Scientific
Research

iBusiness

JOURNAL EDITORIAL BOARD

CONTENTS

iBusiness (iB)

Journal Information

Scientific
Research

Blended Change Management: Concept and Empirical Investigation of Blending Patterns

Michael REISS

BWI, Abt. II, Universität Stuttgart, Keplerstr. 7, Stuttgart, Germany.
Email: lehrstuhl.organisation@bwi.uni-stuttgart.de

ABSTRACT

In coping with the challenges of revolutionary or evolutionary change processes, change managers do not rely on single tools but on toolboxes containing several domains of tools. The impact of toolboxes on change performance depends both on the complementary inter-domain mix and the intra-domain blending of tools. The patterns of blending are investigated both conceptually and empirically with respect to scope, diversity and coupling of tools. Survey results indicate that blending practices are predominantly determined by rational tool evaluation and by task context.

1. Introduction

Failure rates of change projects of 50% and more are a prominent challenge to change managers [1]. Causes of failure have been scrutinized [2–4] and an array of measures to enhance success rates has been developed. There are four interdependent approaches responding to these challenges: The focus of the *functional* approach is optimized change processes, e.g. better timing of pilot and roll-out phases and parallel instead of sequential communication with managers, employees, customers and other stakeholders affected by the change. The organizational role approach emphasizes the installation of change agents or caretakers, such as champions, sponsors, initiators, implementers, and facilitators [5] as well as the installation of committees to enable participation and empowerment. Skills of agents [6], institutions (e.g. dynamic capabilities [7]), and targets (capacities of learning) are a key success factor in this approach. Finally there is a tool or instrumental approach [8,9]. The tool approach represents the arena of change management with a maximum overlap of scientists' and practitioners' handling of change projects. The overall fund of tools is steadily growing [10]. So are toolboxes in terms of configurations of tools for use in specific change projects. This augmentation is partly due to the activities of change consultancies that develop and use tool innovations as a competitive strategy of differentiation (e.g. tool branding).

Most tools support the implementation of a new concept (i.e. "the" change) into an existing context. Together with the strategy, systems and technology, people represent a generic and ubiquitous area of context. The section of a change management toolbox dedicated to influence attitudes and behaviors of people contains four core domains (see Figure 1). This classification is based upon the primary functionality of the tools, i.e. influencing the acceptance of change from employees, managers, customers, public and other groups affected by the transition.

To make people commit to new behavior and thereby accept or even embrace change both the ability for change (skill factor) and the willingness for change (will factor) must be taken care of. The tool domains of information and instruction complementarily determine the level of skill while involvement and integration determine the will of the people affected. Information and communication tools in change management are supposed to specify the content, the objectives, the consequences, and the progress of a change project. Training is expected to develop the requisite personal competencies, and motivation tools to provide a commitment to the change project. Finally, organization tools such as participation and project organization integrate people actively into the change process by giving them roles they are supposed to play (e.g. project managers, mediators, multiplicators, coaches, etc.). Some tools like workshops or wikis are versatile or multifunctional due to their capacity to inform, skill, motivate, and integrate simultaneously.
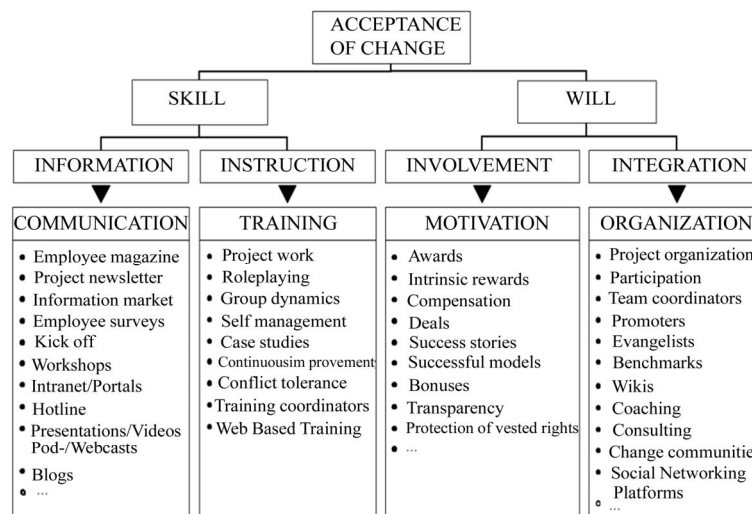
Figure 1. Management tools for enhancing acceptance of change

In addition to the inter-domain mixing of functionally complementary tools (i.e. for information, instruction, motivation, and organization), toolboxes applied in change projects are also the result of intra-domain mixing, i.e. the configuration of functionally identical ("redundant") tools, such as interactive tools like workshops combined with other interactive tools like bilateral talks and negotiations or print media combined with electronic media. Whenever significantly heterogeneous tools are applied simultaneously, this process and the resulting configuration are mostly called "blending" and "blended" respectively, such as blended learning as a mix of face- to-face- and e-learning [11]. Diversity due to a hybrid configuration is a ubiquitous feature of change management in general (e.g. combining revolutionary and evolutionary changes in so called hybrid systems, [12]) and of tool blending in particular:

1) Change managers often utilize both the potential of participative Organizational Development (OD) tools and of "manipulative" marketing tools (e.g. "selling" change by means of awards, slogans, road shows, "deals", and advertising).

2) Besides soft factors such as coaching, success stories, persuasion, role and theatre playing [13], hard factors like Balanced Scorecards or milestone planning are quite frequently used [14].

3) Top-down-interventions (e.g. kick-offs, cascading information processes, corporate blogs) are combined with bottom-up-initiatives (e.g. suggestion schemes, continuous improvement process, communities, social software).

4) Change management relies on face-to-face interactions as well as on virtual interactions via electronic media [15].

5) Communication is based on unilateral broadcasting

and on interactive environments such as team meetings, open space, workshops and video conferences.

6) Acquisition of requisite skills is accomplished by on the job- and off the job-learning.

7) Hybrid tools of "edutainment" and "infotainment" (e.g. business TV, gaming, [16]) help foster learning motivation.

Hence, every configuration of toolboxes implies numerous blending decisions. The denotation of blending emphasizes its hybrid make-up ("combining extremes") whereas the connotation of blending refers to the effectiveness of these mixes, captured in slogans like "best of both worlds-combinations". Blended models outside the management sphere such as blended materials (alloy), beverages or tobacco serve as role models.

2. The Logic of Blending

2.1 Evaluation of Tool Blending: Opportunities and Risks

Augmentation per se is not a guaranty for enhanced performance and does not automatically mean enrichment. Neither is blending per se the reliable cure of weaknesses and drawbacks of some change management tools. A transfer of lessons learned from other blending arenas—modern arenas such as blended or hybrid learning [17], multimedia, multichannel distribution (brick&click-or hybrid companies), or diversity management as well as traditional arenas like carrot & stick-leadership, dual leadership or interactions of intrinsic and extrinsic motivation—indicate that a "more means better" point of view is rather naïve. The message of these lessons reads: blending goes along with several risks of inefficiency or even ineffectiveness, e.g. overload, lack of orientation, high costs and frictions.
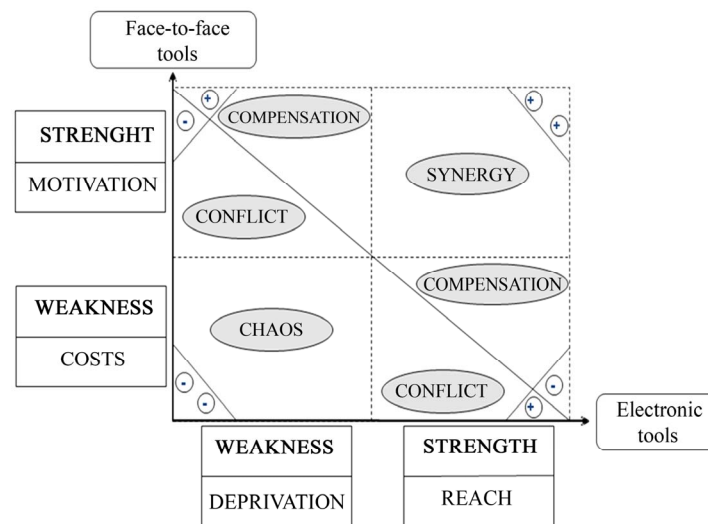
Figure 2. Opportunities and risks of tool blending

To respond to the resulting ambivalent connotation of blending by a simple "blending: bane or boon-approach" turns out to be too superficial. Instead the rational design of toolboxes should follow the principle of "no integration without evaluation". Evaluation serves as a guide for all activities supposed to amplify opportunities and to mitigate risks related to tool blending. Figure 2 illustrates a simplified evaluation of blending face-to-face tools and electronic tools in change management.

In correspondence to the hybrid character of the object of evaluation, i.e. the mix of online and onsite tools, the evaluation model for blended toolboxes is also based on a hybrid design. The evaluation of the two tool clusters is focused on their respective strengths and weaknesses. *Opportunities* are the result of the "productive tension" in a hybrid construction (area above the diagonal in Figure 2) whereas risks come from "unproductive frictions" between the diverse components (area below the diagonal). The evaluation is boiled down to just one typical strength and weakness of each cluster. Face-to-face tools like, for example, group discussions or bilateral discussions often have a strong impact on the motivation by offering the possibility to give immediate and personalized feedbacks and by satisfying social needs. These advantages, however, come along with high costs for traveling to physical meetings, workshops or seminars. Electronic tools like e-mail or intranet portals typically have a broad reach, since they are able to deliver information quickly and easily to geographically dispersed employees, team members, or customers affected by the change. However, this broadcasting bears the risk of social deprivation of the actors involved because emotional feedbacks in the change process are impaired. Thus, electronic tools often leave social needs as well as security

needs unsatisfied and fail in dealing with confusion, anxieties, and other typical effects of change.

