro.html> (1995).

[11] A.M. Geoffrion, G.W. Graves, Multicommodity distribution system design by Bender's decomposition, Management Science 20 (5) (1974) 822–844.

[12] A.M. Geoffrion, R. Powers, Twenty years of strategic distribution system design: An evolutionary perspective, Interfaces 25 (5) (1995) 105–127.

[13] R.W. Hall, Discrete models/continuous models, Omega: International Journal of Management Science 14 (1986) 213–220.

[14] A. Langevin, P. Mbaraga, J.F. Campbell, Continuous approximation models in freight distribution: An overview, Transportation Research – B 30 (3) (1996) 163–188.

[15] D. Li, C. O'Brien, Integrated decision modeling of supply chain efficiency, International Journal of Production Economics 59 (1999) 147–157.

[16] E. Lodree Jr., C.M. Klein, W. Jang, Minimizing customer waiting in two-echelon, production/distribution supply chains, in: Proceedings of the 12th Industrial Engineering Research Conference, Portland, Oregon, 2003.

[17] D.T. Masel, N.A. Pujari, Considering customer waiting time in supply chain integration, in: Proceedings of the 14th Industrial Engineering Research Conference, Atlanta, Georgia, 2005.

La critique de la théorie de l'exploitation

Par Dr. Patrice Cailleba
Groupe ESC Pau, France

Abstract ASSP

Que reste-t-il de Marx, bien après que sa pensée ait séduit ou suborné intellectuels et démagogues ? Nous proposons une réponse : il s'agit de la théorie de l'exploitation. C'est en critiquant la théorie marxienne de l'exploitation et en dépassant les problèmes économiques, politiques et philosophiques qu'elle laisse apparaître que le projet marxien (et non marxiste[1]) pourra prétendre embrasser les défis que lui impose le capitalisme.

Si, d'après Marx, les ouvriers possèdent leur force de travail, encore possèdent-ils quelque chose puisque c'est de cela que les capitalistes les spolient. D'*où le caractère de propriétaire inhérent aux prolétaires*. Or, par l'abolition de la propriété privée, Marx abolit également la propriété du travailleur sur sa propre force de travail. Il le vole une nouvelle fois. En outre, Marx, prisonnier de son temps, ne voit pas que la fabrique n'est pas le seul endroit où la production puisse se faire dans une économie capitaliste.

Par ailleurs, pour les individus qui n'ont pas accès à la propriété, *être contraint de vendre son propre travail vaut mieux que ne pas avoir le droit de le faire* (comme certaines communautés dans certains pays), ne pas réussir à le faire (comme les chômeurs) ou vivre en marge de la société en tirant une maigre subsistance de la délinquance, de la mendicité ou des miettes du patrimoine collectif (comme le *Lumpenproletariat* de Marx).

Au-delà de l'aspect économique, le problème soulevé par la théorie de l'exploitation est le suivant : *La liberté d'un individu, ou plutôt d'une classe, dépend-elle de l'asservissement d'une autre classe ?* Tout d'abord, il est indéniable que *cela peut être vrai*. Et il est encore moins contestable que *cela fut vrai à plusieurs moments de l'Histoire* (pensons aux différentes formes d'esclavage).

Cependant, l'idée que la liberté d'une classe dépende de l'asservissement d'une autre classe demeure source d'interrogations dans une économie aussi complexe (nombre de classes sociales) et mondialisée (inter-dépendance des économies) que la nôtre au XXIème siècle. A la question « qui exploite qui ?», les réponses possibles sont alors légion.

Toutefois, si l'exploitation d'un individu permet à certains d'obtenir ou de conserver une forme de liberté qu'ils enlèvent à d'autres, alors *est-ce que la liberté, pour être réalisée, a besoin de l'absence de liberté ?*

En cas de réponse positive, nous serions en présence de la dialectique hégélienne du maître et de l'esclave. Le problème serait alors de considérer l'apparition d'un tel phénomène : serait-il diachronique ou synchronique ? Ensuite, de philosophique, notre questionnement deviendrait métaphysique. Il faudrait s'interroger sur *la qualité de la souffrance en jeu dans la théorie de l'exploitation* qui renvoie directement à l'eschatologie chrétienne.

1. « Le marxisme est l'ensemble des contresens qui ont été faits sur Marx » in Michel Henry, Marx, tome I : Une philosophie de la réalité, Gallimard, Paris, 1976, réed. 1991, p.9.

En 1857, lors de son *Introduction à la critique de l'économie politique*, Marx a dressé le tableau de l'art grec, ce « paradis perdu ». Aux yeux du philosophe, l'exemple grec pose la question du rapport entre, d'un côté, le fondement social et économique d'une époque historique donnée et, de l'autre, les oeuvres d'art conçues durant cette époque. Marx décèle des strates différentes, des inégalités « entre l'évolution de l'art en général et celle de la société ». L'art est un « jaillissement », la figuration, l'incarnation d'une époque. Dès lors que sa mythologie s'estompe, la réalité connaît une mutation profonde, et tout le genre littéraire de la « terre nourricière », la poésie épique en premier lieu, perd alors de sa substance génitrice.

C'est pourquoi Marx pose la question de la pertinence et de la possibilité de l'Iliade au XIX^{ème} siècle :

> *« Achille est-il possible à l'âge de la poudre et du plomb ? (...) de la presse et surtout [de] la machine à imprimer ? »*[2]

En fait, le philosophe allemand insiste sur le fait que, bien qu'évanoui et formant, pour nous, une sorte de modèle devenu impénétrable à notre regard contemporain, l'art grec nous procure encore un plaisir esthétique. Il en est ainsi de certaines oeuvres artistiques et intellectuelles qui semblent transcender leur époque.

Au-delà de l'art grec, l'exemple que nous propose Marx nous interroge sur l'influence de sa propre philosophie. Que reste-t-il de Marx plus de cent vingt ans après sa mort, bien après que sa pensée ait suborné intellectuels et démagogues et que les régimes qui s'en sont réclamés aient, presque tous, périclité ? Comment peut-on expliquer la rémanence de sa pensée encore aujourd'hui alors qu'elle se révèle être anti-politique[3] ? Bref, quel est le concept tiré de son oeuvre qui semble faire office de factotum dans tous les discours politiques d'inspiration marxienne ? Nous proposons ici une réponse : il s'agit de la théorie de l'exploitation.

1. Critique économique

Dans *ses Théories de la justice*, W. Kymlicka se penche sur la théorie marxienne de l'exploitation afin d'aborder le problème de la justice envisagée par Marx et par le marxisme après lui. Malgré l'excellent tour d'horizon que nous propose Kymlicka, nous pensons que ce dernier, en raison du cadre restreint de son étude, n'a pas cerné dans sa globalité le problème soulevé par la question de l'exploitation telle que l'entend Marx. Kymlicka opère en effet une analyse juste mais partielle en se focalisant essentiellement sur le vol, c'est-à-dire sur l'aspect économique, qu'induit la théorie de l'exploitation.

W. Kymlicka relève que Marx fait du prolétaire la victime d'un vol qui a lieu entre les murs de l'usine. Cette analyse de Marx soulève deux points importants. Tout d'abord, partant de la lapalissade que l'on ne peut voler à quelqu'un que ce qu'il possède en propre, Kymlicka considère que Marx fait du prolétaire un propriétaire puisqu'il est, même dénué de tout, *propriétaire* au moins *de sa force de travail*. Ce faisant, sans s'en rendre compte, Marx généralise la propriété privée à l'ensemble de l'humanité en gratifiant les bourgeois comme les ouvriers d'une propriété[4]. Vouloir priver le prolétaire de sa force de travail revient donc à priver l'individu d'un bien inaliénable et à généraliser la situation d'exploitation dans la mesure où on prend à l'ouvrier le peu qu'il lui reste. Ce qui confine à l'absurdité et politique et économique.

2. I-CEP [1857] in PI I, p.266.
3. Pour plus de détails sur « l'anti-politisme » de la philosophie marxienne, voir Patrice Cailleba, L'individu chez Marx, ANRT, Paris, 2005.
4. Ce qui paraît logique puisqu'il a bien fallu qu'un jour les capitalistes ou autres bourgeois utilisent leurs propres forces de travail (forces physiques brutes et/ou forces intellectuelles) afin de pouvoir exploiter, selon la perspective marxienne, d'autres individus qu'eux-mêmes. De fait, selon Marx, un bourgeois, un capitaliste, n'est qu'un ancien ouvrier agricole.

Ensuite, Marx fait du travail ouvrier en usine l'archétype du travail salarié dans la société capitaliste. Or, une telle conception du travail salarié a difficilement résisté à l'épreuve du temps[5]. Les sociétés capitalistes d'Europe occidentale et d'Amérique du Nord ne sont plus seulement des sociétés industrielles mais sont devenues des sociétés de services.

L'industrialisation semble avoir été un passage d'un mode particulier de production (féodal ou proto-industriel) à un autre (industriel et post-industriel). Certains pays asiatiques (Singapour, Corée du Sud, Taiwan ou le Japon) qui se sont inspirés du mode de développement européen illustrent également le caractère transitoire de cette évolution. Pour autant, cette industrialisation ne constitue en rien une étape obligatoire sur le chemin qui mène à l'établissement d'une société post-industrielle : il existe en effet certains Etats, du Luxembourg aux micro-États des Caraïbes et du Pacifique, qui n'ont pas connu de période d'industrialisation mais qui sont rangés parmi les pays riches[6]. Enfin, le travail ouvrier en usine a lui-même changé. Il est souvent difficile de comparer le travail d'un ouvrier ou d'un employé au début du XXIème siècle en Angleterre ou en France avec ce qu'il était au XIXème dans ces mêmes pays. Un travailleur peut dorénavant choisir plus librement un type de travail où sa tâche n'est pas réduite à une simple fonction machinale[7]. Il peut aussi se former avant et tout au long de sa vie professionnelle. Ainsi, sans amoindrir le caractère encore brutal de l'industrialisation dans certains pays (Chine, Inde, Vietnam, etc.), convient-il de noter que la vision de Marx semble limitée dans le temps.

De plus, cette conception semble également limitée dans l'espace : l'usine n'est pas le lieu idéal typique de la production. En effet, il est difficilement contestable qu'un travailleur puisse exercer une activité rémunérée en dehors de toute structure, bref qu'il puisse être indépendant et travailler « à son compte » en coiffure comme en médecine générale, dans une société industrielle comme dans une société de services. Tous ces points, que nous reprenons rapidement, sont déjà bien connus des critiques économiques de la théorie de l'exploitation[8].

Par conséquent, sans amoindrir le flux des conditionnements qui pèsent sur un individu salarié, que ce soit dans un pays riche mais aussi dans un pays émergent ou en voie de développement, la théorie de Marx n'apparaît pas correcte d'un point de vue économique.

Pour autant, Marx ne nourrissait pas simplement un but économique lorsqu'il a repris cette idée de l'exploitation de l'homme par l'homme pour en faire une théorie. En fait, il y a ajouté une dimension politique et philosophique qui est tout autant contestable et dont découle une conception particulière de la liberté.

5. Du vivant même de Marx, ce dernier était conscient que le travail en usine ne constituait pas la forme de travail la plus répandue dans les pays dits industrialisés (Angleterre, Etats-Unis, France, Allemagne). Cf. le caractère majoritaire de la classe paysanne dans la société française au XIXème siècle que Marx avait lui-même souligné : 18B [1852] in Pl IV, p.532 et Critique du Gotha [1875] in Pl I, p.1427.
6. On pourra objecter qu'il est difficilement imaginable que ces pays aient pu se développer en l'absence totale de révolution industrielle connue par les autres pays avec lesquels ils étaient en relation.
7. Même si le secteur industriel est devenu minoritaire dans les économies des pays de l'OCDE par rapport aux services, on constate toutefois une néo-taylorisation du secteur tertiaire qui se rapproche des tâches répétitive et machinale du taylorisme industriel. Cf. Vincent Grosjean et Corinne Ribert-Van de Weerdt, 'Les modes de management dans un call center et leurs conséquences sur le bien-être des opérateurs' in Note Scientifique et Technique, département « Homme au travail », NS234, Juin 2003.
8. W. Kymlicka, Les théories de la justice, Une introduction, Editions La découverte, Paris, 1999 ; J.A. Schumpeter, Histoire de l'analyse économique, trad. par J.-C. Casanova, NRF, Gallimard, 1983, 3 vol. ; M. Baslé et al., Histoire des pensées économiques, Sirey, 1993 ; Boncoeur J. et Thouément H., Histoire des idéeséconomiques, Nathan, t. 1, Paris, 1989 ; etc...

2. Critique politique

C'est une expérience philosophique banale que de constater que, parce que certains hommes sont libres ou parviennent à le devenir, la liberté est alors chose humaine. La question n'est donc pas de savoir si l'homme en général est libre : il l'est ou, en tout cas, peut l'être puisque certains hommes arrivent à être libre et à le rester. La véritable question est alors pourquoi tous ne parviennent pas à l'être. Le jeune Marx[9] avait remarqué que certains hommes, en combattant la liberté d'autres hommes, combattaient tout au moins pour préserver ou augmenter leur propre liberté[10]. Réduire la liberté des autres, c'est donc une manière d'accroître la sienne. Pour autant, si cette idée de la liberté peut constituer à juste titre un des prodromes de la théorie de l'exploitation, cette dernière n'existe pas en tant que telle chez le jeune Marx. Il faut attendre le *Naturalisme dialectique* (que nous identifions comme correspondant à la période des *Manuscrits* de 1844) pour que l'idée d'exploitation fasse son apparition[11], et le *Matérialisme historique* (pour les oeuvres postérieures à 1845) pour que la théorie de l'exploitation investisse définitivement le champ du politique. Avec le *Matérialisme historique*, c'est cette théorie de l'exploitation qui est passée à la postérité et qu'il nous faut remettre en cause en raison de ce qu'elle implique.