Tool blending attains synergy effects in terms of richness and reach of communication when simultaneous communication via intranet and workshops is applied. Compensation of weaknesses by strengths of the additional tool is reached when the time needed for seminars can be reduced by providing general basic information via electronic media prior to the onsite seminars. One category of risk caused by blending is conflict: Providing redundant content by print and by electronic media may provoke a conflict with project budget restrictions. Incompatibility may even cause chaos, for instance when contradictory content is delivered via electronic media and physical meetings respectively. This may be due to the fact that electronic media are normally more up-to-date whereas print media often distribute obsolete data in turbulent change processes.

2.2 Patterns of Tool Blending

2.2.1 Dimensions of Blending Patterns

To analyze, categorize and design blended toolboxes a multidimensional approach is needed that is based on three parameters of blending:

Scope: This dimension covers the quantitative aspects of blending, i.e. the number of tools incorporated in the blended toolbox and the proportions of blending, i.e. the ratio of percentage of use of the tools in question. 50:50-proportions stand for balanced blending while an 80:20-ratio indicates the dominance of one tool category.

Diversity: A combination of workshops, flyers, meetings and a letter from the CEO in the employees' magazine characterize homogeneous blending since all tools in the list rely on conventional communication via physical

meetings or print media. The level of diversity increases when both face-to-face tools and electronic media (e.g. e-mail, virtual communities and weblogs) are used. Diversity stems from contrast between tools (heterogeneous blending), since electronic media, unlike face-to-face change management, go along with asynchronous communication and lack a direct contact between the participating players.

Coupling: Blending ranges from loose to tight coupling of tools. In the case of loose coupling, change managers pick different tools out of a blended toolbox to be applied in distinct sectors or stages of the change project. By this strictly separated handling, tools can be adjusted to different segments of the context (e.g. different target groups like employees versus temporary manpower, top management versus lower management), preferences of clients (reflecting their respective corporate culture) and modules of the change concept (e.g. redesigned business processes, organization charts, incentives systems, lay-offs). From a rational management point of view this corresponds to the idea of contingent management with respect to tool utilization. Likewise, face-to-face communication in the pilot phase can be combined with electronic communication in the roll out-phase which allows an adjustment to the size of the respective target groups.

Tight coupling is either related to toolboxes in terms of blended menus or blended bundles. Blended menus offer at least two tools (e.g. e-mail or telephone, print media or electronic newsletters, physical workshops or virtual meeting on internet community platforms) as alternative options. Providing menus is client-friendly but quite costly: Since tools are not pre-selected within a contingent change management approach (i.e. loose coupling), the entire range of diverse options has to be provided until employees or clients make their choices.

In blended bundles, tight coupling is performed in a "total" fashion yielding new genuinely hybrid tools that incorporate both genes of their parent tools: Project meetings are not either face-to-face or virtual, but semi-virtual with some team members participating physically, others virtually via videoconferencing. Communication is neither purely top-down nor bottom-up, but takes place in an iterative down-up process.

Each of the three dimensions also serves as a scale to measure the level of hybridity of tool blending. So a broad, balanced scope of heterogeneous and tightly coupled tool bundles represents the maximum challenge for change managers because the performance of the blended toolbox cannot be easily traced back to the strengths and weaknesses of the tool components in question. Focused, unbalanced homogeneous and loosely coupled toolboxes on the other hand are by far easier to understand and to evaluate.

2.2.2 Engineered and Emergent Patterns

Like strategies, structures, systems or other building blocks of management, toolboxes are either the result of deliberate planning (engineered patterns of blending) or the result of unplanned activities (emergent patterns of blending) or possibly a (hybrid) combination of both, in the tradition of hybrid process models like guided evolution, logical incrementalism, organic rationality or organized anarchy.

Engineered patterns of blending are created by a rational planning process. The objective of this process is achieving an optimal opportunity-risk ratio. This is accomplished by aggregating the weighed positive and negative evaluations of blended concepts (see Figure 2). The aggregation has to consider reciprocal interaction effects triggered by the tension between the components of a hybrid toolbox.

Emergent patterns of blending are not determined by rational procedures of evaluation and design. Like in all other fields of management, change managers are not necessarily guided by the rational evaluation of opportunities and risks. There are many other factors that influence blending activities. Some of them are apparently irrational from the standpoint of rational optimization. A prominent example is "change management follows fashion", with "hypes" (quite common in the lifecycle of electronic trends [18]) representing an extreme species of fashion. However, go-with-the-flow behavior in change management may have definitely rational advantages for the situation of the individual manager [19]. The long list of factors influencing the behavior of change managers contains many personal factors such as

1) expertise in change management, e.g. number of change projects managed in the past.

2) personal preferences for tools: preferences may be a matter of familiarity with tools or even an expression of a dogmatic approach when for instance orthodox disciples of OD refuse to adopt Business Process Reengineering tools because of a cultural misfit.

3) lack of skill in handling specific tools.

4) avoidance of risk: Cautiousness may also make change managers refrain from tool blending which very likely provokes some risks (see Figure 2).

5) variety seeking: Openness to new tools and trends. This experimental approach is sometimes reflected in randomly composed toolboxes.

To assess the respective relevance of rational and personal determinants of tool blending and the resulting relevance of engineered or emergent patterns respectively, empirical investigation is required. Existing analyses are non-empirical, empirical but focused on the range of complementary tools or focused on single tools (not on toolboxes) or empirical but case study-based providing evidence that lacks representativeness.

3. Empirical Investigation of Blending Patterns

3.1 Survey Design

To examine the "real world" patterns of tool blending, the department of organizational behavior at Stuttgart University conducted an online survey amongst German, Swiss and Austrian change experts in first quarter of 2008. In addition to this, a weblog (http://www.change-zweinull.de) was installed to enable a virtual sharing and exchange of knowledge. The arena of tool blending examined was the configuration of non-electronic and electronic tools. The tools were clustered into two groups: face-to-face tools (workshops, multiplicators, top management commitment, employee magazines, seminars, brochures/folders/flyers, bilateral talks) and electronic tools (virtual communities/internet forums, intranet portals, information videos, e-mail newsletter, web-based trainings, podcasts/webcasts, individual weblogs, social networking platforms, wikis, and corporate weblogs). The web 2.0 tools investigated are individual weblogs, corporate weblogs, wikis, social networking platforms and podcasts/webcasts. This selection reflects the common web 2.0 tools [20–23].

The majority of respondents were contacted directly via personalized e-mails. The respondents were asked to forward the e-mail to other change managers among their colleagues and clients. Furthermore, a link to the survey was integrated in several electronic newsletters. The project weblog also provided the possibility to take part in the survey. With 305 respondents the return rate (as percentage of the number of mails sent) is 15.5%.

Almost half of the respondents are consultants, less than a fourth of the respondents are academic staff and faculty, and approximately one sixth of the respondents are employed in manufacturing or service companies.

Change managers (people who have already managed change projects) cover almost three fourths of the respondents. Within this group, a majority of change managers has managed between six and 50 projects and can be regarded as well-experienced in change management. Change managers who have managed more than 50 projects account for only 4.3%. A look at the change expertise focusing the spectrum of change categories shows that 45% of all change projects are restructuring projects, followed by strategy shift projects with approximately one third of all projects. 30% of the participating change managers predominantly manage business process reengineering and cultural change projects respectively. Approximately 20% of the change managers are regularly involved in the management of IT implementation projects. The survey investigated mainly a) the incidence of face-to-face and electronic change management tools in change management and b) the existing patterns of blending as well as their determinants.

3.2 Results

Scope of tool blending: The survey supports the assumption that the use of multiple change management tools is standard. More than 70% of the respondents use at least four tools in change management frequently or always. Almost 9% use ten or more instruments at least frequently. Only 35% of the respondents use two or more electronic media frequently or always. When the answers "frequently" and "sometimes" are aggregated, more than 70% of the respondents use at least two electronic tools in change management.

Diversity of tool blending: From the data, three basic types of blended toolboxes can be distinguished with respect to the diversity of blending: Focused toolboxes are used by change managers who concentrate on a particular "core cluster" of tools (here: face-to-face tools). These managers are reluctant to blending and consequently do not use any instrument from the other cluster. In blended toolboxes a distinction between a primary cluster and a secondary cluster is not feasible. The respective change managers do not have a clear preference for one of the two groups but use tools from both groups frequently. Change managers working with ad hoc toolboxes do not use any tool more frequently than "sometimes". Apparently, there is no preference for one cluster of tools among these change managers. Rather, these change managers configure their toolboxes randomly. Table 1 demonstrates the respective frequencies of toolboxes in the sample.

Differentiating the instruments used with respect to the cluster they belong to shows—not surprisingly—that merely one respondent focuses solely on electronic tools, while 31% focus on face-to-face instruments in change management. Blended toolboxes represent the biggest portion in the sample (67%), whereas ad-hoc mixes account for only 2.7%.

The simple assignment of the respondents to one of the two patterns specified by level of diversity (focused and blended mixes) ignores the scope dimension of blending. A valid measure of the hybridity of blending ("blending index") must encompass both diversity and scope. The focused as well as the blended patterns are more hybrid when they are based on a larger number of change management tools. Hence, the scope dimension was differentiated into "narrow", "medium", and "broad" (see Table 2).

Using the blending index, the following types of patterns can be distinguished according to their respective degree of diversity: narrow focused (1), medium focused (2), medium blended (3), and broad blended toolboxes (4).

A look at the frequencies of the different patterns reveals a peculiar result: only 7% of the change managers in the sample put a narrow focus on face-to-face instruments, i.e. use three different tools frequently at the most.

Table 1. Frequencies of blended patterns

Scope (total number of tools used)		Diversity		Total	Ad-hoc toolboxes
		Focused toolboxes	Blended toolboxes		
All instruments	1	2	0	2	
	2	3	0	3	
	3	10	0	10	
	4	17	9	26	
	5	23	16	39	
	6	12	25	37	
	7	0	24	24	
	8	0	26	26	
	9	0	22	22	
	10	0	7	7	
	11	0	8	8	
	12	0	6	6	
	13	0	2	2	
	14	0	1	1	
	16	0	1	1	
Total		**67** (30.5 %)	**147** (67 %)	**214** (97 %)	**6** (2.7 %)

Obviously, a broad scope of tools is essential for change management. Tool blending as opposed to focusing is only practiced when at least four tools are used frequently. In the medium section of the scope dimension (four to six tools) the respondents almost evenly disperse to the two "extreme" categories focused and blended toolboxes, while there exist no focused patterns for seven instruments or more at all. The currently great importance of tool blending is furthermore confirmed by the large portion of broad blended toolboxes (more than six instruments are used frequently).

Coupling patterns: Factor analyses and regression analyses were conducted to find out what context factors lead to specific patterns of blending in general and of coupling in particular. The set of determinants examined that presumably impact coupling patterns contains task

oriented and person oriented factors. Task-oriented determinants analyzed were:
- change categories (the two categories of change projects managed most frequently)
- industry (within which change projects are managed)
- project size in terms of employees affected (number of employees affected)
- project size in terms of project manpower (number of project team members)

Three person-oriented determinants were integrated into statistical analysis:
- expertise (number of change projects managed)
- occupation of the respondents
- blending mindset: Personal assessment of the interaction between face-to-face change management tools and electronic tools, representing some-

Table 2. Typology of blended toolboxes

		scope (focused toolboxes – blended toolboxes)		
		focused	blended	total
diversity (all tools)	narrow (1 to 3 tools)	15 (7 %)	0 (0 %)	15 (7 %)
	medium (4 to 6 tools)	52 (24.3 %)	50 (23.4 %)	102 (47.7 %)
	broad (7 or more tools)	0 (0 %)	97 (45.3 %)	97 (45.3 %)
total		**67** **(31.3 %)**	**147** **(68.7 %)**	**214** **(100 %)**

Statistical measures refer to the recoded variable "blending index". The underlying scale was recoded into 1) narrow focused toolboxes, 2) medium focused toolboxes, 3) medium blended toolboxes, 4) broad blended toolboxes. The arithmetic means have been calculated excluding answers "I cannot assess".
arithmetic mean=3,07; median=3,0; standard deviation =,988

thing like a personal "blending theory". The respondents ranked these tool relationships on a scale ranging from "complementing" to "crowding out" of tools.

Table 3 illustrates that there is a relatively high positive and statistically significant correlation between project size (number of employees affected) and the blending index. Apparently, there is a tendency to focus on few instruments in smaller projects, while large projects trigger extensive blending in the use of change management tools. This corresponds to the definition of loose coupling via context segmentation ("different target groups require different tools"). On the one hand, the complementary use of electronic tools is probably required by the increasing demand for reach in large projects. Such enhancement of change management reach can only be accomplished efficiently by using internet-based media. On the other hand, effectiveness also requires more intensive blending in large-scale projects in terms of large number of employees affected. Along with the number of targeted employees, the diversity in this group of employees also increases. This heterogene-

ity can be dealt with by using a broad range of change management tools to enable individualization of change management activities, i.e. to adjust these activities to the needs and preferences of the respective employees and other target groups.

The project manpower, i.e. the number of team members, also correlates positively with the blending index, although less strongly and less significantly. Partly, this is due to the relationship between the number of employees affected and the requisite size of project teams which (not surprisingly) turns out to be statistically significant. Moreover, the number of team members also has an immediate impact on the blending of instruments: on the one hand, electronic tools are mandatory to warrant the reach of change management activities. On the other hand, the project requires a higher richness of change management toolboxes to cope with the increased need for individualization.

The analysis of correlation revealed another counter-intuitive relationship between a context variable and the blending index: the personal assessments of the interaction between electronic tools and face-to-face change

Table 3. Correlations between context variables and blending index

		Blending index	Ocupa-tion	Number of pro-jects	Industry	Employ-ees af-fected	Project man-power	Interac-tion	Change Catego-ries
Corre-lation	**Blending Index**		–,007 (n.s.)	,013 (n.s.)	–,086 (n.s.)	,289 (,000)	,150 (,039)	–,168 (,024)	,003 (n.s.)
	Occupation			–,398 (,000)	,127 (,069)	–,114 (n.s)	–,127 (0,69)	,069 (n.s.)	,075 (n.s.)
	Number of projects				–,151 (,038)	,096 (n.s.)	,141 (,040)	–,006 (n.s.)	,101 (n.s.)
	Industry					–,140 (,050)	–,116 (n.s.)	,075 (n.s.)	–,021 (n.s.)
	Employees affected						,412 (,000)	–,016 (n.s)	–,077 (n.s.)
	Project manpower							–,030 (n.s.)	–,052 (n.s.)
	Interaction								–,032 (n.s)
	Change categories								

management and the blending index have a slightly negative correlation. In other words, those change managers who assume a harmonic complementary relationship between the two groups of instruments (in Figure 2 above the diagonal) still tend to focus on one group of instruments–and thus do not exploit the opportunities of blended tool boxes. This is most likely caused by context barriers, such as a lack of technical infrastructures, of familiarity with tools, and/or by budget restrictions. Tight budget restrictions are not only a problem in small projects. All varieties of change are currently exposed to a high pressure for efficiency [10]. Also, whenever change managers experience a low degree of acceptance for such media amongst the employees affected, they may refrain from deploying these instruments, although they assume a harmonic relationship with other change management tools.

In addition to correlations between the blending index and task-or person-oriented context factors, interrelations among the context variables were examined. For this investigation, a factor analysis was conducted to discover underlying factor structures. The factor analysis yielded three components (see Table 4) which account together for an explained variance of 53.1%.

The first factor–on which the variables blending index, number of employees affected and number of team members are loading—represents the project size. The structure of this component shows that large projects require electronic change management tools and that these instruments are always used in combination with face-to-face tools. In turn, this structure also shows that the use of blended toolboxes is not based on individual tool preferences of change managers or employees affected. The application of blended toolboxes is rather driven by project requirements—in particular by project size–and thus aims to compensate the weaknesses of many a change management tool (see Figure 2).

The second factor (change expertise)—consisting of the variables occupation and number of change projects—does not contain the blending index. The structure of this factor is plausible: The occupation of the change managers affects the number of projects managed. For example,

                                     iB

Table 4. Component diagramm (rotated component matrix)

	Component		
	1	**2**	**3**
Blending index	,562		,473
Occupation		,778	
Number of projects		–,822	
Industry		,356	
Employees affected	,805		
Project manpower	,703		
Interaction			–,711
Change categories	–,314		,559

Note: extraction method: principal component analysis. Rotation method: Varimax with Kaiser-normalization, explained variance: 53.1%. The results of the rotated component matrix are considered.

Factor 1 (Project size): Blending index, employees affected, project manpower

Faktor 2 (Change expertise): occupation, number of projects

Faktor 3 (Evaluation): Blending index, interaction

consultants are "full-time" change managers and thus typically have ample change experience, while managers in other industries are less experienced in change management.

The third factor (evaluation) captures the assessment of the interaction between face-to-face change management and electronic media as well as the blending index. It does not deliver an explanation for blending practices that is as obvious as the one provided by the first factor. The negative correlation between the two variables (interaction and blending index) has to be explained by context factors not covered by the survey. On the one hand, this correlation can be explained by some additional barriers to the use of blended toolboxes. On the other hand, dynamics of toolbox design can be held responsible for this phenomenon: The majority (204 persons or 87.6%) of the respondents diagnose a crowding-out relationship between face-to-face change management and electronic tools. More than half of these 204 respondents already use blended toolboxes. The statement that the two groups of instruments crowd each other out may derive from the assumption that the future brings a migration from face-to-face change management to electronic media, in other words a step-by-step shift of proportions in favor of new electronic tools. This interpretation gets further support from the respondents' opinion concerning the future relevance of web 2.0-tools in change management toolboxes. While most respon-

dents (64.9%) estimate the current percentage of web 2.0 tools-application in change management to be less than 10%, 82% expect this share to rise in the future. That means the vast majority of the participating experts expects an increasing importance of electronic media in change management.

4. Conclusions and Outlook

There is evidence provided by the survey that blended change management is a reality reflected in blended toolboxes. The majority of change managers advocate tool blending. Only a minority concentrate on familiar face-to-face tools. Tool blending is predominantly guided by rational considerations: task features are more relevant than personal preferences (such as affinity to technology or conservative tendencies towards familiar instruments) or experience. The scope, diversity and coupling of tool blending primarily depend on project size: Large-scale projects drive the use of several different change management tools whereas small projects tend to be focused on face-to-face change management. The identified blending patterns are generic; they are in particular not influenced by specific change categories. The opportunities of blended toolboxes assessed by the experts are currently not sufficiently exploited. In particular, situational restrictions in the application of change management tools constrain the diffusion of new electronic tools, although the experts assume a harmonic complementary relationship between these tools and face-to-face change management activities. The experts' assessments outline a tool scenario that is characterized by blended toolboxes with an increasing share of electronic change management tools. This trend gets more momentum when new hybrid tools such as augmented reality [24,25] will be applied in the training of interpersonal and not only technical skills. Unlike virtual reality (e.g. avatars), these hybrid tools operate on an extremely tight coupling of virtuality and reality. Simulations of typical constellations in change processes (e.g. conflict resolution, post-merger integration, coopetitive arrangements within networks) that are normally dealt with in physical role playing and business theatre [13] up to date could profit considerably from these sophisticated blended tools.

REFERENCES

[1] IBM Global Business Services, Making Change Work, 2008, https://www.ibm.com/services/de/bcs/pdf/2007/making_change_work.pdf.

[2] Kotter, J. P., "Leading change, Why transformation efforts fail," Harvard Business Review, No. 73, pp. 59–67, 1995.

[3] Gerkhardt, M. and Frey, D., "Erfolgsfaktoren und psychologische Hintergründe in Veränderungsprozessen," Zeitschrift für Organisationsentwicklung, No. 4, pp. 48–59, 2006.

[4] Greif, S., Runde, B., and Seeberg, J., "Erfolge und Mis-

serfolge beim, Change Management," Hogrefe, Göttingen, 2004.