Au-delà de l'aspect économique, le problème soulevé par la théorie de l'exploitation est le suivant : *La liberté d'un individu, ou plutôt d'une classe, dépend-elle de l'asservissement d'une autre classe ?* Ou plutôt, l'asservissement de certains hommes s'explique-t-il par l'action volontaire d'autres hommes qui deviennent ainsi libres ? De cette conception manichéenne – un mal pour un bien – dépend tout le problème de la liberté, tel qu'il est envisagé par la théorie marxienne de l'exploitation. Il faut ainsi comprendre ce qu'une telle affirmation induit. Si la liberté d'un individu, et plus généralement d'une classe, dépend de la subordination complète d'une autre classe, nous devons examiner la réalité historique pour vérifier pareille assertion.

Tout d'abord, il est indéniable que cela peut être vrai. Et il est encore moins contestable que cela fut vrai à plusieurs moments de l'Histoire tant il est difficile de ne pas relever dans l'histoire de chaque nation un, ou plutôt plusieurs moments, où la prospérité d'un groupe d'hommes se fait au détriment d'autres hommes qui ne sont pas volontaires et qui n'ont aucun autre choix. Les exemples sont légion : les esclaves des Grecs et des Romains dans l'Europe Antique assuraient le confort des oligarques grecs et des patriciens romains ; les serfs au Moyen Age assuraient l'aisance matérielle et spirituelle des seigneurs ; les esclaves indiens et noirs en Amérique permettaient aux grands propriétaires fonciers d'exploiter les richesses naturelles d'un nouveau continent, les intouchables indiens formaient la base d'un système de castes, etc. Cependant, l'histoire ne se réduit pas à l'affrontement de deux classes antagonistes comme l'écrit Marx dans le *Manifeste*. L'Antiquité romaine ne se résumait pas à une classe de patriciens qui exploitait des esclaves. Il existait des classes intermédiaires (paysans, soldats, commerçants, etc.) qui composaient cet entre-deux social et représentaient la grande majorité de la population active comme il l'a constaté quelques fois à regret[12]. Il en fut pareillement par la suite et ce jusqu'au début du XXème siècle. L'histoire ne se limite pas à l'antagonisme de deux classes, l'une exploitante, l'autre exploitée.

Pourtant, malgré l'abolition de l'esclavage au XIXème siècle et les démentis que l'on peut apporter à une telle conception, la réalité de l'exploitation transparaît ici et là. En effet, une forme plus insidieuse de travail obligatoire rendu volontaire s'est immiscée dans les sociétés dites « modernes ». Il est évident qu'on continue actuellement à trouver des cas précis et particuliers où un tel phénomène subsiste : dans des pays émergents (certaines régions de Chine et du Brésil…), dans des pays en voie de développement (comme dans certains pays africains) mais aussi dans les pays riches (l'ensemble de l'économie parallèle qui regroupe les ateliers de travail « au noir », les filières d'immigration clandestines en Europe ou en Amérique du Nord, etc.).

9. Nous appelons Jeune Marx ses écrits antérieurs à L'Idéologie Allemande écrit en 1845.
10. « La liberté est à ce point essentielle aux hommes que même ses adversaires la réalisent, tout en combattant sa réalité… Nul homme ne combat la liberté ; tout au plus combat-il la liberté des autres. » Gazette Rhénane, Sur la liberté de la presse [Mai 1842] in Pl III, p.166.
11. Dans le Naturalisme dialectique du Marx de 1844-1845, si l'aliénation au genre est générale (tous les hommes sont des êtres génériques conscients, ils sont tous également aliénés), certains hommes – les non-propriétaires – le sont pourtant plus que d'autres – les propriétaires industriels et fonciers.
12. 18B [1852] in Pl IV, p.532 et Critique du Gotha [1875] in Pl I, p.1427.

Malgré ces tristes exemples, l'idée que la liberté d'une classe dépende de l'asservissement d'une autre classe demeure simpliste et inopérante dans une économie aussi complexe et mondialisée que la nôtre au XXIème siècle. Si la liberté des uns dépend de l'exploitation des autres, la vraie question devient alors : Qui exploite qui ? Qui est le coupable ?

Si l'histoire est l'histoire de la lutte des classes, chaque pays du monde voit l'affrontement de deux classes antagonistes, donc l'exploitation de l'une par l'autre. Or, nous avons vu auparavant combien une telle affirmation était sujette à caution et ce, pour Marx lui-même. Si on peut penser qu'une telle compréhension de la société paraît encore plus ou moins justifiée dans un pays en voie de développement, il en est différemment des pays riches où l'éclatement des classes sociales[13], la diversité des formes de travail salarié, l'existence des couvertures sociales (bien que largement inégales entre pays riches et malgré le fait qu'elles puissent être assimilées, par certains – inclus certains marxistes[14] –, à une nouvelle forme d'exploitation) et la protection des travailleurs amoindrissent largement cette « exploitation ». Pour autant, si d'aucuns considèrent encore qu'il existe une exploitation des classes ouvrières par les classes bourgeoises, il nous faut alors élargir notre critique à l'échelle mondiale pour confondre une conception plutôt simple des échanges.

Selon Marx, il existe dans tous les pays une lutte des classes. Les pays riches, comme les pays pauvres, se composent, d'un côté, d'une classe bourgeoise ou dominante et, de l'autre côté, d'une classe ouvrière ou d'autres classes dominées qu'ils sont censés exploiter. Les échanges internationaux qui se tissent à l'échelle mondiale amènent à se poser ainsi une nouvelle question quant à une exploitation possible déjà entrevue par Marx : la classe bourgeoise ou dominante d'un pays riche exploite-t-elle la classe dominante ainsi que la classe ouvrière ou dominée d'un pays pauvre ? L'étude de plusieurs cas de figure typiques devrait nous aider à démonter une théorie trop grossière.

A. *La classe dominante d'un pays riche exploite la classe dominée de son propre pays ainsi que les classes d'autres pays.*

Telle est la vulgate marxienne qui se base sur la notion de plus-value qu'elle justifie en retour dans le même temps. Cependant, dans le cadre de relations inter-nations, il arrive que la classe dominante d'un pays riche exploite non seulement la classe dominée du même pays mais aussi la classe dominante et dominée d'un autre pays, un pays relativement moins riche, voire tout simplement pauvre. Tel est le cas classique de la théorie de l'exploitation défendue par le marxisme. Du fait des lois « élucidées » par Marx – loi de la baisse tendancielle du taux de profit, récurrence des crises économiques, paupérisation croissante des classes sociales –, le capitalisme bourgeois est obligé d'élargir ses débouchés et d'améliorer ses perspectives économiques. Dès lors, la classe bourgeoise cherche à l'étranger les richesses naturelles et les nouveaux marchés qui lui permettront de consolider sa situation dominante. Lénine, en 1916, dans *L'impérialisme, le stade suprême du capitalisme* n'a fait que reprendre cette idée à Marx, qui l'avait lui-même trouvée en se penchant sur la politique coloniale de l'Angleterre.

Cette exploitation peut s'opérer au détriment de la classe dominée nationale et à l'avantage d'une classe dominante ou dominée d'un pays relativement moins riche. A l'époque de Marx, par exemple, les lois sur l'importation de céréales en Angleterre, l'abolition des *Corn Laws*, mirent fin, au milieu du XIXème siècle, au monopole des céréaliers britanniques sur le marché

13. On peut difficilement parler « d'une » classe bourgeoise, comme on peut difficilement parler de « la » classe ouvrière qui composerait la majorité de la société. Cette place est plutôt occupée par la classe dite « moyenne » qui, elle-même, est largement stratifiée. Cf. la nomenclature de l'INSEE des Professions et Catégories Socioprofessionnelles (PCS) qui propose 189 professions, 32 CSP (Catégories Socio-Professionnelles), réduites à 6 PCS sur http://www.insee.fr/fr/nom_def_met/nomenclatures/prof_cat_soc/pages/pcs.htm

14. L'exploitation moderne telle qu'elle est pensée par certains marxistes contemporains inclut dans sa tentative apologétique les « prestations sociales, services collectifs utilisés gratuitement ou à prix réduit » comme appartenant à un « système d'ensemble à l'échelle sociale tout entière » qui conforte cette exploitation dans un « capitalisme monopoliste d'État ». Bref, ce qui est pour le commun des mortels considéré comme la possibilité d'une libération, le reflet d'une société plus juste, est considéré par le marxisme contemporain comme un instrument aliénant de plus : « Dans la phase actuelle du capitalisme monopoliste d'État et si l'on retourne dans les pays de capitalisme développé, l'essentiel des changements dans le processus d'exploitation réside dans la socialisation capitaliste et dans le rôle nouveau de l'État. Le salaire demeure certes toujours le prix de la force de travail mais aujourd'hui le salaire direct qui tend de plus en plus à être fixé par des procédures collectives ne correspond plus, à lui seul, à la valeur de la force de travail. Aussi la structure actuelle du capital variable comprend-elle des éléments salariaux et non salariaux (salaire, prestations sociales, services collectifs utilisés gratuitement ou à prix réduit). Les dépenses de reproduction de la force de travail ne se trouvent donc plus entièrement régulées par le marché. A la phase du capitalisme monopoliste d'État, l'exploitation du travail constitue, de plus en plus, un système d'ensemble à l'échelle sociale toute entière. Par ailleurs l'élargissement à l'ensemble de la nation des sources de l'accumulation capitaliste par la fiscalité, l'inflation, le développement des diverses formes d'épargne forcée, etc., marque le rôle croissant de l'appareil d'État dans l'intensification de l'exploitation capitaliste et le pillage des couches monopolistes. » Guy Caire in Bensussan G. – Labica G., Dictionnaire critique du marxisme, Quadrige / P.U.F., Paris, 1999, p.437 (souligné par nous).

anglais et causèrent la ruine de nombre d'entre eux au profit des exploitations céréalières des colonies, en particulier des exploitations indiennes. Ce cas relève du marxisme classique. L'ouverture à la concurrence étrangère de certains pans de l'économie au détriment d'une classe déterminée suit souvent une telle logique. En fait, dans pareil cas, il n'est pas aisé de trancher entre les gains retirés (bénéfice d'une plus forte concurrence mais aussi d'un plus grand accès à des biens moins chers en plus grand nombre pour le grand public) et les pertes induites (appauvrissement d'une classe, perte d'un savoir-faire, risque environnementaux et/ou sanitaires, mise en péril d'un secteur de l'économie nationale, voire de plusieurs si les secteurs sont interdépendants, etc.). L'analyse à réaliser est alors au cas par cas et doit faire l'économie de toutes les démagogies.

B. La classe dominée d'un pays riche exploite la classe dominée et la classe dominante d'un pays pauvre.

Un tel cas est le fruit d'une économie contemporaine où le capitalisme d'État, aidé et/ou influencé par l'action menée par certaines classes sociales dominantes, mais aussi dominées, peut protéger un secteur donné de son économie. Dès lors, ce secteur particulier peut recevoir une protection indue et injuste par rapport aux secteurs équivalents dans d'autres pays moins riches qui ne peuvent protéger de la même manière leur économie et avec lesquels il est en concurrence. Il suffit ici de penser à l'agriculture européenne et nord-américaine subventionnée par des états riches et désireux de protéger leur agriculture rendue largement dépendante de leurs soins. Nous avons là des classes agricoles qui passent pour être dominées économiquement, mais qui conservent un pouvoir politique notoire leur permettant de défendre leur droit au détriment d'autres classes agricoles, celles d'autres pays riches ou, surtout, d'autres pays pauvres[15] . L'échec du sommet de l'Organisation Mondiale du Commerce de Cancún en 2003, du fait, entre autres, de la pression des pays en voie de développement, a permis de mettre un frein, pour un temps seulement, à cette situation. Un tel échec laisse entrevoir la possibilité qu'une classe, ou un groupe de pays, a de renverser une situation d'exploitation. Toutefois cela est rarement le cas. La multiplicité des accords (accords ACP, accords de Cotonou) en faveur de l'agriculture occidentale subventionnée illustrent parfaitement l'inverse[16].

C. *La classe dominée et la classe dominante d'un pays pauvre peuvent exploiter la classe dominée et la classe dominante d'un pays riche.*

Nous avons vu auparavant que la classe dominée et/ou dominante d'un pays pauvre pouvait exploiter ou, tout du moins, connaître des termes de l'échange favorables au détriment de la classe dominée d'un pays riche (cf. abolition *Corn Laws* en Angleterre au XIXème siècle). De manière générale, cela peut être également vrai : un pays riche peut se trouver « exploité » par un pays relativement plus pauvre dans certaines situations données que la fin du XXème siècle a permis de faire éclore. Ce qui apparaît comme une forme nécessaire de solidarité internationale peut être alors interprété comme une forme *toute relative* d'exploitation.