[5] Cawsey, T. F. and Deszca, G., "Toolkit for organizational change," Sage, Thousand Oaks, 2008.

[6] Wunderer, R. and Bruch, H., "Umsetzungskompetenz, Diagnose und Förderung in Theorie und Unternehmenspraxis," Vahlen, München, 2000.

[7] Helfat, C. E., *et al.*, "Dynamic capabilities: Understanding strategic change in organizations," Malden, Blackwell, MA, 2007.

[8] Hughes, M., "The tools and techniques of change management," Journal of Change Management, Vol. 7, No. 1, pp. 37–49, 2007.

[9] Orthey, F. M., "Tools, tools, tools, Ein Beitrag über die 'Vertoolisierung' und 'Baumarktmentalität' in der Trainings- und Beratungsarbeit," Zeitschrift für Organisationsentwicklung, No. 4, pp. 73–75, 2007.

[10] Capgemini, "Change management 2003/2008, significance, strategies, trends," 2008, http://www.capgemini.com.

[11] Thorne, K., "Blended learning: How to integrate online & traditional learning," Kogan Page, London, 2003.

[12] Blom, H. A. and Lygeros, J., "Stochastic hybrid systems," Theory and Safety Critical Applications, Springer, Berlin, 2006.

[13] Cornelissen, J. P., "What are we playing at? Theatre, organization, and the use of metaphor," Organization Studies, Vol. 25, No. 5, pp. 705–726, 2004.

[14] Klarner, P. and Raisch, S., "Organisationalen Wandel messen, Erfolgreiche Veränderungen setzen geeignete Messinstrumente voraus, Ein Überblick zu Messinstrumenten des organisationalen Wandels und ihrer Einsatzgebiete," Zeitschrift für Organisationsentwicklung, No. 4, pp. 4–13, 2007.

[15] Pitt, L., Murgolo-Poore, M., and Dix, S., "Changing change management: The intranet as catalyst," Journal of Change Management, Vol. 2, No. 2, pp. 106–114, 2001.

[16] Lewis, M. and Maylor, H., "Game playing and operations management education," International Journal of Production Economics, No. 105, pp. 134–149, 2007.

[17] Fong, J., Reggie, K., and Wang, F. L., "Hybrid learning and education," Springer, Berlin-Heidelberg, 2008.

[18] Gartner Group, "Understanding hype-cycle," 2008, http://www.gartner.com/pages/story.php.id.8795.s.8.jsp.

[19] Kieser, A., "Moden und Mythen des Organisierens," Die Betriebswirtschaft, Vol. 56, No. 1, pp. 21–39, 1996.

[20] Anderson, P., "What is web 2.0? Ideas, technologies and implications for education," 2007, http://www.jisc.ac.uk/media/documents/techwatch/tsw07 01b.pdf.

[21] Ebersbach, A., Glaser, M., and Heigl, R., "Wiki: Web collaboration," Springer, Berlin, 2006.

[22] McAfee, A., "Enterprise 2.0. the dawn of emergent collaboration," MIT Sloan Management Review, Vol. 47, No. 3, pp. 20–28, 2006.

[23] O'Reilly, T., "What is web 2.0? Design patterns and business models for the next generation of software," 2005, http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/0 9/30/what–is–web–20.

[24] Milgram, P. and Kishino, F., "A taxonomy of mixed reality visual displays," IEICE Transactions on Information Systems, Vol. E77–D, No. 12, December 1994.

[25] Billinghurst, M. and Kato, H., "Collaborative augmented reality," Communications of the ACM, Vol. 45, No. 7, pp. 64–70, 2002.

Scientific
Research

On Entrepreneurship, Intentionality and Economic Policymaking

Félix-Fernando MUÑOZ[1], María-Isabel ENCINAR[2], Carolina CAÑIBANO[3]

[1,2]Universidad Autónoma de Madrid, Madrid, Spain; [3]Universidad Rey Juan Carlos, Madrid, Spain; [1,2,3]Instituto de Investigaciones Económicas y Sociales (IIES) "Francisco de Vitoria", Madrid, Spain.
Email: felix.munoz@uam.es

ABSTRACT

Within evolutionary economics, entrepreneurship is seen as the main force of economic change, as the agency of self-transformation within restless capitalist economic systems. Therefore, a truly evolutionary perspective on economic policy-making must consider the significance and scope of entrepreneurship. On the basis of such a perspective, it might be possible to assess future outcomes of economic evolution under different policy measures related with, for instance, stimulating entrepreneurship as a policy that would provide the seeds for recovery from a slump in an economy. In this short note, our main claim is that the very nature of entrepreneurship implies the recognition of the role played by entrepreneurs' intentions, their tendency towards transforming goals and agents' spaces of action. Recognition is possible due to a more systematic analytical integration of these elements into a theory of entrepreneurship based on a 'production of action' conception (vs. the standard framework based on a 'technology of choice'). This analytical vision sheds light on how economic policymaking should be implemented to stimulate entrepreneurship.

Keywords: *Entrepreneurship, Intentionality, Evolutionary Economic Policymaking*

1. Introduction

Any theory of economic policymaking in an evolutionary perspective should examine the significance and scope of entrepreneurship. This interest is partially compatible with developing a solid theoretical base for an evolutionary theory of economic policymaking [1]. There are also 'practical' reasons: in periods of economic slump, politicians look for economic policies and instruments that pull the economy out of recession. In their efforts to do so, they usually claim that crises open up new periods of economic opportunities. Some of them even point to entrepreneurship and policies that stimulate entrepreneurship as one of the forces for providing the seeds for recovery.

In any case, many important questions arise. For instance, what is the role of entrepreneurship in relation to the emergence of new opportunities? What do such opportunities involve? How does 'entrepreneurship' explain 'economic change'? If we take for granted that entrepreneurship plays a substantive role in the explanation of structural change, what implications would it hold for economic policymaking? Might there be a trade-off between the forces that are promoted by entrepreneurs and stimulated by politicians?

In order to deal with questions like these, an approach that integrates 'entrepreneurship' and 'economic policymaking' is needed. In this paper, we present an approach that may identify certain fundamental common elements at the base of the *links* between entrepreneurship, the generation of novelties, socio-economic self-transformation processes and economic policymaking. This approach, the so-called *action plan approach* [2], pre-requires a theoretical treatment of intended action.

The main thesis we propose here is that the categories of intentionality–such as belief, goal, intention, collective intentionality, etc. [3,4]–are necessary for a substantive explanation of entrepreneurship. The argument is consistent with the role these categories of intentionality play in cognitive sciences, artificial intelligence and social philosophy, etc. and in the explanation of individual and collective behaviour and the emergence of institutions [5–7]. Moreover, the goals and intentionality of agents play an essential role in the explanation of the emergence of novelties and evolving capabilities, institutions and learning processes. Insofar as entrepreneurial and political actions are special instances of human action, the categories of intentionality apply for substantive expla-

nations of such entrepreneurship and policymaking.

What are the implications of the above ideas for economic policymaking? The application of categories of intentionality for the explanation of entrepreneurship and its consequences for economic policymaking do not imply a new catalogue of political economy or new specific means for political action. The main implication is that we need a redefinition of the general context in which economic policymaking is implemented. The action plan approach provides such a general context and basically consists of moving from of a conception of economics as a technology of choice to economics as a conception of production of action. Our claim here is that the latter conception becomes a condition of possibility for implementing a substantive approach to entrepreneurship and to policymaking: a theory of entrepreneurship developed on a 'production of action' basis should consider the fact that new goals of action may emerge, the hierarchical ordering of goals may change, goals that have (or have not) been reached may (or may not) be removed from or replaced in agents' plans, etc. as a result of entrepreneurial experimentation. Moreover, we pose that agents' rationality depends on the goals and motivations they pursue and this claim is valid for politicians and entrepreneurs alike. Thus, what directs (entrepreneurial, political, etc.) human activity is not only economic calculus, but also the possibility of developing a true open rationality: the rationality of the unexpected in a context of radical uncertainty [8,9]. Economics as a conception of production of action is the compatible analytical context for this 'true open rationality'.

The paper is organised as follows; Section 2 points out the relationship between entrepreneurship and economic change, where the nexus between the two is novelty; Section 3 discusses the implications of the treatment of entrepreneurship from a 'production of action' point of view; in this context, Section 4 considers entrepreneurship and economic policymaking; the paper ends with concluding remarks.

2. Entrepreneurship and Economic Change

Significant contributions that address entrepreneurship lead to the identification of a common key element which lies at the basis of the links between entrepreneurship, self-transformation and the endogenous generation of novelties. This key element is that entrepreneurs do plans which involve transforming goals, and attempt to execute them interactively within the economic system. In doing so, they transform the socio-economic system.

How do they accomplish this function? Basically, entrepreneurs transform agents' spaces of action within an economic system. Entrepreneurs do this through a continued action linked to what we designate as entrepreneurs' goal dynamics: it is the entrepreneurs' intention to transform the human and physical environment that sur-

rounds them according to the goals (objectives) they have set, and which they have previously imagined and valued more than other alternatives. Together with the associated intentions and the means/actions the agents consider necessary for their fulfilment, these goals configure their action plans and once they are being deployed in interaction with the plans (and actions) of the other agents in the economy, they produce their effects in and transform the reality. Thus, intentionality and goals take on a central role in the explanation of the economic change.

2.1 Intentionality and Novelty

Although the importance of intentional action has been recognised in the evolutionary literature [19,20], evolutionary economics generally proceeds in its models and theories as if the goals pursued by agents were given. The evolutionary tradition has focused mainly on the role of knowledge to explain entrepreneurial behaviour. For some economists the distinctive function of entrepreneurs consists of generating [21], organising [22], and using [23] knowledge. Other contributions also give an important role to entrepreneurs' imagination [24] and creativity, etc. Until recently the analysis of the role played by agents' intentionality and the goals they pursued in the development of new capabilities and new patterns of behaviour, etc. had been postponed [25]. However, a deeper analysis shows that all these approaches to entrepreneurship literature are compatible with the recognition of the role played by entrepreneurs' intentions, with their focus on transforming goals.

Furthermore, this recognition should turn a more systematic analytical integration of these elements (means, actions, goals, intentions, knowledge, imagination and practical rationality, etc.) into a theory of entrepreneurial action and economic change. An approach that allows for a theoretical treatment of intended action seems to be a pre-requisite for exploring the very nature of entrepreneurship. The underlying idea is that agents' intentionality is a necessary condition for a substantive explanation of entrepreneurship. This idea is consistent with the role played by the categories of intentionality, such as belief, goal, intention and collective intentionality, etc,, in cognitive sciences, artificial intelligence and social philosophy, etc. to explain individual and collective behaviour and the emergence of institutions [4,5] and socio-economic systems [26,27]. Additionally, goals and intentionality play an essential role in explaining the emergence of novelties and evolving capabilities [28], institutions [29] and learning processes [30].

Thus, entrepreneurship does require a tendency towards a transforming goal. As these goals must be the goals of somebody, this is the reason for the existence of 'creator personalities' [31]. The label 'creator personality' designates a locus for novelty. Novelty goes together

with a transforming intention which results in the breaking of symmetries and the introduction of jumps, which becomes visible through the existence of constructive impulses, through the ability of being necessarily alert to discover and being necessarily ready to act on the basis of thoughts not held by others, etc. From this point of view, entrepreneurship can be considered as a transforming impulse and the will that points action towards the generation of change. These transforming impulses and the collective interactions (both on and off markets) generate the variety that fuels the evolutionary processes of diffusion, selection and retention. Schumpeter's [31] example of Mantegna's innovations could be interpreted as a conscious and individual act undertaken by the painter; but it could also be shown as a precursor of the 'Renaissance style'.

2.2 Entrepreneurship, Action Plans and Novelty

Economic self-transformation processes involve learning processes, as well as the emergence of completely new actions that cannot be explained only by means of mere knowledge acquisition [32]. In this context, an 'action plan' is the projective ordering of means to achieve goals located in the imagined future. The very nature of action plans is the projective character of the ordering that is involved. The action plans individuals elaborate are idiosyncratic. They can also be plans that coordinate the action and goals of many people: a plan for a trip, a company's business plan, a plan by the European Commission to achieve the objectives of the Treaty of Lisbon. An action plan can include routine patterns of behaviour, strategic designs and monitoring and valuation procedures.

A plan can also refer to its goals at several points in the future, represent hierarchical dependencies among goals and actions with as many analytical moments in time as may be required, as well as alignments of goals with other individuals' plans. Its projective character refers not only to the fact that historic time (and timing) play central roles in explaining human action, but also that actions and goals need to be imagined before they are deployed by agents[1].

Agents choose their goals of action on the basis of psychological, social, and cultural factors, as well as ethics and beliefs [33], etc. Agents constitute their action

plans using their imagination [34], taking into account that the goals they pursue are located in an imagined future [35]. Accordingly, it could be said that agents 'invent' the future on which they focus their actions. This idea is valid whether we consider objectives in the short, mid or long term. The opportunities for acting in a specific way (entrepreneurial action, policymaking, etc.) are not hidden somewhere in the external reality, waiting to be discovered by entrepreneurs or policymakers, but they 'emerge' initially in agents' minds regardless of the fact that at some time in the future they may be embodied in a written document or an organizational form, etc. Action plans are an open analytical representation of agents' projective action, in which means and goals are not given but rather *produced* by the agents themselves. At each moment in time, an action plan may be interpreted as a template or 'guide' for action that projectively *connects* elements of a different nature: something the agent wants to achieve (goals) with the actions and means the agent 'knows' afford him/her success[2].

Searching for novelty in economics sometimes corresponds to the perception of opportunities to get better results than those achieved through actions deployed in the past and present [37]. If this were the case, how would the action plan framework contribute to the study of entrepreneurship? The answer is that the action plan approach makes it possible to point out where novelties can be located.

Novelties operate in economic systems because economic agents (individuals and organisations) incorporate them into their spaces of representation when they configure their action plans and, in particular, when they settle their goals, thereby producing choice ex novo[3]. 'Rational choice is an inadequate explanation for behaviour, because neither the empirical premises nor the objectives of behaviour can be logically derived' [38]. As Loasby points out, following Hume's dictum, the search for novelty cannot be rational: 'no kind of reasoning can give rise to a new idea'. Creating opportunities for choice is, in the first place, producing new objectives of action. When agents incorporate new goals and intentions into their plans, they trigger the process of discovering new means to achieve these goals. If we consider novelty in goals as the most general case, we can treat other particular cases as novelties in means, given objectives and given means, etc.

Agents' action plans require the definitive abandonment of the timeless framework of the 'technology of choice'. The paradox of a timeless approach as an analytical basis for the explanation of processes that are necessarily deployed in time is solved through the dynamic openness of the actions and goals pursued by agents. The classical definition of economics offered by Robbins is essentially correct, but it is not sufficient.

---

[1]Imagination plays a central role in this approach since the projective character of action plans implies imagining a future course of action in order to reach one or more goals.

[2]Accordingly, an action plan may be interpreted as a very special (or rather, complex) system [36]; the elements connected in the system are of two different kinds; on the one hand, we have the action/means and, on the other, the goals. Actions/means are always linked to a goal, whereas goals may also be connected to each other. The goals within this system introduce the direction of an action, a direction that leads into the (imagined) future.

[3]Thus, action plans are the carriers of novelties.

3. Entrepreneurship from a Conception of 'Production of Action'

As already pointed out, entrepreneurship consists of setting a new goal the agent *desires*, triggering the production of action the agent deems necessary to achieve the goal, and thus, if successful, giving rise to unheard-of possibilities for the agent. From the viewpoint of a theory of entrepreneurship, it is not enough to consider the possibility of agents (entrepreneurs) learning new ways for achieving given ends. The true challenge for a theory of entrepreneurship consists of agents' admission of the endogenous generation of new goals, which finally give rise to new spaces of action for the agents themselves.

3.1 Explaining Entrepreneurs' 'Production of Action': Novelty and Economic Change

Moving from a conception of economics as 'technology of choice' to a conception of economics as 'production of action' reveals the decisive role of entrepreneurship, novelty and agents' intentionality in the explanation of economic change. The entrepreneur's most important role is the production of new courses of action; in other words, producing new economic situations. This is the very nature of entrepreneurship in the context of 'production of action': entrepreneurship requires a focus on a transforming goal. This focus involves learning processes but it can also imply the emergence of completely new actions that are not explained solely by mere knowledge acquisition processes[4].

    The entrepreneur—the 'creator personality', the *locus* of novelty—is both a 'maximizer' and a 'rational' agent. However, contrary to the neoclassical entrepreneur, in our approach the entrepreneur sets out the formal achievement of goals, which may be completely new goals that are hierarchically superior to all the other goals, etc., and for these reasons goals act as the *norm* of their own action plans. The entrepreneur defines his action plans with the setting of a goal he/she desires: the entrepreneur wants to *produce* that goal. From the 'production of action' point of view, we can also pose that the entrepreneur is rational: human action, *qua* rational, within human constraints, is intended action; there must be goals (reasons) *for* acting [40]. Like the other agents in the economy, entrepreneurs decide what their goals of action are (and what they are not) and which place they should be given in their action plans. These decisions are regardless of what their goals or actions are with or without a price.[5] The conception of *production of action* is analytically compatible with this true open rationality, where entrepreneurship plays a substantive role in the explanation of structural change. The projective links between entrepreneurs' goals and actions and their interactive deployment imply the endogenous generation of novelties and self-transformation.

    However, the identification of novelty is of little interest in itself if the consequences in terms of economic change are not explored. Let us briefly consider the logical links for the following thesis: *entrepreneurship is a source of economic change*. If economic change is 'dynamic endogenous structural change capable of inducing or generating novelties'; if structural change refers to processes that transform these structural elements; if novelty refers to the occurrence of something that has not previously taken place within any of these elements; and if novelty could be produced by entrepreneurs as a result of their goals, dynamics and transforming intentions; then entrepreneurship generates economic change. The very entrepreneurial function consists of changing the agents' spaces of action through a transforming impulse (linked to novelty) that points action towards the generation of change. As already suggested, transforming impulses in the entrepreneurs' action plans generate the variety (and collective interactions on and off markets) that fuels evolutionary processes.

    Assuming that entrepreneurship plays a substantive role in the explanation of structural change, could we point to the implications for economic policymaking from an evolutionary perspective?

4. Entrepreneurship and Economic Policymaking

Witt [1] claims the need for the rigorous incorporation of the common assumptions of evolutionary economics on agents' behaviour and other framing conditions into normative analysis[6] As Witt points out, some of the characteristics inherent to the processes of public choice and public intervention, which should be part of the normative analysis [41] of the processes of economic changes, are bounded rationality, the endogenous generation of new factual and normative knowledge and social interaction in competitive social environments. The political economy of actual policymaking, the analysis of policy instruments (for given ends), and the debate on policy goals and their legitimization (Witt [1]; italics added) must account for the possibility of changing

---

[4]Take for instance the interesting case of Grameen Bank: it has reversed conventional banking practice by removing the need for collateral and it has created a banking system based on mutual trust, accountability, participation and creativity. Yunus, the founder of Grameen Bank considered that financial resources ought to be made available to the poor under terms and conditions that are both appropriate and reasonable. Those ideas were at the origins of the microcredit system and they modify the contents and forms of the spaces of agents' action and, consequently, generate new realities. [39]

[5]An entrepreneur that maximizes his/her profits and a 'socially responsible' entrepreneur (like Grameen Bank) would differ in the specific prescriptive content of the hierarchically superior goals in their respective action plans: for the former, the maximum difference between his/her revenues and their costs; for the latter, a 'socially responsible' aim.

[6]Otherwise, economic policymaking would always be an appendix or a strange element to economic analysis.

knowledge constraints. However, what can be said about transforming changing goals (the key element of entrepreneurship) in relation to economic policymaking?