En dehors des politiques d'aide au développement qui amènent les pays riches à contribuer unilatéralement aux financements de projets (éducatifs, sociaux, économiques, urbains, etc.) dans les pays en voie de développement via les Nations Unies, il existe d'autres types de poli-

15. Au passage, on note que la description de cette situation où une classe particulière, la classe agricole des pays riches, fait de son intérêt privé un intérêt national pour défendre ses « droits » se rapproche aisément de la conception marxienne de l'État : un État de classe à la solde d'intérêts privés. Pour autant, nous sommes loin d'un véritable libéralisme qui se résumerait à l'absence de barrières douanières, de subventions et d'aides fiscales, etc., bref d'un « laissez faire, laissez aller » agricole qui plonge ses racines dans la physiocratie. Nous sommes davantage dans la situation d'un capitalisme d'État protectionniste que dans un véritable capitalisme d'essence libérale.
16. The Economist, The sun sets on Cotonou, Finance & Economics, print edition, 3 Janvier 2008.

tiques de plus grande envergure et d'un plus grand suivi. A l'échelle européenne par exemple, chaque pays contribue en fonction de sa richesse globale, son Produit Intérieur Brut, au financement de projets communautaires. Les pays riches comme l'Allemagne et la France contribuent davantage que d'autres pays comme la Grèce ou l'Espagne par exemple. En outre, les pays relativement moins riches reçoivent davantage que les pays les plus riches de l'Union Européenne. Au final, à l'exception notoire de la France, les pays les moins avancés de l'Union perçoivent davantage qu'ils ne donnent. On peut considérer ainsi que les aides communautaires, les subventions agricoles, le financement de programmes de développement (culturels, éducatifs, sociaux, logistiques, etc.) accordés, entre autres, au Portugal, à l'Irlande et à la Grèce font de ces pays, en particulier des classes qui reçoivent la manne européenne (population agricole, ouvrière, etc.), les exploitants des pays riches européens (en particulier l'Allemagne)[17]. Dans ce cas, nous pouvons conclure qu'une classe exploitée d'un pays relativement pauvre peut exploiter la classe exploitante et exploitée d'un pays riche puisque c'est sur la richesse fiscale nationale qu'est prélevée la contribution à la construction européenne. La différence avec les autres situations d'exploitation est que cette dernière est « volontaire », même si on peut arguer que des décisions sont prises et seront prises à l'avenir à une majorité qualifiée comme le propose le mini-traité actuellement en cours de ratification. Cela obligera alors certains pays européens à financer des programmes dont ils ne veulent *peut-être* pas et créera dès lors une forme particulière d'exploitation.

Nous sommes conscients du caractère surprenant d'une telle argumentation qui consiste à lier exploitation et solidarité. Ceci paraît étrange dans la mesure où le principe de solidarité européenne étant à l'origine volontaire, il ne saurait être réduit à une forme d'exploitation. Pour autant, il faut bien distinguer la volonté d'un gouvernement d'un côté et les conséquences de la mise en place d'une telle politique de l'autre. Effectivement, la réallocation de ressources financières d'un poste budgétaire de politique intérieure (subventions à la reconversion de certains sites industriels, aides à l'emploi à un chômeur par exemple) vers un poste budgétaire liée à la politique extérieure (financement de projets européens par exemple) peut se faire – et se fait concrètement – au détriment d'une partie de la population nationale qui aurait reçu les subsides de son Etat dans le cas inverse. La pertinence du raisonnement se mesure alors à l'aune de la spoliation de la population. Ceci est encore le cas dans la situation suivante.

D. *La classe dominée d'un pays riche peut exploiter une autre classe dominée dans son pays d'origine, mais aussi la classe dominante de ce même pays, voire la classe dominante d'un autre pays riche.*

Nous revenons ici sur un thème qui concerne la protection sociale et qui, dans notre réflexion, prend forme sous l'aspect d'une forme d'exploitation relative. On peut, en effet, considérer que, lorsque des personnes sans activité professionnelle perçoivent des allocations chômage ou autres (R.M.I. e.g.), il s'instaure une relation d'exploitation vis-à-vis de la frange de la population salariée qui, elle, travaille et se voit grevée d'une part conséquente de ses revenus : que cette population laborieuse, selon la théorie marxiste, soit la classe dominante (elle se fait alors ponctionner par le biais des cotisations patronales) ou qu'elle soit la classe dominée (la ponction a alors lieu par le biais de cotisations salariales).

Un tel raisonnement peut être reconduit pour des personnes handicapées, en arrêt maladie, en arrêt maternité ou tout simplement à la retraite. Il peut être également reconduit pour une situation où la classe dominée d'un pays riche exploite une classe dominante d'un autre pays riche par le biais d'une redistribution fiscale favorable, etc.

La théorie de l'exploitation semble alors faire obstacle à l'idée de solidarité nationale. Ou bien c'est peut-être le concept même de justice dans sa conception redistributive qui semble lié à une forme nécessaire d'exploitation marxienne.

17. En 1999, l'Allemagne contribuait à hauteur de 21 millions d'Euros et percevait 9,7 millions d'Euros. De même le Royaume Uni contribuait à hauteur de 11 millions d'Euros et percevait 5,7 millions d'Euros. La Grèce, elle, recevait 4,9 millions d'Euros (contre une contribution de 1,3), l'Espagne 6,2 millions d'Euros (contre 12,8) et l'Irlande 1 million d'Euros (contre 2,8), etc. La France, quant à elle, contribuait pour un montant à peine supérieur au montant qu'elle percevait (autour de 12 millions d'Euros). Cf. The Economist, Europe in Figures, Pocket, fifth edition, London, 2002, pp.196-197.

> Au final, comme nous l'avons repris auparavant à Kymlicka, il est de plus en plus évident que les défavorisés et les exploités de la théorie marxienne ne sont pas toujours les mêmes. On le comprend, il ne s'agit plus simplement d'un capitaliste, d'un bourgeois qui exploite un pauvre ouvrier besogneux. Cette question (économique autant que politique) est très complexe et il ne s'agit certainement pas ici d'y répondre de manière péremptoire. Elle dépasse de loin le cadre d'un article de quelques pages. Il s'agit encore moins d'esquisser un simple avis tranché qui ferait la nécessaire économie de la prudence et de la rigueur. Nous avons simplement voulu donner quelques points qui mettaient à mal les idées reçues en la matière. Car s'il est indéniable qu'une situation d'exploitation peut apparaître ponctuellement (à l'intérieur d'une nation, entre plusieurs nations et enfin entre des classes différentes ou identiques entre différentes nations), il y a beaucoup à dire que la situation soit généralisée et linéaire (exploitation d'une classe « ouvrière » par une classe « bourgeoise »). Il s'agit donc d'une pétition de principes qui réduit de manière caricaturale une réalité qui tient davantage d'un patchwork composé de classes sociales différentes et interdépendantes que de relations sociales unidimensionnelles.

Dès lors, les marxistes ont beau affirmer que les situations B, C et D ne font que renverser la tendance, ne sont que des exceptions à une loi générale qui serait l'exploitation marxienne, *il est impossible de prétendre aisément que la liberté d'un groupe, d'une classe, dépend de l'exploitation d'un autre groupe*. La théorie de l'exploitation montre par conséquent ses limites d'un point de vue politique. C'est, en effet, une vision idéologique que de situer la liberté du côté des dominants ou du côté des dominés, si on ne peut pas établir avec certitude qui domine qui, qui exploite qui dans une économie mondialisée voire, tout simplement, dans une société où la multiplicité des classes sociales, la diversité des intérêts privés et l'existence d'un État providence forment un maelström difficilement sondable et rendent les choses largement plus complexes qu'auparavant.

3. Critique philosophique

La critique de la théorie marxienne de l'exploitation nous amène à quitter l'aspect politique de cette théorie pour nous concentrer sur son aspect philosophique que l'on peut traduire de la manière suivante : si l'exploitation d'un individu permet à certains d'obtenir ou de conserver une forme de liberté qu'ils enlèvent à d'autres, alors *est-ce que la liberté, pour être réalisée, a besoin de l'absence de liberté ?*

Si la réponse à cette question devait être positive, nous pourrions gloser sur la dimension dialectique de la liberté : pour qu'elle se réalise quelque part, il faut nécessairement qu'elle soit absente autre part. Nous serions ainsi en présence, une nouvelle fois, de l'idée hégélienne de l'unité des contraires et de l'influence du maître d'Iéna sur Marx. De philosophique, notre questionnement deviendrait métaphysique. Le problème serait alors de considérer l'apparition d'un tel phénomène : serait-il diachronique ou synchronique ? S'il s'agissait d'un phénomène diachronique, cela signifierait qu'il y a eu, dans l'histoire, des périodes de pleine liberté qui ont succédé ou précédé à d'autres périodes dont la liberté était absente. En fait, un court examen historique permet de considérer le caractère saugrenu d'une telle approche. La guerre, l'exploitation ou l'asservissement de certaines populations apparaissent comme le fil rouge de toute l'histoire humaine. Il ne semble pas y avoir eu une seule période de paix, ou de coexistence pacifique hors de tout asservissement, au niveau planétaire depuis que l'homme existe. Mais,

en fait, si la liberté est dialectique, nous sommes amenés à croire qu'un tel phénomène s'actualise de manière synchronique, à savoir que sa vraie dimension n'est pas temporelle, mais spatiale : pour n'importe quelle période, nous pouvons constater que des individus sont libres et que d'autres ne le sont pas. L'examen de cette situation nous oblige alors à revenir sur les points déjà cités qui concernent l'aspect politique et économique qu'induit une telle conception de la liberté, fille de la théorie marxienne de l'exploitation. Ce faisant, il semble très difficile de trancher au vu de ce que nous avons écrit sur ce sujet.

Pour autant, il convient de nous arrêter sur un point. La conception de la liberté qu'implique la théorie de l'exploitation – ce qui est liberté ici s'explique par l'absence de liberté ailleurs – n'est pas nouvelle. Elle semble même reprise à l'économie politique classique que Marx a tant fustigée. Effectivement, à la fin du XVIIIème siècle, A. Smith avait déjà fait part d'une conception similaire, au point de vue économique, avec sa parabole de la loterie :

> *« Dans une loterie parfaitement égale, ceux qui tirent les billets gagnants doivent gagner tout ce que perdent ceux qui tirent les billets blancs. Dans une profession où vingt personnes échouent pour une qui réussit, celle-ci doit gagner tout ce qui aurait pu être gagné par les vingt qui échouent. »*[18]

Selon la parabole smithienne, tout le monde ne peut être heureux et libre en même temps dans l'économie libérale. Certains doivent souffrir et perdre pour que d'autres gagnent et soient libres. Il y aurait donc un jeu à somme nulle qui s'appliquerait à la liberté. Toutefois, il ne s'agit pour Smith que d'un modèle économique particulier. Il n'est en rien révélateur de la conception de la liberté selon le philosophe écossais. Ce dernier utilise l'idée de la main invisible, qui correspond le mieux, nous pensons, à sa conception de la liberté qui n'est pas une conception dialectique ou métaphysique. Il s'agit plutôt d'une conception particulière selon laquelle la poursuite de l'intérêt personnel est censée bénéficier à tous en se fondant dans l'intérêt général. Bref, la liberté d'un individu, la réalisation de ses projets, la satisfaction de ses besoins n'entrent d'aucune manière en contradiction avec l'intérêt, c'est-à-dire le bonheur et la liberté, d'un autre individu. Mieux, elle lui est profitable car la liberté individuelle s'accroît lorsqu'elle est partagée par le plus grand nombre. Voilà une conception qui semble bien aux antipodes de la théorie marxienne de l'exploitation.

Toutefois, la conception de Marx n'est pas si différente de la conception smithienne. Elle s'y oppose évidemment en ce qui concerne la définition de la liberté dans la société contemporaine, i.e. dans la société séculaire capitaliste. Pourtant, si l'on considère la conception marxienne de la liberté dans la société communiste, selon son *Matérialisme historique*[19] et son *Naturalisme dialectique*[20], on constate que Marx a une conception proche de la main invisible smithienne à laquelle se résume, somme toute, la tradition libérale. Effectivement, dans la société communiste, la liberté de l'un est consubstantielle de la liberté de l'autre, la première enrichissant la seconde. Une fois bannie l'exploitation, il ne peut plus y avoir de relation qui se fasse au détriment des uns et à l'avantage des autres. La liberté d'un homme accroît celle de l'autre et réciproquement. Dans ce cas-là, il s'agit alors d'une conception bien moderne de la liberté.

18. A. Smith, Recherches sur la nature et les causes de la richesse des nations [1776], I, VIII, trad. G. Garnier [revue par A. Blanqui], GF-Flammarion, Paris, 1991, t. I, p.180.
19. « C'est seulement dans la société communiste que l'épanouissement original et libre des individus n'est pas un vain mot, car il dépend des liaisons entre les individus, liaisons qui consistent partie dans les conditions économiques, partie dans la nécessaire solidarité du libre épanouissement de tous, et enfin dans le mode d'activité universel des individus sur la base des forces productives existantes. » IA [1845] in PI III, p.1321.
20. « (…) chacun de nous [s'affirmera] doublement dans sa production, soi-même et l'autre. 1. Dans ma production, je réaliserais mon individualité, ma particularité ; j'éprouverais, en travaillant, la jouissance d'une manifestation individuelle de ma vie, et, dans la contemplation de l'objet, j'aurais la joie individuelle de reconnaître ma personnalité comme une puissance réelle, concrètement saisissable et échappant à tout doute. 2. Dans ta jouissance ou ton emploi de mon produit, j'aurais la joie spirituelle immédiate de satisfaire par mon travail un besoin humain, de réaliser la nature humaine et de fournir au besoin d'un autre l'objet de sa nécessité. 3. J'aurais conscience de servir de médiateur entre toi et le genre humain, d'être reconnu et ressenti par toi comme un complément à ton propre être et comme une partie nécessaire de toi-même, d'être accepté dans ton esprit comme dans ton amour. 4. J'aurais, dans mes manifestations individuelles, la joie de créer la manifestation de ta vie, c'est-à-dire de réaliser et d'affirmer dans mon activité individuelle ma vraie nature, ma sociabilité humaine [Gemeinwesen] » M44 [1844] in PI II, p.33.

De ce fait, concernant sa conception de la liberté, on peut considérer que Marx a d'abord repris la parabole d'un modèle économique libéral (la loterie smithienne) qu'il a adaptée aux différents modes de production (de l'Antiquité à la société bourgeoise) au travers de la théorie de l'exploitation qui est, selon lui, la trame de l'Histoire. Puis, il a proposé de résoudre, *via* la révolution et l'avènement de la société communiste, cette éternelle exploitation de l'homme par l'homme, à l'aide d'une philosophie politique téléologique dans laquelle il a réinscrit, *in fine*, la liberté dans sa conception libérale.

Si, jusqu'à présent, la liberté des uns s'est réalisée au détriment de celle des autres, si, selon Marx[21], toutes les classes exploitées dans l'Histoire ont su changer leur situation afin de devenir, à leur tour, des classes exploitantes, le prolétariat, classe exploitée de la société bourgeoise, abrogera cette loi d'airain. En effet, par son action révolutionnaire, le prolétariat ne se contentera pas d'une action égoïste à la fin de laquelle il deviendra libre en exploitant d'autres classes. En réalité, le prolétariat libèrera la société dans son ensemble. Dès lors, la liberté des uns ne se fera plus au détriment des autres mais tous les hommes seront libres en même temps[22]. Clairement, si la réalisation de la liberté a été négative jusqu'à présent dans la mesure où elle requérait, pour rendre un individu libre, la privation de liberté d'un autre individu, la société communiste donnera une acception pleinement positive à la liberté en lui donnant son sens libéral originel[23]. Nous constatons alors que nous sommes en présence d'une *transsubstantiation* : de négative, l'essence de la liberté devient positive. Tout comme l'homme auquel elle est indéfectiblement liée, la liberté change d'essence au cours de la révolution prolétarienne.

Cependant, notre réflexion nous oblige à voir au-delà de la simple répétition d'un concept métaphysique (la transsubstantiation) concomitant à la conception de l'homme chez Marx[24]. Si la liberté, comme l'homme, voit son essence se transformer automatiquement, ce qui apparaît en filigrane est le caractère nécessaire d'une liberté négative, ou exclusive car partielle, qui précèderait l'avènement d'une liberté positive, ou totale. Car, effectivement, en prétendant avoir reconnu dans l'histoire les lois d'un mouvement dialectique[25], en faisant de la détermination l'une des caractéristiques principales de l'Histoire, la démonstration de Marx se résume tout simplement au fait que l'exploitation à l'oeuvre dans le monde est inévitable afin de pouvoir être dépassée. Dès lors, il semble exister, pour le philosophe allemand, une nécessité et une justification à la souffrance des hommes.

«*Son déclin [celui de la bourgeoisie] et le triomphe du prolétariat sont également inévitables.*»[26]

Dans la perspective inéluctable de la société communiste, le *Matérialisme historique* nous amène ainsi à considérer la théorie de l'exploitation sous l'angle métaphysique d'une *qualité de la souffrance*, d'une valorisation de la misère humaine par Marx : ceux qui sont exploités ne souffrent pas en vain car ils finiront par vaincre tous les obstacles sur leurs routes avant de connaître le bonheur, c'est-à-dire l'abondance matérielle et spirituelle, dans la société communiste.

La qualité de la souffrance est un concept purement chrétien[27]. En appliquant ce concept au phénomène politique, Marx conserve l'idée de destin et y adjoint l'idée que cette souffrance est provoquée par

21. « Toutes les classes qui s'assurèrent autrefois le pouvoir ont cherché à consolider leur position déjà acquise en assujettissant la société tout entière aux conditions de leur pratique. » MPC [1848] in PI I, p.172. Selon Marx, l'exemple typique de cette situation est celui qu'offre la Révolution française qui a fait de la classe bourgeoise, ancienne classe féodale exploitée, la classe exploitante de la société capitaliste.
22. « Jusqu'à présent, toutes les révolutions ont toujours laissé intact le mode des activités ; il s'y agissait seulement d'une autre distribution de ces activités (…). En revanche, la révolution communiste, se dressant contre le mode traditionnel des activités, se débarrasse du travail et abolit la domination de toutes les classes en abolissant les classes elles-mêmes (…) » IA [1845] in PI III, p.1123 ; « L'ancienne société bourgeoise, avec ses classes et ses conflits de classes, fait place à une association où le libre épanouissement de chacun est la condition du libre épanouissement de tous. » MPC [1848] in PI I, p.183.
23. Or, c'est ce sens originel que Marx qualifie, lorsqu'il le trouve chez les économistes libéraux classiques, de bourgeois. Marx parle ainsi de « conceptions bourgeoises de liberté, de culture, de droit, etc. » MPC in PI I, p.178.
24. Voir notre étude, L'individu chez Marx.
25. « (…) le but final de cet ouvrage est de dévoiler la loi économique du mouvement de la société moderne » Préface du Cap. [1867] in PI I, p.550. Nous avons vu plus haut quand et comment Marx parle de « méthode dialectique » à son sujet (Postface à la 2nde éd. allemande du Cap., Livre I [1867] in PI I, p.556 sq.).
26. MPC [1848] in PI I, p.173. Cette phrase est reprise dans la conclusion du Cap., Livre I [1867] in PI I, p.1240.
27. Même si on peut penser que Platon avait déjà énoncé cette idée en affirmant qu'il vaut mieux subir l'injustice que la commettre (« Je suis convaincu, moi, que, toi et moi et tous les hommes, nous pensons tous que c'est un plus grand mal de commettre l'injustice que de la subir et de n'être pas puni que de l'être » Platon, Gorgias, GFFlammarion, Paris, 1967, 473d-474c, p.208), ce n'est véritablement qu'avec le Nouveau Testament, en particulier avec la Passion du Christ, que ce thème fait son irruption dans l'histoire des idées.

un autre homme. Le Christ, qui distinguait l'intériorité religieuse de l'extériorité politique[28], n'était pas allé si loin. Marx, lui, ajoute l'idée d'exploitation, à savoir que la liberté d'une personne est conditionnée par l'asservissement d'une autre, et inversement que la souffrance d'une personne est le produit de facteurs exogènes à sa propre personne et endogènes à la volonté d'une autre. Pour un malheur personnel, il y a toujours la jouissance d'un autre, une *schadenfreude*[29] exacerbée.

Ce faisant, cette qualité de la souffrance pose deux problèmes. Le premier problème est un problème métaphysique qui devient rapidement un problème religieux. En donnant à la question « pourquoi les hommes souffrent-ils ? » une réponse unique, une perspective monolithique *via* la révolution prolétarienne, Marx ne se pose guère en philosophe, mais plutôt en une forme de prophète. Réduire la diversité des situations d'exploitation à une exploitation uniforme, donner à la mosaïque des souffrances humaines une seule explication, ce n'est pas faire acte de pensée. La justification de cette souffrance – par la mise en avant de la lutte des classes, de la théorie de l'exploitation et de la perspective émancipatrice et expiatoire de la société communiste – semble rigide et simpliste.

Le second problème que pose cette qualité de la souffrance est un problème quasiment insoluble : «comment mesure-t-on la souffrance individuelle et comment la compare-t-on objectivement ?» De même que, selon Descartes, la raison est inégalement partagée par tous bien que chacun s'en croit suffisamment pourvu pour juger, de même la souffrance est « vécue », « expérimentée » par tous, bien que de manière inégale. Pourtant, il est indiscutable que certains souffrent plus que d'autres, comme il est indiscutable que certains utilisent mieux leur raison que d'autres. Car l'égalité ne peut se fonder que sur un rapport : si nous sommes tous égaux au vu de nos droits fondamentaux, nous ne sommes égaux que sous des rapports particuliers de la vie quotidienne : âge, sexe, emploi, origines, etc. On ne peut donc rendre compte de « l'égalité » de la souffrance, *i.e.* de sa comparaison et de sa mesure, qu'avec de grandes précautions. Faut-il prendre une échelle sociale et comparer la misère des gens au vu de leur situation de classe (origines sociales), de leur situation personnelle (environnement familial), etc. ? Faut-il comparer la souffrance endurée dans les pays riches (chômage, exclusion, handicaps de toutes sortes) avec celle endurée dans les pays pauvres (famines, maladies, corruption et exactions de l'État, conflits ethniques, etc.) ? A ceci, vient s'ajouter l'échelle temporelle : est-il possible, voire simplement pertinent, de comparer l'exploitation ou la souffrance présente avec celle que pouvaient connaître nos prédécesseurs ? De telles comparaisons, de telles mesures de la souffrance sont difficiles à mener. Affirmer que ceux qui sont exploités sont ceux qui souffrent et, inversement, que ceux qui souffrent sont ceux qui sont exploités est, par conséquent, de nature simpliste. Ajouter que ceux qui souffrent et sont exploités ont raison devient dangereux parce que cela ne peut reposer sur aucun fait avéré (il y a toujours quelqu'un qui souffre davantage) ou rationnel (cela est difficilement de l'ordre d'une démonstration). Bref, une telle affirmation justifie toutes les violences. La misère n'est ni une élection, ni une épreuve méritant automatiquement récompense dans un autre monde (la Jérusalem céleste ou la société communiste) où s'exercera une vengeance terrifiante (l'Apocalypse ou la Révolution) qui précédera le bonheur universel d'êtres humains débonnaires.

Ainsi pouvons-nous comprendre sans difficulté que cette qualité de la souffrance soit à l'origine du succès d'une telle idéologie politique auprès des masses populaires. Les marxistes ne s'y sont pas trompés qui ont su reprendre et développer cette idée pour accroître leur influence. A l'exemple de Marx, ils ont fait de l'épisode révolutionnaire français l'archétype même de l'évènement qui devait justifier l'espoir placé – lire la foi – dans leur idéologie. Si des hommes exploités avaient su, un jour, abolir leur situation d'esclaves pour devenir libres, alors cela pouvait et devait être répété ailleurs. En guise d'illustration, il est aisé de rappeler qu'autour de 1830-1840, il y eut, en Europe, une véritable conversion à la Révolution française parmi les intellectuels. Marx, qui n'avait pas initié le mouvement, sut en recueillir les fruits. Pareillement, un siècle plus tard, la Révolution russe, présentée comme une révolution « prolétarienne », déclencha une vague de sympathie dans les populations ouvrières et intellectuelles entre les deux guerres mondiales (même si cela dura moins longtemps). Ainsi, semble-t-il que, pendant une trop longue période, pour être « bon », il fallut être révolutionnaire.

28. « Rendez à César ce qui est à César et à Dieu, ce qui est à Dieu. » Nouveau Testament (Version Jérusalem), Evangile selon Saint Luc, §20, 35.
29. Schopenhauer décrivait ce sentiment comme la « réjouissance méchante » : elle consiste à se réjouir, à tirer jouissance du malheur d'autrui.

Au fil des ans, cette qualité de la souffrance appliquée à l'idéologie politique marxienne est devenue le fond de commerce de toutes les démagogies, puis la vulgate de tous les totalitarismes. Ces totalitarismes, qu'ils soient de « gauche » ou de « droite », s'appuient, de manière équivalente, sur la souffrance ressentie et partagée par une part de la population (communauté ouvrière, raciale, religieuse, etc.) afin de légitimer l'usage de la force et le non respect des règles démocratiques. La communauté à laquelle ils s'adressent est une communauté qui se sent, avec plus ou moins de raisons, exploitée, abâtardie, bref, en souffrance. Pour gagner à leur cause le plus grand nombre, ces totalitarismes partent ainsi du postulat que ceux qui souffrent sont ceux qui ont raison. Cet appel à la souffrance motive et excuse donc l'action violente, source de nouvelles souffrances, celles-ci voulues et excusées car rendues nécessaires par la réalisation d'une société plus juste, d'un « monde meilleur ».

Dès lors, la présence de cette qualité de la souffrance chez Marx nous interpelle. Considérer l'analogie qui fait de Marx, ré-inventant le communisme, l'ultime prophète d'Israël, et du prolétariat, le nouveau peuple élu[30], ne manque pas de pertinence. Si Marx a pu souhaiter que sa philosophie ne soit pas confondue avec une religion, s'il a pu écrire qu'il souhaitait favoriser la discussion, la critique, etc., il reste que sa philosophie postérieure à *L'Idéologie allemande* trouve à son origine des postulats qui ne laissent que peu de doutes sur le caractère religieux de sa philosophie[31]. Les oeuvres philosophiques postérieures à 1844 ne sont pas entièrement convaincantes quand il s'agit d'expliquer pourquoi et comment elles voient le monde, l'homme et la société qui l'a vu naître, c'est-à-dire pour quelles raisons l'homme souffre et quels sont les remèdes que l'Histoire va lui apporter. Marx se contente alors d'énoncer une vérité à laquelle il faut croire, comme l'a souligné Claude Lefort[32]. On peut alors parler de révélation pour Marx et d'une véritable métaphysique religieuse puisque peu de choses sont vérifiables (l'exploitation, la lutte des classes, l'inéluctabilité de la société communiste à venir, etc.). Comme Marx l'a écrit lui-même dans le *Manifeste*, « rien n'est plus facile que de poser un vernis socialiste sur l'ascétisme chrétien »[33].

Partant, le « Maure » n'a fait que développer l'idée d'une qualité de la souffrance en l'enveloppant dans une philosophie politique qui la travestit. Le philosophe allemand est donc le premier penseur politique chrétien, comme l'a écrit M. Henry.