4.1 Considering the Requirements for Evolving Policymaking

The main challenge for policymaking is how to deal with the agents' 'production of action' in the (socio-) economic system. For instance, adaptive policymaking processes should imply promoting learning and taking care of flexible market performance; policymaking should also promote the search for novel action (Witt [1]: 80). Decision-making, the fundamental pillar of policymaking, which includes the exchange of information, the revision of data and the evaluation of the different alternatives (policy goals), etc., in short, the role of the policy-maker, is considered within the limits of public choice theory or, more generally, political economy as the result of the behaviour of a 'maximizer', an agent that does not stop until he/she finds the best option or as the result of a 'satisfier' [42] in the case of an agent that behaves with limited information and bounded rationality.

However, in the context of political economy as it is usually presented in textbooks and mainstream economics, both separate interests and defects in voting rules, institutions and markets, etc. are seen as 'policy failures'. Thus, it could be considered that in the context of orthodox political economy the relationship between economic and political processes (and therefore, between economics and policymaking) is one of mere juxtaposition (Figure 1). Economics and policies are set side by side in such a way that the anomalies conventional economics (from a conception of 'technology of choice') reveals certify the presence of failures within the processes of valuation and choice by agents. Conventional economic theorizing does not provide a sufficiently analytical base for integrating behaviours based, for instance, on non-consequentialist motivations, changing preferences or, more generally, (endogenous) changing goals[7]. Reducing the gap between evolutionary economic theorising and policymaking, or rather, integrating theorizing and policymaking would require theory to consider the very fact that new goals may arise, the hierarchical ordering of goals may change, goals that have been achieved may be removed from action plans and goals that have not been achieved may be replaced with other goals, etc.

[7]In this context, it is legitimate to question the mere existence of a concept such as 'political economy'.

[8]'Fully rational' in the sense commonly used of the term 'rational' in standard economics.

[9]Together with learning and creativity, etc. It is very important take into account that not all novelty is a result of intended actions or of the action plans themselves, but the result of the interaction of intended action. Thus, novelty may emerge as an unexpected consequence of interaction.

With regard to the purpose of this paper, our claim at this point is that one necessary condition for developing a consistent theoretical base for an evolutionary or Schumpeterian theory of economic policymaking is that evolving policymaking can only be approached from a systemic analysis in which the economic action of the agents involved in the system is essentially a projective action [2]. It is in this kind of system where, as already mentioned, entrepreneurship, intentionality and novelty are central elements for the analysis. Indeed, this kind of system requires adaptive policymaking [43], given the impossibility of setting fully rational[8] socio-economic policy designs and policy experimentation and variety, which are essential for individual and social adaptability. These requirements are better recognized in dynamic spaces of action without a preset time horizon and where the genuine multidimensionality of decision-making (by policymakers) means that that agents may continuously deploy learning and experimentation processes. This is possible if we analytically open the means/actions and the goals of action. This is the only logical approach compatible with a vision of policymaking in a truly evolving economy, in which time has real meaning. We refer to economics as a theory of the 'production of (human) action'.

4.2 Entrepreneurship as a Challenge for Evolving Economic Policymaking

If novelty can be produced by entrepreneurs as a result of their goals, dynamics and transforming intentions[9], then entrepreneurship causes economic change and poses *new challenges* for policymaking. 'Evolving economy' and 'economic policymaking' are not juxtaposed concepts; there is a comprehensive relationship between the two. This relationship arises from an approach to economic theory in which policymaking depends, among other causes, on the formulation of new goals by entrepreneurs and policy–makers alike. This analytical approach would make it possible to redefine economic policymaking in order to stimulate (and, to some extent, help organize)
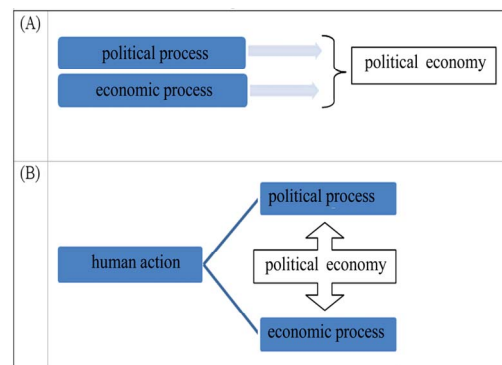

Figure 1. Juxtaposition vs. integration of 'economic' and 'political' processes.

iB

Table 1. Modes of theorizing and policy implications, (*) As special cases, the action plan approach integrates the above types of theorizing, (**) More research is needed

	Main characteristics	Economic policy implications	'entrepreneurship' and 'economic policymaking'
General Equilibrium Theory	Given means, ends and expectations / technology of choice	No room for economic policy	juxtaposition
Neoclassical economics		Welfare economics	
Keynesian economics	Given means, ends & knowledge/ changing expectations	Keynesian economic policy (monetary and fiscal)	
Austrian Economics	Given ends / growing knowledge / changing expectations	Laissez faire	
Evolutionary economics	Learning processes / structural change / but given ends	Industrial, technological and research policy / promoting competition	some degree of integration
Schumpeterian economics		innovation policy; credit policy	
Action plan approach[(*)]	Learning processes / structural change / changing goals	promoting innovation [(**)]	integration

entrepreneurship by, for example, fostering the creativity of individuals. However, the fact that the *specific content* of goals, which depend on the individuals' beliefs, knowledge, experience and intentionality, etc., may produce their effects through the actions of the individuals themselves also implies the recognition of and challenge for policymaking. Agents can devise, create and imagine, etc. new courses of action by forming new spaces of action.

It is in this conception of spaces of action permanently renewed by the entrepreneur's action and the transforming goals he/she wants to achieve, where the new challenges for policymaking are to be found: how to manage processes that promote and channel transforming goals. The policymaker's role should be reconsidered within a context in which the agents' rationality depends on the goals and intentions they pursue and in which it is necessary to channel the innovative goals as a key element for continuous economic change.

From this perspective it would even be possible to state an evolutionary efficiency criterion within an economic system when agents' intentionality is being actualised (materialised) through agents' actions: because of the efficiency of the connections between means/actions and goals, intentions turn out to be actual facts in which goals are being produced. Otherwise, some agents (or all agents) may perceive the fact that they are not achieving their goals, fulfilling their expectations or actualizing (materialising) their intentions as a signal of a certain incompatibility with the actions (i.e. means deployed and timing, but also incompatibilities of goals) carried out and that would merit a (more or less detailed) revision. Thus, bagents may interpret that their action is rationed and, therefore, the performance of the system is below expectations from their own point of view. This kind of judgment applies to policy-makers and entrepreneurs in an economic system. Moreover, an improvement of the performance of the economic processes would require the realignment of agents' goals or, in the terms used in this paper, the revision of the individual intentionality of the agents involved in the system.

5. Concluding Remarks

The mere existence of entrepreneurial forces operating in the economic system is a fundamental and permanent challenge for economic policymaking. It is impossible to conceive any economic system without entrepreneurial action. A first implication of this approach for policymakers is that it is necessary to abandon all mechanicistic approaches to economic policymaking if we want policies that stimulate entrepreneurship, allow the generation of variety and selection (through competitiveness) and retain superior forms of organisation and technology [44]. Necessary conditions include an environment that stimulates creativity, imagination and market experimentation (under a system of true competition), etc. and, of course, a non-confiscatory fiscal policy that does not discourage entrepreneurial efforts and healthy credit and monetary policies that allow economic calculus [10,40] and, therefore, the deployment of entrepreneurial action are. However, these conditions are necessary but not sufficient.

The paradoxes of timeless perspectives (as an analytical basis for the explanation of policymaking applied to processes that necessarily unfold in the time) are resolved by the dynamic opening of the actions and goals pursued by agents in the context of a conception of the economy as 'production of action'. Here, the concepts of (economic and political, etc.) intentionality play a key role.

Of course our contribution does not cover the entirety of such vast issues. The aim of this paper is simply to locate the point of connection between economic poli-

iB

cymaking and entrepreneurship (a favourite concept of Schumpeter) and this is possible under the common focus of a general theory of human action, where, as Mises stressed, economics is the most developed branch. The theoretical and practical implications of such a theory are almost evident. Thus, we feel that our approach to entrepreneurship and its implications for economic policymaking deserve a place in the research agenda of the twenty-first century.

REFERENCES

[1] Witt, U., "Economic policy making in evolutionary perspective," Journal of Evolutionary Economics, Vol. 13, No. 2, pp. 77–94, 2003.

[2] Rubio de Urquía, R., "La naturaleza y estructura fundamental de la teoría económica y las relaciones entre enunciados teórico-económicos y enunciados antropológicos," In Rubio de Urquía, R., Ureña, E. M. and Muñoz, F. F., Eds., "Estudios de Teoría Económica y Antropología," AEDOS-Unión Editorial-IIES Francisco de Vitoria, Madrid, 2005.

[3] Searle, J. R., "The construction of social reality," The Free Press, New York, 1995.

[4] Searle, J. R., "Rationality in action," The MIT Press, Cambridge MA, 2001.

[5] Grosz, B. J. and Hunsberger, L., "The dynamics of intention in collaborative activity," Cognitive Systems Research, Vol. 7, pp. 259–272, 2006.

[6] Metzinger, T. and Gallese, V., "The emergence of a shared action ontology: Building blocks for a theory," Consciousness and Cognition, Vol. 12, pp. 549–571, 2003.

[7] Baldwin, D. A. and Baird, J. A., "Discerning intentions in dynamic human action," Trends in Cognitive Sciences, Vol. 5, No. 4, pp. 171–178, 2001.

[8] Knight, F., "Risk, uncertainty and profit," Houghton Mifflin, Boston, 1921.

[9] Shackle, G. L. S., "Epistemics and economics," Cambridge UP, Cambridge, UK, 1972.

[10] Schumpeter, J. A., "The theory of economic development," Transaction Publishers, Harvard, 1934.