30. M. Henry, Marx, t.I : Une philosophie de la réalité, Tel, Gallimard, Paris, p.144.
31. Au-delà de la pensée marxienne, nous pouvons nous arrêter sur les traits communs et obvies qu'a pu revêtir, au cours du XXème siècle, le marxisme avec la pratique du christianisme : hiérarchie pyramidale des différents partis communistes, stricte discipline (et leur pendant avec les nombreuses excommunications), pratique du secret, rituels inamovibles et indiscutables, ainsi qu'intériorisation de ces valeurs sacrées que sont le dévouement, la fidélité aux dogmes, l'adoration des premiers « Pères » avec ses Saints laïques (Bebel, Lénine, Kautsky, etc.), etc.
32. C. Lefort, Essais sur le politique, XIXè-XXè siècles, Seuil, Points, Essais, Paris, 1986, pp.195-212.
33. MPC in PI I, p.184. Si ce passage fait directement référence au socialisme féodal et non pas au socialisme scientifique que prétend promouvoir Marx, M. Rubel cite un extrait d'un des derniers discours publics de Marx dans lequel il semble que le philosophe allemand fut lui-même conscient du parallèle : « L'ouvrier doit saisir un jour la suprématie politique pour asseoir la nouvelle organisation du travail ; il doit renverser la vieille politique soutenant les vieilles institutions, sous peine, comme les anciens chrétiens qui l'avaient négligé et dédaigné, de ne voir jamais leur royaume de ce monde » (Discours de La Haye [1872], cité par M. Rubel in PI IV, p.1698). On retrouve la même inspiration dans un travail rédigé par Engels, Contribution à l'histoire du christianisme primitif : « Tous les deux, christianisme et socialisme ouvrier, prêchent une délivrance imminente de l'esclavage et de la misère ; le christianisme situe cette délivrance dans l'au-delà d'une vie après la mort, dans le Ciel, le socialisme dans ce monde-ci, dans une transformation de la société. L'un et l'autre sont poursuivis et traqués, leurs adeptes sont proscrits, soumis à des lois d'exception, les uns comme ennemis du genre humain, les autres comme ennemis du Reich, ennemis de la religion, de la famille, de l'ordre social. Et malgré toutes les persécutions, et même directement servis par elles, l'un et l'autre se fraient victorieusement leur chemin. Trois siècles après sa naissance, le christianisme est reconnu comme religion d'État de l'Empire romain, et en moins de soixante ans le socialisme a conquis une position telle que sa victoire est absolument assurée » (Werke 22, p.449). Cité par M. Rubel in PI IV, p.1697. Il faut croire que le caractère religieux de la philosophie marxienne parût bien assez tôt puisque Proudhon, en réponse à la demande (lettre du 17 mai 1846 in Correspondance, Tome I) de Marx de faire partie d'une correspondance communiste, exprima des réserves sur le dogmatisme de Marx : « Ne nous faisons pas les chefs d'une nouvelle intolérance, ne nous posons pas en apôtres d'une nouvelle religion, cette religion fût-elle la religion de la logique, la religion de la raison. » Cf. Rubel, Chronologie in PI I, p.LXIX.

Bibliographie partielle

ALTHUSSER Louis, *Pour Marx*, Paris: Maspero, 1965.

ARON R. Philosophie critique de l'histoire. Paris: Julliard, 1987.

ARON Raymond. *Le Marxisme de Marx.* Paris: de Fallois, 2002.

BALIBAR Etienne, *Politique et philosophie avant et après Marx,* Galilée, Paris, 1997.

BALIBAR Etienne, *La philosophie de Marx*, La découverte – Repères, 1993.

BENSUSSAN Gérard – LABICA Georges, *Dictionnaire critique du marxisme,* (sous la direction de), Paris, Quadrige / PUF, 1999.

BIDET Jacques – KOUVELAKIS Eustache, *Dictionnaire Marx Contemporain*, (sous la direction de), Actuel Marx Confrontation, PUF, Paris, 2001.

BIDET Jacques, 'Manifeste du Parti Communiste', in *Etudier Marx* (Paris: CNRS), 1985.

BIDET Jacques, *Théorie de la modernité* suivie de *Marx et le marché*, PUF, collection Questions, Paris, 1990.

CAILLEBA Patrice. *L'individu chez Marx*. Paris: ANRT, 2005.

HABERMAS Jürgen, *Après Marx*, Traduit de l'allemand par J.R. Ladmiral et M.B. de Launay, Paris, 1997.

HENRY Michel. *Marx, tome I : Une philosophie de la réalité* ; *Marx, tome II : Une philosophie de l'économie*. Paris: Gallimard, 1976.

MARX Karl. *OEuvres : Economie I*. 2a ed. Paris: Gallimard – La Pléiade, 1977

MARX Karl. *OEuvres : Economie II*. Paris: Gallimard – La Pléiade, 1968

MARX Karl. *OEuvres : Philosophie*. Paris: Gallimard – La Pléiade, 1982

MARX Karl. *OEuvres : Politique I*. Paris: Gallimard – La Pléiade, 1994

MARX Karl – ENGELS, Friedrich. *Correspondance – Tome I : novembre 1835 – décembre 1848*. Paris: Editions Sociales, 1971.

Les abréviations

Pour la plupart des ouvrages utilisés, un certain nombre d'abréviations ont été retenues.

OEuvres de Marx	**Abréviations**
Thèse de doctorat sur la Différence de la philosophie naturelle chez Démocrite et chez Epicure	Thèse
Critique de la philosophie du droit de Hegel	M43
A propos de la Question juive (Annales franco-allemandes)	QJ
Introduction à la Critique de la philosophie du droit de Hegel (Annales francoallemandes)	I-M43
Manuscrits de 1844	M44
La Sainte Famille	LSF
Articles pour le Vorwärts !	Vorwärts !
Thèses sur Feuerbach	TsF
L'Idéologie allemande	IA
Misère de la Philosophie	MPh
Manifeste du Parti Communiste	MPC
Les luttes des classes en France 1848-1850	LdC
Le 18 Brumaire de Louis Bonaparte	18B
Introduction à la Critique de l'Economie Politique	I-CEP
Principes d'une Critique de l'économie Politique	Grundrisse...
Critique de l'Economie Politique	CEP
Tous les opuscules où il est fait référence en titre à l'Association Internationale des Travailleurs	AIT
Le Capital	Le Cap.
La Guerre Civile en France	GCF
Critique du programme du parti ouvrier allemand	Critique du Gotha

Les citations sans nom d'auteur concernent toutes Marx. Les textes cités dans l'édition Rubel de la Pléiade sont notés PI suivi du numéro de volume (de I à IV en chiffre romain). De même, pour tous les extraits de sa riche correspondance, nous nous sommes contentés d'indiquer Correspondance, puis le Tome.

The Role of Formal and Informal Chinese Institutions in Shaping the Entrepreneurial Landscape

Par Jean-Michel Quentier
IRMAPE
Ecole Supérieure de Commerce de Pau (ESCPAU) France

Abstract

This paper views entrepreneurship in China as a legitimacy seeking process. We use institutional theory as a lens to understand the pattern of entrepreneurship in China. We also examine how rapidly changing formal and informal institutions are likely to bring change in entrepreneurship in the country.

Keywords

China, institutions, entrepreneurship, legitimacy

Introduction

While the Chinese Communist Party (CCP) maintained political dominance thanks mainly to entrepreneurs that were highly innovative in carrying out state plans to build industries and launch new enterprises (Yang, 2002), "reemergence" of private entrepreneurship (Guiheux, 2006) in the country is a recent phenomenon. In recent years, indeed, there has been a rapid growth of entrepreneurship in China. As of 2005, China had about 24 million small independent companies and the number is growing at 15-20% annually (Loyalka, 2006). Small to medium sized enterprises (SMEs) account for 75% of new jobs (Loyalka, 2006). For most of these enterprises, entrepreneurship is considered to play a very important role (Daokui Li *et al.*, 2006).

International differences in entrepreneurship are typically framed as a result of cultural differences across countries. The existing studies on entrepreneurship thus narrowly focus on culture (Busenitz *et al.*, 2000) and a number of them have linked Hofstede's (1980) cultural dimensions -especially individualism-to examine a country's propensity to engage in entrepreneurial activities.

The findings of studies focusing on individualism are, however, inconsistent and confusing. While some commentators argued that collectivism is negatively related with entrepreneurial activities (Takyi-Asiedu, 1993), others have found a limited correlation between countries' levels of individualism and entrepreneurship (Acs 1992; European Network for SME Research 1996; Mueller and Thomas, 1997). Similarly, Harper (2003) rejected the idea that people in individualist cultures are more entrepreneurial than those in collectivist cultures. He, however, proposed that developing countries characterized by collective culture may have different social organizations that may impact on entrepreneurship patterns (Harper, 2003). Likewise, Holt (1997) found that Chinese and U.S. entrepreneurs did not differ significantly in terms of their emphasis on individualism and acceptance of uncertainty (Hofstede, 1980). It can thus be concluded that private entrepreneurship either encourages individualism and acceptance of uncertainty or attracts people with values comparable to entrepreneurs (Morris and Schindehutte, 2005).

Other researchers point to declining differences in development across cultures (Baumol, 1986) and global diffusion of entrepreneurial institutions (Gereffi and Hempel, 1996) such as IPO markets in countries considered to be characterized by nonentrepreneurial cultures (Edmundson *et al* 1996). These studies provide support for the notion that Hofstede's (1980) measures of culture are far from sufficient to describe cross-country differences in entrepreneurial activity. In China's case, close state control can be arguably attributed to the failure of apparently-abundant Chinese entrepreneurship (Moore, 1997). These examples imply that in addition to cultural differences, entrepreneurship is determined by a number of other forces.

Perspectives on entrepreneurship in the social sciences include focus on the role of economic, political and legal institutions in facilitating or hindering entrepreneurship (Djankov and Roland, 2006). Referring to China, Guiheux (2006) argues that entrepreneurship in the country can be attributed to initiatives coming from society (informal institutions) and the setting up of a new legal framework (formal institutions). To broaden our understanding of international similarities and differences in entrepreneurship, researchers have emphasized the necessity to study a broader set of institutions (Busenitz, *et al.*, 2000) and the complex interaction between different kinds of entrepreneurs and the institutional environment in which they are embedded (Yang, 2002).

Clearly, thus there are under explored issues on how institutions with opposing and different perspectives are shaping entrepreneurship in China. The purpose of our study is to fill this research void. To more fully understand factors driving Chinese entrepreneurship landscape, in this article we integrate and apply findings in literatures on institutional theory. We propose a framework, which identifies very clear contexts and attendant mechanisms associated with institutional influence on entrepreneurship in China.

In the remainder of the paper, we first briefly review the theoretical foundation. Then, we translate the theories within the context and limits of China and attempt to explain the nature of Chinese institutions that influence the country's response to private entrepreneurship with some propositions. The final section provides conclusion and implications.

1. Theoretical Framework

1.1 - Entrepreneurship as a legitimacy seeking process

It is important to note that institutional theory is described as "a theory of legitimacy seeking" (Dickson *et al.*, 2004, p. 81). Isomorphism is arguably positively related to legitimacy (Deephouse, 1996). In an attempt to gain legitimacy, organizations thus adopt certain behaviors irrespective of whether such behaviors increase organizational efficiency (Campbell 2004; George *et al.*, 2006).

Organizations that are able to acquire legitimacy from external institutional actors, on the other hand, are likely to gain resources as well as maintain control over the environment (George *et al.*, 2006). Put differently, an organization can increase its chance of survival and/or growth by engaging in actions that are approved by powerful institutional actors, used by organizations that are perceived to be 'successful' (Newman, 2000), or have the backing and approval of professions in their industry or field (Aldrich 1999; Baum and Oliver 1991; Meyer and Scott 1983; Ruef and Scott 1998; Sitkin and Sutcliffe, 1991). Non-isomorphic responses that deviate away from "established structures, practices, and utterances of other actors in the environment" (George *et al.*, 2006), on the other hand, are likely to face resistance.

Institutional influence on entrepreneurship becomes an admittedly complex process (Dickson *et al.*, 2004), when organizations have to derive legitimacy from multiple sources. In China's context, entrepreneurs have to acquire legitimacy from the state regulations, local authorities and bureaucrats, business partners, employee (Yang, 2002) and the society.

1.2 - Scott's institutional pillars and entrepreneurship in China

Table 1 presents institutional influence on entrepreneurship in China in terms of three institutional pillars proposed by Scott (1995, 2001): regulative, normative and cognitive. Regulative institutions consist of "explicit regulative processes: rule setting, monitoring, and sanctioning activities" (Scott, 1995). They are related to regulatory bodies and the existing laws and rules that influence offshore outsourcing. Normative components introduce "a prescriptive, evaluative, and obligatory dimension into social life" (Scott 1995). Practices that are consistent with and take into account different assumptions and value systems of national cultures are likely to be successful (Schneider, 1999). Cognitive institutions represent culturally supported habits that influence the pattern of entrepreneurship. The nature of entrepreneurship is a function of a cognitive institutions related to history, culture, ideology, and social attitudes (Doucouliagos, 1995). In most cases, they are associated with cognitive legitimacy concerns that are based on subconsciously accepted rules and customs as well as some taken-for-granted cultural account related to entrepreneurship (Berger and Luckmann, 1967). Cognitive programs affect the way people notice, categorize, and interpret stimuli from the environment.