[11] Kirzner, I. M., "The meaning of the market process," Routledge, New York, 1992.

[12] Penrose, E., "The theory of the growth of the firm," Basil Blackwell, Oxford, 1959.

[13] Baumol, W., "Entrepreneurship: Productive, unproductive, and destructive," Journal of Political Economy, Vol. 98, No. 5, pp. 893–921, 1990.

[14] Baumol, W., "The free-market innovation machine," Princeton UP, 2001.

[15] Dopfer, K., "The evolutionary foundations of economics," Cambridge UP, Cambridge, 2005.

[16] M. Casson, "The Entrepreneur," Totowa: NJ, Barnes and Noble Books, 1982.

[17] Casson, M., *et al.*, "The oxford handbook of entrepre-

neurship," Oxford UP, NewYork, 2006.

[18] Ripsas, S., "Towards an interdisciplinary theory of entrepreneurship," Small Business Economics, Vol. 10, pp. 103–115, 1998.

[19] Nelson, R. R., "Bounded rationality, cognitive maps, and trial and error learning," Journal of Economic Behavior & Organization, No. 67, pp. 78–89, 2008.

[20] North, D. C., "Understanding the process of economic change," Princeton UP, Princeton, 2005.

[21] Harper, D. A., "Entrepreneurship and the market process," Routledge, London, 1996.

[22] Loasby, B. J., "Knowledge, institutions and evolution in economics," Routledge, London, 1999.

[23] Hayek, F. A., "The use of knowledge in society," American Economic Review, Vol. 35, pp. 519–530, September 1945.

[24] Loasby, B. J., "The imagined, deemed possible," In Helmstädter, E. and Perlman, M., Eds., "Behavioral norms, technological progress and economic dynamics," University of Michigan Press, Michigan, 1996.

[25] Cañibano, C., Encinar, M. I., and Muñoz, F. F., "Evolving capabilities and innovative intentionality: Some reflections on the role of intention within innovation processes," Innovation: Management, Policy and Practice, Vol. 8, No. 4–5, pp. 310–321, 2006.

[26] Muñoz, F. F., Encinar, M. I., and Cañibano, C., "Is economic evolution blind? The role and dynamics of intended action in economic processes," International Schumpeter Conference, Rio de Janeiro, 2008.

[27] Rubio de Urquía, R., "Estructura fundamental de la explicación de procesos de 'autoorganización' mediante modelos teórico-económicos," In R. Rubio de Urquía, F. J. Vázquez and F. F. Muñoz, Eds., "Procesos de autoorganización," IIES Francisco de Vitoria-Unión Editorial, Madrid, 2003.

[28] Langlois, R. N., "The dynamics of industrial capitalism. schumpeter, chandler, and the new economy," Routledge, London, 2006.

[29] Nelson, R. R., "What enables rapid economic progress: What are the needed institutions?" Research Policy, Vol. 37, No. 1, pp. 1–11, 2008.

[30] Dosi, G., Nelson, R. R., and Winter, S. G., Eds., "The nature and dynamics of organizational capabilities," Oxford UP, Oxford, 2000.

[31] Schumpeter, J. A., "Development," Journal of Economic Literature, Vol. 43, pp. 108–120, 2005.

[32] Encinar, M. I. and Muñoz, F. F., "On novelty and economics: Schumpeter's paradox," Journal of Evolutionary Economics, Vol. 16, No. 3, pp. 255–277, 2006.

[33] Metcalfe, J. S. and Foster, J., "Evolution and economic complexity," Edward Elgar, Cheltenham, 2004.

[34] Loasby, B. J., "Imagination and order," DRUID Summer Conference, 2007, www.druid.dk.

[35] Lachmann, L., "Vicissitudes of subjectivism and the di-

lemma of the theory of choice," In D. Lavoie, Ed., "Expectations and the meaning of institutions, Essays in economics by Ludwig Lachmann," Routledge, London and New York, pp. 218–228, 1994.

[36] Potts, J., "The new evolutionary microeconomics, complexity, competence and adaptive behaviour," Edward Elgar, Cheltenham, 2000.

[37] Shane, S. and Venkataraman, S., "The promise of entrepreneurship as a field of research," The Academy of Management Review, Vol. 25, No. 1, pp. 217–226, 2000.

[38] Loasby, B. J., "The evolution of knowledge: Beyond the biological model," Research Policy, Vol. 31, No. 8–9, pp. 1227–1239, 2002.

[39] Yunus, M., "Vers un monde sans pauvreté," Jean-Claude Lattès, Paris, 1997.

[40] Mises, L., "Human action: A treatise on economics," Yale UP, New Haven, 1949.

[41] Pelikan, P., "Why economic policies need comprehensive evolutionary analysis," In P. Pelikan and G. Wegner, Eds., "The evolutionary analysis of economic policy," Edward Elgar, Cheltenham, 2003.

[42] Simon, H., "Reason in human affairs," Stanford University Press, Stanford, 1983.

[43] Hodgson, G. M., "Economics & Utopia," Routledge, London, 1999.

[44] Metcalfe, J. S., "The entrepreneur and the style of modern economics," Journal of Evolutionary Economics, Vol. 14.

iB

◆❖ Scientific
◆❖ Research

Economic Freedom and Foreign Direct Investment: How Different are the MENA Countries from the EU[1]

José CAETANO, António CALEIRO

Departamento de Economia & CEFAGE-UE, Universidade de Évora, Évora, Portugal.
Email: jcaetano@uevora.pt, caleiro@uevora.pt

ABSTRACT

The risk perceived by investors is crucial in the decision to invest, in particular when it concerns a foreign country. The risk associated to any (foreign) investment is a multi-faceted element given that it reflects many aspects that are relevant to (foreign) investors, such as the level of transparency, corruption, rule of law, governance, etc. In this paper we consider the level of economic freedom, as provided by the "Heritage Foundation", for the most recent years, in order to analyse how is this measure of risk related to the inward foreign direct investment performance index, as provided by the UNCTAD. Given the subjectivity of risk an appropriate methodology consists on using fuzzy logic clustering, which is applied in the paper in order to verify how different the MENA region is from the set of EU-member states. The results show that economic freedom and inward FDI are positively associated, in particular in the cluster of countries that present a higher economic freedom. Of particular interest is the result that some MENA countries belong to the same cluster of most of the EU-countries.

1. Introduction

Since the 1990s the literature has been paying more attention to the importance of the quality of institutions and of economic freedom for the countries economic development. Economic freedom means the degree to which a market economy is in place, where the central components are voluntary exchange, free competition, and protection of persons and property. O'Doriscoll *et al.* [1] define economic freedom as "the absence of government coercion or constraint on the production, distribution or consumption of goods and services beyond the extent necessary for citizens to protect and maintain liberty itself". In these terms, the economic freedom could be a key factor accounting for economic growth [2].

The incentives that economic actors face are determined in large part by the institutions in place, which can

be more or less efficient. Furthermore, sustained high growth rates imply eventually great wealth, and so in the long term the economic freedom that increases growth can also be expected to increase wealth. Despite this fact, there are theoretical reasons to expect a positive relation between economic freedom and economic growth but does empirical evidence confirm this link?

A number of studies have corroborated those expectations, with varying strengths and in different forms. For instance, Adkins *et al.* [4] find that the level of economic freedom at the beginning of the growth period does not contribute significantly to explaining growth, but that positive changes in economic freedom do so. Yet, other studies conclude that the initial level of economic freedom is also positively related to growth [5]. In any case, the issues included in economic freedom should be taking into account once policies try to promote economic development.

Since the share of the developing countries in the global foreign direct investment (FDI) flows has been Rising[2], there is also a growing interest in study the de-

---

[1]The suggestions and remarks of an anonymous referee are gratefully acknowledged.

[2]From 1980 to 2004, the share of developing countries in the world FDI flows evolved by 25% to 44%, which means a significant evolution. However, data for 2007 displayed an erosion of this share for 27.3% [3].

terminants of this kind of flows. In fact, the literature has accepted that FDI can provide additional resources for developing countries, by which they could improve their economic performance and factor productivity, through the diffusion of technological progress and the boost of domestic investment [6].

In this context some studies have analyzed the importance of economic freedom in the FDI performance in developing countries, especially in what concerns some aspects of a country's trade policy, its banking and finance services and its property right protection [7]. Likewise, Gwartney *et al.* [8] suggested that the key ingredients to economic freedom include freedom to compete, voluntary exchange, and protection of person and property.

In order to uncover the factors that matters for foreign investment flows, it is necessary to distinguish the following types of investment: market seeking; resource seeking; efficiency seeking. Thus, the new wave of globalization has led to a reconfiguration of the ways in which multinationals pursue these various types of FDI, and changed the motives for investing abroad. Dunning [9] sustain that the FDI in developing countries has been shifting from market and resource seeking investments, to more efficiency seeking investments. Some authors argue that the relative importance of the traditional market related factors (wage costs, infrastructure or macroeconomic policy) no longer hold and suggest that less traditional determinants have become more important, like institutions or economic freedom [10].

The paper considers the Middle East North Africa (MENA) countries, vis-a-vis the European Union (EU) ones, for the following reasons. FDI flows to the MENA region have been relatively low when compared to the EU and to other developing and emerging countries [11]. Some characteristics of the MENA countries could entail an important constraint for the inward FDI performance. This region is highly anchored on oil, which weakens the economic base, has high unemployment rates, displays a weak regional economic integration and the capital and financial markets persist undeveloped.

Yet, some countries in the region are witnessing a new era in privatization, bank regulation and market-oriented financial institutions, making the need to look at the role of other determinants even more pertinent. The analysis of MENA institutional systems that influence economic freedom appears to be attractive since a significant number of this countries is been experiencing institutional reforms. Moreover, the Euro-Mediterranean Partnership Agreement, along with the progressive elimination of trade barriers, has boosted trade relations and some countries have liberalized their investment regulatory framework, creating particular regimes for FDI. Taking

into account these facts and the relatively scarce empirical research on FDI in MENA countries, we consider being important to study this subject.