Institutional component	Institutional forces	Effect on entrepreneurship in China
Regulative	Entrepreneurial ventures perceived as potential threats to the CCP regime (Yang 2002).	Resistance to substantive actions to promote entrepreneurship and even some regressive changes.
	Ineffective enforcement of laws related o private property (Yang 2002).	Private entrepreneurs depend to a large extent on informal norms and networks for security.
	Marriage between entrepreneurship and party membership.	Substantive measures to encourage private entrepreneurship.
Normative	Entrepreneurs were considered as "getihu" or “cadres" (Hsu 2006; Mourdoukoutas 2004).	Private entrepreneurship lacked social legitimacy.
	Society's attitude toward business increasingly favoring entrepreneurship, especially related to high technologies (Hsu 2006; Han and Baumgarte 2000).	Increased social legitimacy to private entrepreneurship.
Cognitive	Culture of complacency, conformity and risk aversion (Mourdoukoutas 2004).	Hinders discovery and exploitation of new market opportunities. The Western level of risk is unacceptable in China (Harwit 2002).
	Entrepreneurs are sensitive to the communist regime and the society, which resist ownership of private property (Economist 2006).	Discourages entrepreneurship.
	Perceived “glass ceiling” for ethnic Chinese employees at multinationals (Browne 2004).	They prefer starting their own businesses.
	Rapid increase in the returns of overseas Chinese.	Emergence of a club culture promoting innovation, and risk taking (Wang 2001).

Table 1: Understanding institutional influence on the Chinese pattern of entrepreneurship

1.3 - Diffusion of entrepreneurship among the CCP

Institutionalists use the concept of diffusion to refer to the spread of institutional principles or practices among a population of actors (David and Foray, 1994; Strang and Meyer, 1993). Prior researchers have also noted typical patterns of events and relationships in institutionalization and institutional changes (Lawrence *et al.*, 2001). New values, norms, practices and ideas are first recognized, and then accepted by relatively few actors and then widely diffused (Leblebici *et al.*, 1994; Meyer and Rowan, 1977). Finally, they reach to a stage of saturation and complete legitimation (Zucker, 1987). In this regard, it is important to note that, although the Chinese Communist Party's (CCP) policies and formal legal institutions regarding private entrepreneurship have gone through stages: strict prohibition, tolerance, accommodation, and encouragement (Peng, 2004), there has been a lack of complete legitimation to entrepreneurship. For instance, property rights in China aren't yet well-defined and fully binding (Mourdoukoutas, 2004). Private enterprises are thus never sure which resources are their property and under their command and for what length of time (Mourdoukoutas, 2004).

1.4 - Progressive and regressive changes

The work of Paul D. Bush (1983, 1987, 1989, 1994), among others, provides us with additional insights into the temporal dimension of institutional changes. According to this approach, the process of institutional adjustment is broken down into two phases: Phase I involves ceremonial encapsulation, and Phase II involves regressive or progressive changes. These phases are conceptually similar to what North (1990) refers as the “cause-effect-cause" model.

In the first phase, the ideas of entrepreneurship and private ownership are ‘encapsulated’ within the existing value structure (Bush, 1994). In this phase, such decisions are mainly based on technical grounds (Tolbert and Zucker, 1983). This stage involves no change in the value structure of the community. The idea of entrepreneurship is introduced without disrupting the patterns of power, status, and other forms of existing differential advantages (Bush, 1994).

According to North’s (1993) "shared mental models", institutional evolution is influenced by the feedback process by which human beings perceive and react to changes in their environment. Phase II is a result of reaction to entrepreneurship and starts when entrepreneurship demonstrate new instrumentally warranted possibilities for correlating behaviors, which are not sanctioned by the community's traditional pattern of values (Bush, 1994). While decisions in the early phase are based on technical grounds, legitimacy pressures play a critical role in the latter phase (Tolbert and Zucker, 1983). That is, organizations must strive to achieve social legitimacy in addition to technological and operational efficiency to prosper in their environment (Abernathy and Chua, 1996).

In a progressive change, the new instrumentally warranted patterns of behavior displace ceremonially warranted patterns of behavior (Bush 1994). The sustainability of progressive institutional changes requires "minimal dislocation". Put differently, the incorporation of the new instrumental behavior has a minimally disruptive effect on other instrumental patterns of behavior in the community (Bush 1994).

An unintended consequence of entrepreneurship is disruption in existing power structure. In a regressive change, new instrumentally warranted behavior is suppressed and additional patterns of ceremonially warranted behavior are instituted to secure the suppression (Bush, 1994).

1.5 - Institutional changes crested by entrepreneurship

Entrepreneurs may also bring institutional changes as they stimulate new isomorphic pressures, modify the nature of existing pressures or replace one type of pressure by another. Structural changes institutions can be explained in terms of coercive, normative and mimetic isomorphism (DiMaggio and Powell, 1983). Scott (1995, 2001) describes these forces in terms of regulative, normative and cognitive processes respectively (Table 1).

2. Institutions and Entrepreneurship in China

2.1 - Progressive and regressive changes

The conservative faction in the Chinese Communist Party considers entrepreneurial ventures as potential threats to the party's dominance, ideology, administrative authority and moral standards (Yang, 2002). The leftist opposition leaders have thus employed China's rising income gap and increasing social unrest to criticize and justify measures against private entrepreneurship (Kahn, 2006). Some analysts argue that the delay in granting full rights to private entrepreneurs reflects "ideological rigidity and institutional inertia against changes" (Peng, 2004).

Chinese leftist leaders thus perceive improved legal institutions as potential challenges for legitimacy to the CCP regime (Potter 2004). Consequently, Chinese legal institutions related to entrepreneurship have been victims of political ideology (Yang, 2002). Following the Tiananmen events in 1989, the conservative faction's actions severely impacted private entrepreneurs. Estimates suggest that the number of private enterprises reduced by 50% that year (Ling, 1998).

We can, however, argue that institutional actors bringing regressive changes in Chinese institutions will weaken over time. Why might this be the case? First, as noted above, although many Chinese government officials and policymakers consider China's integration with the global market associated with significant socioeconomic costs, they cannot openly reject global integration (Heer, 2000). To gain legitimacy from international institutions such as the World Trade Organization (WTO), China is required to respect private entrepreneurship and ownership of private property.

Second, entrepreneurs are being openly accepted into the CCP's inner circle. The CCP's policies and formal legal institutions encourage entrepreneurship (Peng, 2004). The CCP in 2002 changed its bylaws to allow entrepreneurs to become members (Loyalka, 2006). In a speech on July 1, 2001, during celebrations of the party's 80th anniversary, President Jiang Zemin acknowledged the benefit that capitalists bring to the economy (Hoogewerf, 2002). He also handed party membership to a capitalist and one of China's most respected private companies and the first private company to list on a foreign stock exchange (Pomfret, 2001). In another instance, in January 2003, the CCP appointed Yin Mingshan, one of China's wealthiest private entrepreneurs, as deputy chairman of an advisory body to the government of Chongqing municipality, the first private businessman in China to be awarded such a high position (The Economist March 29, 2003). Although some analysts argue that the seemingly impressive position carried "no real power", optimists argue that these entrepreneurs will give the private sector a more powerful voice in policymaking (The Economist March 29, 2003) and further weaken forces contributing to regressive changes in institutions influencing private entrepreneurship. In line with these arguments, the following propositions are presented:

> ***Proposition 1a:*** *In China, forces contributing to progressive changes in institutions related to entrepreneurship and private property will strengthen over time.*
>
> ***Proposition 1b:*** *In China, forces contributing to regressive changes in institutions related to entrepreneurship and private property will weaken over time.*

2.2 - Development of entrepreneurship and substantiveness in laws protecting private property: a progressive change

Chinese communities arguably have a greater cultural disposition toward entrepreneurship (Waldinger *et al.*, 1990). Traditionally, regulative institutional such as insecurity of property rights and close state control hindered entrepreneurship in China (Djankov and Roland, 2006; Moore, 1997). Some commentators argue that the traditional culture of entrepreneurship will strengthen efforts to create the rule of law in China (Shuli, 2004).

A good example to illustrate how Chinese entrepreneurs are influencing regulative institutions related to entrepreneurship is the changes in laws related to intellectual property rights (IPR) in recent years. With rapid increase in the creation of IP by local firms, these firms are actively participating in substantive measures that could help strengthen the country's IP regime. According to the Chinese Supreme Court, in 2005, over 16,000 civil cases and 3,500 criminal cases related to IP rights violations were handled by Chinese courts and more than 2,900 people were jailed (Culpan, 2006). The number of cases involving IPR protection including patents, trade secrets and counterfeit goods increased by 21percent in 2005 (AFX News Limited, 2006). It is also important to note that 95 percent of China's IPR related cases in 2005 were brought by Chinese companies (Culpan, 2006).

The Chinese nano technology industry provides a visible example to illustrate how local IP creation leads to substantive actions to protect IPR. Chinese scientists are capable of producing carbon nanotubes 60 times faster than their U.S. counterparts (Stokes, 2005). The Nanometer Technology Center established in Beijing is actively involved in protecting IPR (Singer *et al.*, 2005).

A rapid increase in domestic IP creation has led to more substantive measures to protect IPR. Under China's recently enacted new piracy laws, buyers of pirated goods can be fined 5-10 times the value of the goods and manufacturers face jail time and equipment confiscation (Kanellos, 2002). The government has provided a significant empowerment to regulatory agencies involved in IPR issues such as the State Administration of Industry and Commerce, the State Administration of Press and Publications, the intellectual property right office and the State Pharmaceutical Administration (Yang, 2002). Similarly, China announced its plans to open special centers in 50 cities by 2006 to handle IPR infringement complaints as well as to provide consulting services (MacLeod, 2006). In sum, we argue that:

> ***Proposition 2:*** *The development of entrepreneurship will provide pressure to enhance regulative institutions related to private property and entrepreneurship.*

2.3 - Formal and informal institutions and implementation capabilities

Prior research has indicated that the degree to which ideas such as private ownership are translated into local practices is a function of implementation capacities (Scott *et al.*, 2000). The relative strength of state institutions in implementing and enforcing laws related to entrepreneurship, thus, can make or break governments' policies and firms' innovation strategies (Spencer *et al.*, 2005).

Implementation capabilities are largely influenced by informal institutions which do not necessarily change at the same rate as formal institutions. North (1990) noted that "although formal rules may change overnight as the result of political and judicial decisions, informal constraints embodied in customs, traditions, and codes of conduct are much more impervious to deliberate policies". In the Chinese entre-

preneurship landscape, there has been a rapid shift in formal institutions related to entrepreneurship. Following the 1978 economic and political reforms, China enacted thousands of new laws to protect private property and IP (Pei 1998; Meredith, 2003); and abolished or amended many laws in these areas to comply with World Trade Organization (WTO) obligations (Hughes, 2005).

Nonetheless, informal institutions are not changing at the same rate as formal institutions. Mao arguably developed a critique of capitalism, private property, and inequality as well as material interest (Guiheux, 2006). During the Mao era, private entrepreneurship was virtually eradicated and was a political taboo in China (Loyalka 2006; Peng, 2004). Entrepreneurship was "shunned" in China as late as the 1980s (Wehrfritz and Seno, 2003). After decades of socialism, the idea of respecting the constitutional rights of private entrepreneurs has been slow to diffuse among various institutional actors in China (Nee, 1989).

Chinese incubators thus lack proper mindsets in assisting and guiding private entrepreneurs (Harwit, 2002). Private enterprises also complain about difficulties in dealing with state-owned banks for loans and other state agencies as well as harassment and extortion by local cadres, tax officers and other government officials (The Economist 2002; Yang, 2002).

As discussed above, until a few years ago, when private entrepreneurs' rights were violated, they lacked legal protection (Yang, 2002). The situation is, however, changing rapidly. Like in a number of other Asian economies, China is "shifting from top-down, state-directed technology policies to more flexible, market-oriented approaches that foster innovation and entrepreneurship" (Segal, 2004) and is adopting policies that actively encourage entrepreneurship (Schramm, 2004).

Among many examples that illustrate such a trend, let us consider some. Sender (2000) documents a story related to state-run China Telecom's complaint against two entrepreneur brothers who started offering callback services in Fujian as an alternative to China Telecom's monopoly and high charges. The courts weren't convinced the brothers had violated any laws and ruled against China Telecom. Recently, Pfizer successfully went against a major Chinese ministry-level government agency to defend its Viagra patent (Boswell and Baker, 2006). Across these two examples we see how state-owned monopolies' and government agencies' control and power are declining. It is proposed:

> ***Proposition 3:*** *Over time, informal institutions and enforcement capabilities will change to catch up with the formal institutions.*

2.4 - Shift from double entrepreneurship towards legal entrepreneurship

Institutional boundaries for economic activities are not well defined in emerging economies such as China. Exploitation the regulative uncertainty and the weak rules of laws has arguably become an important form of entrepreneurship in China (Kolko, 1997). Yang (2002) refers this phenomenon prevalent in many developing economies as "double entrepreneurship" which entails maximizing economic rewards and minimizing sociopolitical risks. In a rapidly changing environment like that of China, entrepreneurs find attractive economic niches from outside the current institutional boundaries (Yang, 2002). For instance, entrepreneurs depend on relations with government bureaucrats to obtain a license to enter and remain in a business (Mourdoukoutas, 2004). At the same time, because of ineffective legal enforcement of private property rights, they have to acquire political and administrative protection or depend upon informal norms for security (Yang, 2002).

In many developing countries, starting a business entails overcoming a significant amount of red tape (Schramm, 2004). In China, one way to overcome bureaucratic red tape for businesspeople has been to be close to the CCP in order to gain advantages and preferential treatments (Guiheux, 2006). A membership in CCP would give an entrepreneur easier access to loans and official protection and discourages the entry of new players in the market (Guiheux, 2006). In addition, entrepreneurs also spend time and energy in forming and maintaining "guanxi" (networks) and cultivating ties with officials (lobbying) through gifts or bribery (Yang, 2002). In China, the factors discussed above limit an entrepreneur's ability to pursue genuine ideas and business opportunities (Mourdoukoutas, 2004). In sum, whereas entrepreneurship in the West is about identifying profitable opportunities, in China, "the ability to form an alliance" with those who control the financial, physical or human resources is critical to succeed (Krug, 2004).

Institutional environment shapes private entrepreneurs' motivation to enter into politics. For instance, one study found that the probability of an entrepreneur's political participation decreases by 8% to 20% from the mean when institutional indices related to markets and market-supporting institutions improve by one standard deviation (Li *et al.*, 2006). Improvement in market-supporting institutions or transformation of a socialist economy into a market economy can thus be an important force in converting double entrepreneurship into legal entrepreneurship (Yang, 2002). In recent years, Chinese regulative landscape has undergone significant improvement in rule setting and monitoring activities. Such an environment is expected to promote legal entrepreneurship. Based on above discussion, the following proposition is presented:

> ***Proposition 4:*** *Over time, with the development in regulative institutions, the Chinese entrepreneurship landscape will shift from double entrepreneurship to legal entrepreneurship.*

2.5 - Societal perception of entrepreneurs

The perception of entrepreneurship in China is drastically different from market economies (Mourdoukoutas, 2004). Mao arguably developed a strong critique of capitalism, private property, income and wealth inequality and material interest (Guiheux, 2006). During the Mao era, private entrepreneurship was thus virtually non-existent and was a political taboo (Loyalka, 2006; Peng, 2004). Traditionally entrepreneurship was not the most desired one for China's best and brightest and was limited to people with criminal records that found it difficult to find other jobs (Nair, 1996). Policies that were reminiscent of the Chinese Communist Public Goods Regime (Solinger, 1995) thus discouraged private entrepreneurship. Most entrepreneurs are still considered as "selfish, avaricious peddlers", or "getihu" (Hsu, 2006).

Private entrepreneurship in China is the result of the "market track" of the Chinese "dual-track approach", in which economic agents were allowed to participate in the market at free-market prices provided they fulfill their social obligations (Lau *et al.*, 2001). However, as late as the 1990s, Chinese societies had highly negative perception of those trying to build their own company (Harwit, 2002). Entrepreneurs are still considered members of the working class striving for China's development rather than risk takers (Mourdoukoutas, 2004). In an ethnographic study conducted in the Chinese city of Harbin, Hsu (2006) found some entrepreneurs were understood as "cadres" and were judged by their ability to provide socialist benefits for their employees, rather than by their success at generating profits.

Faced with such societal perceptions, China's entrepreneurs are also sensitive to the society and the communist regime that still resist ideas related to the ownership of private property (Economist, 2006). Accumulating a huge amount of wealth is thus still a "delicate subject" in China (Hoogewerf, 2002).

As noted above, although China is still characterized by a significant integration of state and social organization (Moore, 1997), attitude toward businesses and private entrepreneurship and a business career is rapidly shifting in the positive direction (Han and Baumgarte 2000, Nair, 1996). Entrepreneurs have started to get more respect in the society (Loyalka, 2006). Some Chinese leaders have also provided validity to entrepreneurship. Even Deng said: "To be rich is to be glorious" (Nair, 1996). Hsu's (2006) finding also provided support for such a trend. Educated entrepreneurs running high-tech businesses are seen as highly respected good businesspeople (Hsu, 2006). The above leads to the following:

Proposition 5a: The societal perception of private entrepreneurship will be more positive over time in China.

Proposition 5b: Private entrepreneurs in China will have a better cognitive assessment of their occupation over time.

2.6 - Inflow of overseas Chinese and the influence on entrepreneurship

"Social remittances" associated with immigrants in the form of various resources such as ideas, behaviors, identities and social capital play critical roles in promoting immigrant entrepreneurship in the home country (Levitt, 1998). In this regard, it is important to note that much of the new entrepreneurship in China can be attributed to an increasing number of overseas Chinese educated abroad that are returning home, some with significant entrepreneurship experience in industrialized world. Estimates suggest that there are about 200,000 Chinese who have returned to the country after working or studying abroad (Loyalka, 2006). Moreover, because of perceived "glass ceiling" for ethnic Chinese employees at multinationals (Browne, 2004), more and more Chinese prefer to start their own businesses.

Before proceeding further it is important note that researchers argue that it would be erroneous to assume the existence of a generic Chinese culture. There is arguably a major difference between the social organization and risk taking behavior of Chinese that have stayed in China for their whole life and Overseas Chinese (Moore, 1997).

First, consider proper Chinese. Some commentators argue that Chinese culture does not encourage independent thinking (Friedman, 2005). Compared to managers in the West, Chinese managers are more likely to avoid uncertainty, less likely to exhibit innovations and possess low degree of self-determination and risk taking (Holt, 1997, p. 490; Anderson *et al*., 2003). Chinese managers also tend to be conformists, adhering to standard rules and procedures, rather than to personal insights based on their professional experiences (Mourdoukoutas, 2004). Entrepreneurship as discovery and exploitation of market opportunities and the introduction of new products and processes arguably are incompatible with China's culture of complacency and conformity (Mourdoukoutas, 2004). Some observe that entrepreneurship in China is arguably about copying products invented and innovated in other countries (Loyalka, 2006; Mourdoukoutas, 2004).

Overseas Chinese with educations and entrepreneurship experience in the industrialized world, on the other hand, tend to be more similar to managers from the Western world. A rapid rate of returns of Chinese with education and entrepreneurial experience in the industrialized world is thus likely to promote risk taking and innovativeness in China.

Overseas Chinese returnees are likely to influence the Chinese entrepreneurship landscape through a number of mechanisms. The Overseas Chinese in Asia have earned reputation for developing complex and dense social organizations and institutions (Wu and Wu 1980; Moore, 1997). In recent years, Chinese returnees have started developing such institutions in the mainland. Wang (2001) observed the evolution of a "club culture" in China, which has stimulated interaction among various ingredients of entrepreneurship promoting innovation and risk taking. Existence of such a culture is especially evident in industrial and high-tech parks of the country (Loyalka, 2006). Overseas Chinese have contributed in producing synergies and to thicken existing Institutional (Amin and Thrift 1995; Keeble, Lawson *et al.*, 1999).

While lifelong working for big enterprises is considered to be the most sought after career in Japan (Muller *et al.*, 2004), employment in big state owned factories is ideal for many Chinese. People who have spent most of their life in such careers may not like the idea of entrepreneurship (Loyalka, 2006). Likewise, thanks to elite cultivation" in China's IT education, over half of IT graduates pursue senior technological or managerial positions after a couple of years of work rather than pursuing entrepreneurial ventures (Kharbanda and Suman, 2002). There is thus no culture to promote entrepreneurship in China. Successful entrepreneurial spin-offs from Chinese returnees may promote risk taking behavior among Chinese.

On a more speculative basis, we can argue that Chinese returnees may also change other components of Chinese institutions such as the Chinese VC landscape, which currently discourages risk taking. In China, most VC funds are linked to the government and can be considered as a loan (Harwit, 2002). Enterprises that are able to obtain VC funds feel an obligation not to lose the resources. Moreover, an incubator losing the government owned money also becomes a target of official criticism. Chinese government VC funds thus cannot accept the Western level of risk taking (Harwit, 2002). There is already evidence of a significant inflow of VC in China thanks to dense networks of overseas Chinese (UNDP, 2001). Estimates suggest that overseas Chinese control assets worth trillions of dollars. The discussion in this section is summarized as:

> ***Proposition 6:*** *The inflow of overseas Chinese into the mainland is positively related to similarity of Chinese entrepreneurship pattern with that of the Western world in terms of (a) risk taking; (b) product innovations.*

3. Discussion and implications

In this paper, we employed institutional theory to examine the pattern of Chinese entrepreneurship. Notwithstanding, the existence of a Chinese entrepreneurial culture, regulative institutions traditionally severely obstructed the growth of private entrepreneurship in China. Nonetheless, Chinese regulative institutions are undergoing fundamental and extraordinary shifts. Such shifts directly as well as indirectly through informal institutions are likely to facilitate entrepreneurship in China. From the above discussion, we can draw a number of implications.

With the change in formal institutions, profiles of entrepreneurs are likely to change. Strengthening rule of law and increased regulative certainty, for instance, are likely to encourage legal entrepreneurship and discourage double entrepreneurship. We can expect that profiles of individuals who are likely to be successful in legal entrepreneurship are likely to different from those that are likely to be successful in double entrepreneurship. Given that China is rapidly moving toward a market economy, "institutional entrepreneurs", that are skillful at dealing with government officials and public opinion are likely to face risks in a changed economy (Daokui Li *et al.*, 2006).

As Veblen (1915) argued, transition from state ownership toward private ownership and a market economy can lead to changes in cultural and behavioral characteristics. Economic motives and behaviors of individuals are also likely to change (Karayiannis and Young, 2003). Entrepreneurial activities themselves can lead to "new cognitive paradigm", which are likely to change the ingredients of entrepreneurship in China. As noted above, some entrepreneurs are still considered as selfish and in some cases the success is measured in terms of their contribution to the society. The lens through which the society views entrepreneurship and measures of success of entrepreneurs may change over time.

As noted above, rapid rise in the inflow of Chinese with education and experience in entrepreneurial ventures in Western countries is also like to bring significant changes in "way of thinking" related to entrepreneurship. An increasing number of Chinese educated abroad are returning home. In 2002 alone, over 18,000 Chinese graduated from foreign universities returned to China. This figure was 47% more than in 2001, double that in 2000, and over three times the figure for 1995 (Lynch, 2003). Chinese with experience in entrepreneurial ventures in the U.S. are thus expected to have a more positive attitude towards risk taking. As mentioned above, there has been an emergence of "club culture", which has stimulated interactions among entrepreneurs and other professionals (Wang, 2001). Such interactions with Chinese returned from overseas, proper Chinese entrepreneurs are likely to develop more positive attitude towards risk taking.

References

Abernathy, M.A. and Chua, W.F. (1996). "A Field Study Of Control System 'Redesign': The Impact Of Institutional Processes On Strategic Choice", *Contemporary Accounting Research*, 13, 569–606.

Acs, Z. J. (1992). "Small business economics: A global perspective", Challenge, 35(6): 38-44.

Amin, A., and N. Thrift (1995). "Globalisation, institutional thickness and the local economy", *in Managing Cities: The New Urban Context*, P.Healey, S. Cameron, I S. Davoui, S. Graham and Ali Madani-Pour (Eds), 92-108, John Wiley, Chichester

Aldrich, H. (1999). *Organizations evolving*, Newbury Park, CA: Sage

Anderson, Alistair R; Jin-Hai Li, Richard T Harrison, and Paul J A Robson (2003). "The increasing role of small business in the Chinese economy," *Journal of Small Business Management*, (Jul), 41(3), 310.

Baum, J. and C. Oliver (1991). "Institutional linkages and organizational mortality", Ad*ministrative Science Quarterly*, 36, 187–218.

Baumol, WJ. (1986). *Entrepreneurship and a century of growth*, J Bus. Venturing 1, 14145

Berger, P.L., Luckmann, T. 1967. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. New York: Doubleday.

Boswell, C. and J. Baker (2006). "Securing value in China", *ICIS Chemical Business*, (October 2), 1 (37), 31-32

Browne, Andrew (2004). "Chinese Recruit Top Executives Trained Abroad", *Wall Street Journal*, Nov 30, B.1

Busenitz, Lowell W., Carolina Gomez and Jennifer W Spencer (2000). "Country institutional profiles: Unlocking entrepreneurial phenomena ", *Academy of Management Journal*, (Oct), 43(5), 994-1003.

Bush, Paul D. (1983). "An Exploration of the Structural Characteristics of a Veblen-Ayers-Hoster Defined Institutional Domain", *Journal of Economic Issues*, 17(3), 35-66.

Bush, Paul (1987). "The Theory of Institutional Change", *Journal of Economic Issues*, 21, 1075-1116.

Bush, Paul D. (1989). "The concept of "progressive" institutional change and its implications for economic policy formation", *Journal of Economic Issues*, 23(2), 455-464.

Bush, Paul D. (1994). "The pragmatic instrumentalist perspective on the theory of institutional change", *Journal of Economic Issues*, 28(2), 647-657.

Campbell, John L. (2004). *Institutional Change and Globalization*, Princeton, NJ: Princeton University Press.

Culpan, T. (2006). "Industry awaits IP court in China", *Billboard*, (April 22), 118(16),14.

Daokui Li, David, Junxin Feng, Hongping Jiang (2006). "Institutional Entrepreneurs", *American Economic Review*; (May), 96(2), 358-362.

David, Paul A., and Dominique Foray (1994). "Dynamics of Competitive Technology Diffusion Through Local Network Structures: The Case of EDI Document Standards", *in Evolutionary Economics and Chaos Theory: New Directions in Technology Studies*, L. Leydesdorff and P. Van den Besselaar (Eds.), Pinter, London and New York, 63-78.

Deephouse, D. L. (1996). "Does Isomorphism Legitimate?", *Academy of Management Journal*, 39, 1024-1039.

Dickson, Marcus, BeShers, Renee and Gupta, Vipin (2004). "The impact of societal culture and industry on organizational culture: Theoretical explanations", *in Culture, leadership, and organizations : the GLOBE study of 62 societies*, Robert J. House, Paul J. Hanges, Mansour Javidan, Peter W. Dorfman, and Vipin Gupta (Eds.), Thousand Oaks, Calif. : Sage Publications.

DiMaggio, Paul J., and Powell, Walter W. (1983). "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields", *American Sociological Review*, 48, 147-160.

Djankov, Simeon, Yingyi Qian, Roland, Gérard, Zhuravskaya, Ekaterina (2006). "Who Are China's Entrepreneurs?", *American Economic Review*; (May), 96(2), 348-352.

Doucouliagos, Chris (1995). "Institutional bias, risk, and workers' risk aversion", *Journal of Economic Issues*, (Dec), 29(4), 1097-1118.

The Economist (US) (2002) *Not in the club; China's private sector.*(Being a private-sector entrepreneur in China has its own special problems), (Oct 12).

The Economist (US) (2003). *The communist entrepreneur; Face value.*(Yin Mingshan: Chinese entrepreneur--and communist), (March 29), 366(8317).

The Economist (2006). *China's uneasy billionaire*, (Feb 4), 378(8463).

Edmundson G, Flynn J, Glasgall J. (1996). "Come home little startups: Europe wants its own NASDAQ to capture more IPOs", *Business Week*, February 26

European Network for SME Research (ENSR). 1996. The European observatory for SMEs (4th annual report). Zoetermeer, The Netherlands: ENSR

Friedman, Thomas L.(2005). "From Gunpowder to the Next Big Bang", *New York Times*, Late Edition (East Coast), Nov 4, A.27.

Gereffi G, Hempel L. (1996). "Latin America in the global economy: Running faster to stay in place", *NACLA Report on the Americas*, 29(4), 18-27.

George, E., Chattopadhyay, P., Sitkin, S.B., and Barden, J. (2006). "Cognitive underpinnings of institutional persistence and change: a framing perspective", *Academy of Management Review*, 31(2), 347-385.

Guiheux, Gilles (2006). "The Political "Participation" of Entrepreneurs: Challenge or Opportunity for the Chinese Communist Party?", *Social Research*, (Spring), 73(1), 219-246.

Han, Vo Xuan and Roger Baumgarte (2000). "Economic reform, private sector development, and the business environment in Viet Nam", *Comparative Economic Studies*, (Fall), 42(3), 1-30.

Harper, David A. (2003). *Foundations of Entrepreneurship and Economic Development*, New York: Routledge, 276

Harwit, Eric (2002). "High-technology incubators: Fuel for China's new entrepreneurship?", *The China Business Review*, (Jul/Aug), 29(4), 26-29.

Heer, P. (2000). "A house divided", *Foreign Affairs*, 79,18-24.

Hofstede, G. H. (1980) Culture's consequences: International differences in work-related values. Newbury Park, CA: Sage

Holt, D. G. (1997). "A Comparative Study of Values among Chinese and U.S. Entrepreneurs: Pragmatic convergence between contrasting cultures" Journal of Business Venturing, 12(6), November

Hoogewerf, Rupert (2002). "A new breed of Chinese entrepreneur: Louis Zan Shengda is one of China's richest businessmen--and a prime target for private bankers moving into the country. Rupert Hoogewerf speaks to him about his success, and his investment plans for the future. (Chinese Entrepreneurs)", *Wealth Management*, (Summer), 24(2).

Hsu, Carolyn (2006). "Cadres, Getihu, And Good Businesspeople: Making Sense Of Entrepreneurs In Early Post-Socialist China", Urban Anthropology & Studies of Cultural Systems & World Economic Development, (Spring), 35(1), 1-38.

Hughes, N. C. (2005). "A trade war with China?", *Foreign Affairs*, (Jul/Aug), 84(4), 94.

Kahn, Joseph. (2006). "A Sharp Debate Erupts in China Over Ideologies", *New York Times*, (Late Edition (East Coast)). Mar 12, 1.

Kanellos, M. (2002). "Software comes of age: Piracy crackdown pays off", *CNET News.com*, (July 11), http://news.com.com/2009-1001-940335.html

Karayiannis, Anastassios D, Allan E Young. (2003). "Entrepreneurial Activities in A Veblenian Type Transition Economy", *American Economist*, (Fall), 47(2), 40.

Keeble, D., Lawson, C., Moore, B., Wilkinson, F. (1999), "Institutional thickness in the Cambridge region", Regional Studies, Vol. 33 pp.319-32.

Kharbanda, V. P. and Suman, Y. 2002. Chinese initiative in the software industry – Quest to leap. *Current Science*, Vol. 83, No. 12, 25 December 2002, 1455

Kolko, Gabriel (1997). "Privatizing communism: Politics and market economics in Russia and China", *World Policy Journa*l, (Spring), 14(1), 23-24.

Krug, Barbara (2004). *China's Rational Entrepreneurs: The Development of the New Private Business sector*, London: RoutledgeCurzon.

Lau, L. J., Y. Qian, G. Roland. (2001). "Reform without losers: An interpretation of China's dual-track approach to transition", Journal of Political Economy, 108(1), 120-143.

Lawrence, Thomas B., Winn, Monika I. and Jennings, P Devereaux (2001). "The Temporal Dynamics of Institutionalization", *The Academy of Management Review*, 26(4), 624-644.

Leblebici, H., Salancik, G. R., Copay, A., and King, T. (1991). "Institutional change and the transformation of interorganizational fields: An organizational history of the U.S. radio broadcasting industry", *Administrative Science Quarterly*, 36, 333-363.

Levitt, Peggy (1998). "Social remittances: Migration driven local-level forms of cultural diffusion", *The International Migration Review*, (Winter), 32(4), 926-948.

Li, Hongbin; Lingsheng Meng; Junsen Zhang (2006). "Why Do Entrepreneurs Enter Politics? Evidence from China", *Economic Inquiry*, (Jul), 44(3), 559-578.

Ling Zhijun (1998). "Chenfu: 1989-1997 Zhongguo Jingji Gaige Beiwanglu" (Ups and downs: Memorandum of China's economic reform during 1989-1997). Shanghai: Tongfang Chuban Zhongxin.

Loyalka, Michelle Dammon A. (2006) Chinese Welcome for Entrepreneurs, *Business Week Online*; 1/6, 1.

Lynch, D.J. (2003). "More of China's best, brightest return home ; Nation's brain drain days over as it becomes land of opportunity", *USA Today*, Mar 7, B.01

MacDonald, Scott B. (2006). "A Tale of Two Indias (economic development)", *Society*, 72(6), 43-48.

Meredith, R. (2003). *(Microsoft's) Long March, Forbes*, (February 17), 171(4), 78-86.

Meyer, J. W., and Rowan, B. (1977). "Institutionalized Organizations: Formal Structure as Myth And Ceremony", *American Journal of Sociology*, 83, 340-363.

Meyer, J. W., and Scott, W. R. (1983). *Organizational environments: Ritual and rationality*, Beverly Hills, CA: Sage.

Moore, Mick (1997). "Societies, polities and capitalists in developing countries: A literature survey", *The Journal of Development Studies*, (Feb), 33(3), 287-363.

Morris, Michael and Minet Schindehutte (2005). "Entrepreneurial Values and the Ethnic Enterprise: An Examination of Six Subcultures", *Journal of Small Business Management*, (Oct.), 43(4), 453-479.

Mourdoukoutas, Panos (2004). "China's Challenge", *Barron's*, (Feb 16), 84(7), 37.

Muller, Christian, Takao Fujiwara, and Cornelius Herstatt (2004). "Sources of Bioentrepreneurship: The Cases of Germany and Japan", *Journal of Small Business Management*, (Jan), 42(1), 93.

Mueller, S. L. and Thomas, A. S. (1997). "National culture and entrepreneurial orientation: A nine-country study". Paper presented at the annual meeting of the Academy of Management, Boston.

Nair, S R. (1996). "Doing business in China: It's far from easy", *USA Today*, (Jan), 124(2608), 27-29.

Nee, Victor. (1989). "Peasant Entrepreneurship and the Politics of Regulation in China." Pp. 169-207 in *Remaking the Economic Institutions of Socialism: China and Eastern Europe*, Victor Nee and David Stark (Eds.), Stanford, Calif.: Stanford University Press

Newman, K. L. (2000). "Organizational transformation during institutional upheaval", *The Academy of Management Review*, 25(3), 602-619.

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge, MA: Harvard University Press.

North, D. C. (1993). "Douglass C. North-Autobiography", http://www.nobel.se/economics/laureates/1993/ north-autobio.html

Pei, M. (1998). "Is China democratizing?", *Foreign Affairs*, 77, 68-82.

Peng, Yusheng (2004). "Kinship Networks and Entrepreneurs in China's Transitional Economy", *The American Journal of Sociology*, (Mar), 109(5), 1045-1074.

Pomfret, John. (2001). "New Deal For China's Capitalists; Businessmen Join Party But Run Their Own Show"; *The Washington Post*, Jul 3, A.12.

Potter, P. B. (2004). "Legal reform in China: institutions, culture, and selective adaptation", *Law & Social Inquiry*, 29(2), 465-495.

Ruef, M., and Scott, W. R. (1998). "A multidimensional model of organizational legitimacy: Hospital survival in changing institutional environments", *Administrative Science Quarterly*, 43, 877–904.

Schneider, A., (1999). "US neo-conservatism: Cohort and cross-cultural perspective" *The International Journal of Sociology and Social Policy*. 19(12) 56-86.

Schramm, Carl J. (2004). "Building Entrepreneurial Economies", *Foreign Affairs*, (Jul/Aug), 83(4), 104.

Scott, R. (1995). *Institutions and organizations*, Thousand Oaks: CA: Sage.

Scott, R. (2001). *Institutions and organizations*, Thousand Oaks: CA: Sage.

Scott, W. Richard, Ruef, Martin, Mendel, Peter J. and Caronna, Carol A. (2000). *Institutional Change And Healthcare Organizations : From Professional Dominance To Managed* Care, Chicago : University of Chicago Press

Segal, Adam (2004). "Is America Losing Its Edge?", *Foreign Affairs*, (Nov/Dec), 83(6), 2.

Sender, Henry (2000). "China's future: Entrepreneurs", *Far Eastern Economic Review*, (Jan 27), 163(4), 73.

Shuli, Hu (2004). "Enrons of China ; Scoops: They reveal a business culture that is now teetering on the brink between cronyism and capitalism", *Newsweek*, Dec 1, 67.

Singer, P. A., Salamanca-Buentello, F. and Daar, A. S. (2005). "Harnessing Nanotechnology to Improve Global Equity", *Issues in Science and Technology*, 21(4), 57-64.

Sitkin, S. B., and Sutcliffe, K. M. (1991). "Dispensing legitimacy: Professional, organizational, and legal influences on pharmacist behavior", in Research in the sociology of organizations, P. Tolbert and S. Barley (Eds.), 8, 269–295. Greenwich, CT: JAI Press.

Solinger, D. J. (1995). "China's urban transients in the transition from socialism and the collapse of the communist urban public goods regime", *Comparative Politics*, 27(2), 127-146.

Spencer, Jennifer W, Thomas P Murtha, Stefanie Ann Lenway. (2005). "How Governments Matter to New Industry Creation, Academy of Management", The Academy of Management Review, (Apr), 30(2), 321.

Strang, David and Meyer, John W. (1993)."Institutional Conditions for Diffusion", *Theory and Society*, 22, 487-511.

Stokes, B. (2005). "China's high-tech challenge", *National Journal*, (July 30), 37(31), 2448-2457.

Takyi-Asiedu, S. (1993). "Some Socio-Cultural Factors Retarding Entrepreneurial Activity in Sub-Saharan Africa," *Journal of Business Venturing*, 8(2), 91-98.

Tolbert, P. S. and Zucker, L. G. (1983). "Institutional Sources of Change In The Formal Structure Of Organizations: The Diffusion Of Civil Service Reform, 1880-1935", *Administrative Science Quarterly*, 28, 22-39.

Waldringer, R., H. Aldrich, and R. Ward (1990). Ethnic Entrepreneurs Immigrant Business in Industrial Societies, London: Sage, 1990 in Doucouliagos, Chris. Institutional bias, risk, and workers' risk aversion Journal of Economic Issues. 1995, 29(4), 1097-1198.

Wang, J. (2001). *Innovative spaces: Enterprise clusters and regional development* [in Chinese], Beijing: Peking University Press. UNDP 2001. *Human Development Report*, New York: United Nations Development Program, http://www.undp.org/hdr2001/completenew.pdf, accessed 11 July 2001.

Wehrfritz, George and Alexandra A. Seno. (2003). "Chugging Along; No one saw China as one of the big 'engines' that could drive the world out of a slump, until now", *Newsweek*, (International ed.), Jun 23, 42.

Yang, Keming (2002) "Double entrepreneurship in China's economic reform: An analytical framework", *Journal of Political and Military Sociology*, (Summer), 30(1), 134-148.

Zucker, L. G. (1987). "Institutional theories of organizations", *Annual Review of Sociology*, 13, 443-464.

PROGRAMME GRANDE ÉCOLE, GRADE MASTER

GROUPE ÉCOLE SUPÉRIEURE DE COMMERCE DE PAU

Campus Universitaire 3, rue Saint-John Perse - B.P. 7512 - 64075 PAU Cedex - France
Tél. 33 (0)5 59 92 64 64 - Fax : 33 (0)5 59 92 64 55 - info@esc-pau.fr - www.esc-pau.fr
ESCPAU Bangalore Office : SP Ramanna N°29, 9th Main Shivanagar Bangalore 560 010 - India
ESCPAU Washington Office : 6601 Rivercrest Court. Bethesda, MD 20816 -2150 - USA

Membre de la Conférence des Grandes Écoles et du Chapitre des Grandes Écoles de Management - Membre de la Fondation Nationale pour l'Enseignement de la Gestion des Entreprises (FNEGE) - Membre Fondateur de Pyrénées-Océanes Campus - Membre de l'European Foundation for Management Development (EFMD) - Membre de l'Association to Advance Collegiate Schools of Business (AACSB) - Membre du Consejo Latinoamericano de Escuelas de Administración (CLADEA) - Membre du Consejo de Rectores de las Universidades Chilenas (CRUCH) - Membre du Centre de Management Euro Amérique Latine (CMEAL) Membre de la Conférence des Recteurs et des Principaux des Universités du Québec (CREPUQ).

PUBLICIS ACTIV RC B 337 934 483