There is a vast literature on the determinants of FDI and the empirical studies differ in terms of the variables, methodologies, the type of FDI and the countries included. In general, variables affecting the FDI flows can be classified into two categories: market-oriented variables and institutional-oriented variables. In this study, our emphasis is the institutional-oriented variables, especially in the economic freedom issues.

In this paper we use the Index of Economic Freedom, provided by the Heritage Foundation, for the most recent years, in order to analyse how is this measure related to the Inward Foreign Direct Investment Performance Index, as provided by the UNCTAD. Given the subjectivity of economic freedom, an appropriate methodology consists on using fuzzy logic clustering [12], which is applied in the paper in order to verify how different is the MENA region from the set of EU-member states. The aim is to investigate whether there are region-specific factors in the economic freedom that are significant for FDI performance. The accomplishment of this objective adds to the literature given the methodology that is applied in the paper.

The rest of the paper is structured as follows. Section 2 briefly presents the empirical literature linking economic freedom issues and inward FDI, emphasizing the research on the MENA countries. Section 3 comments some descriptive statistics on the economic freedom and foreign direct investment in the MENA region and in the EU members. Section 4 explains the methodological aspects related to the data and the fuzzy logic technique analyses how economic freedom exerts influence on the FDI. Given that a certain level of perceived economic freedom can, in fact, be subject to different subjective evaluations by investors, the paper uses a fuzzy logic approach in order to determine conceivable clusters in the space economic freedom-FDI, which is done in Section 5. Section 6 presents some concluding remarks.

2. Literature Review

The literature is mainly dedicated to study the impact of the economic freedom on inward FDI flows. For a number of reasons, the transparency in economic policies is an essential issue for investors, especially for foreigners. The lack of these conditions imposes extra costs to the firms, linked to the lack of information about activities or even future intentions of some governmental departments. Thus, the selection of investment location is, sometimes, biased for the presence of non-economic elements. So, a steady and actively legal framework against the corruption and promoting economic freedom can, in fact, represent a factor of attractiveness for FDI.

The positive interaction between economic freedom and FDI atractiveness is due to, in the first place, the fact that free markets promote a better factor allocation and stimulate the productivity and the investments profitability. In the second place, since FDI involve significant sunk costs (particularly the greenfield), investments become very sensitive to the degree of stability and security offered by the legal protection system of the intellectual property rights. So, the existence of clear and predictable economic policies related to liberalizing regimes of investment and trade can be powerful instruments in the way to attract FDI flows [13].

The OLI paradigm [14] is a milestone reference in the theoretical and empirical approaches of the FDI. This paradigm sustains that firm decisions in relation to foreign markets depends on the economic and institutional conditions in home and host countries. In concrete, the decision to invest in a foreign country needs the firm boast, simultaneously, three types of advantages: ownership (O), location (L) and internalization (I). The ownership advantages reveal to be a basic condition for that the firm explore it in any market. Also, the choice of the location is conditional on the existence of structural market imperfections or from specific factor endowments, being mostly relevant the risk that firm incurs when dislocating to an unknown market. Finally, firms internalize their own markets of intermediate goods, whenever the costs of transaction in the markets surpass the coordination costs that the company supports for the internal accomplishment of this type of activities.

Later, the new concept of "capitalism of alliances", based in the mutual trust, commitments and the contractual obligations between partners, widens the original scope of the OLI Paradigm [15]. In this sense, reciprocal trust may be a key instrumental issue for the firms' potential success. The inclusion of economic freedom issues turned to be considered in an explicit form, given its impacts on the confidence level of the agents (see [16]). This Paradigm has been important to understand the multinationals behaviour, its usefulness being able to be strengthened by the inclusion of the freedom and its impacts on FDI. In fact, this issue basically affects the location dimension and it motivates firms to reduce the degree of uncertainty associated with its entrance in a foreign market.

The linkages between FDI flows and political risk and institutions are explored by Busse and Carsten [17] for a large sample of 83 developing countries, taking into ac-

count 12 different indicators for the period 1984 to 2003. They found that the investment profile, internal and external conflict, ethnic tensions and democratic accountability are significant determinants of FDI flows. Across different econometric models, the relative magnitude of the coefficients for these political indicators are largest for government stability and law/order, suggesting that changes in these components are greatly relevant for investment decisions of multinationals.

A more recent study is provided by Dumludag *et al.* [19], who investigate the relationship between FDI flows and institutions in several emerging markets, em- ploying a panel data approach from 1992 to 2004. The socio-political variables include juridical system, corruption, investment profile, government stability, economic, social and political risks. Those authors wrap up that institutional variables are important, particularly corruption, investment profile and government stability.

Despite those approaches, the impact of institutional differences between the home and the host countries has been little researched so far. Yet, in a recent study, using a database provided by the French Ministry of Finance network in 52 countries and the Fraser Institute database, Bénassy-Quéré *et al.* [20] examine the role of institutions in the both host and source country by estimating a gravity equation for bilateral FDI stocks that includes governance indicators. The analysis provides abundant evidence to carry on the idea by which institutions do matter whatever the countries development level. In fact, results show that inward FDI is positively affected by public efficiency, which includes tax system, transparency and lack of corruption, security property rights and the easiness to create a business.

In sum, literature recognizes the importance of institutional variables in empiric studies, providing support for the idea that an efficient legal and social framework promotes economic freedom and reduces uncertainties. So, most of the studies conclude that the protection of intellectual property rights, low corruption levels, enforcement mechanisms and political stability influences positively the FDI inward flows and the economic growth[3]. In fact, when these conditions do not exist in a country, foreign investors can face particularly high costs in establishing an operation and inhibit FDI inflows.

Despite the lack of research on determinants of FDI in MENA countries, recent studies have analysed this issue by using different methodologies and data sets. All these studies share the idea that FDI for these countries is low when compared with other developing countries. In addition, most of them concentrate on the importance of the institutional issues for the FDI inflows in these countries, concluding that institutions are vital to explain the poor performance of the MENA region in attracting FDI.

An early analysis is performed by Kamaly [21], who uses a dynamic panel model for the period 1990 to 1999.

---

[3]Literature has also been paying attention to the relationship between economic freedom and corruption. Graeff and Mehlkop [18] identify a stable pattern of aspects of economic freedom influencing corruption that differs depending on whether countries are rich or poor. So, despite there is a strong relation between economic freedom and corruption, this relation depends on a country's level of development and, contrary to what is expected, they find that some types of regulation reduce corruption.

iB

In this study, economic growth and the lagged value of FDI/GDP were the only significant determinants of FDI flows to the MENA region. However, this approach, as are most other studies on FDI in developing countries, does not cover the recent period and uses a small sample, thus raising questions about the consistency and efficiency of the coefficients of the dynamic model. Also, it does not consider the institutional factors that affect FDI flows to the MENA region.

By using a fixed effects panel data model for the period 1975 to 1999, Onyeiwu [22] compares 10 MENA countries with other developing countries, including in the study institutional aspects that may affect FDI flows to the region. He concludes that corruption is, in general, significant for all the developing countries and, in the case of the MENA countries; it is the only significant variable in explaining FDI inflows. However, the author uses government expenditure over GDP as proxy for corruption, which might not be the appropriate measure for this variable.

Chan and Gemayel [23] study the relation between macroeconomic instability and FDI in the MENA region. They employ two dynamic panel data models using two groups: one with 19 MENA countries and the other with 14 EU countries as well as Canada and USA. Their results show that the instability has a much stronger impact on FDI than risk itself, being this particularly important for the MENA region. However, the study suffers from the weak consistency of the coefficients in the dynamic models, because the sample data is not large enough to be confident on the results and the applied estimation methods are not the appropriate ones for obtaining consistent estimates in a dynamic panel data model.

Other assessment on the influence of quality of institutions on trade and FDI in MENA countries is developed by Méon and Sekkat [24], who includes data from 1990 to 1999, covering a large number of countries, including some MENA countries. They use some proxies the quality of institutions, namely corruption, political risk and governance. The results show a significant relationship between political risk and inward FDI, but failed to find clear evidence of a significant relationship between corruption and FDI flows. In fact, they employ different indicators of corruption and conclude that the results are sensitive to the index used to measure corruption. In the same line, applying the Kaufmann *et al.* [25] governance indexes, Daniele and Marani [26] look into the role of the quality of institutions on FDI, through a cross sectional regression analysis for 129 countries to the period 1995-2004, concluding that institutions are crucial to explain the performances of countries in attracting FDI.

Kobeissi [27] performed a testing on the impact of some non-traditional factors on foreign investment in MENA countries, focusing on factors such as governance,

legal environment, and economic freedom, based on the indicators provided by the Heritage Foundation. The results reveal a consistent support for the positive impact of governance, legal system and economic freedom on the FDI flows in the MENA region, but the governance showed the most significant results followed by legal system and then economic freedom. The relatively lower importance of the last two variables could be due to the fact that investors from different countries have varying degrees of tolerance for imperfections in the host country's investment environment.

Ferragina and Pastore [28] examines FDI flows from the EU to two neighbouring regions: Central and Eastern Europe (CEE) and South Mediterranean (MED) countries (including some MENA countries), to verify whether there was any diversion effect on FDI flows following the CEE integration in the EU. They use a gravity type model and a panel data approach to study the determinants of bilateral FDI flows for the period 1994-2004. Among the explanatory variables are included some institutional and economic freedom issues. They conclude that there is no evidence of FDI diversion, but results also highlight that governance is highly significant with positive sign and the current and capital account restrictions are both negative and highly significant.

Finally, in a fresh study, Onyeiwu [11] uses a logit and cross-country regressions, for 61 MENA and non-MENA countries, to examine whether scarce in- vestment in knowledge, technology, and human capital by MENA countries explains their sub-optimal FDI profile. Results from both models suggest that investment in knowledge and technology is not significant for the MENA country's ability to attract an optimal level of FDI. To the contrary, openness of the economy, GDP per capita and political risks are more important to attract this kind of flows. So, one implication for MENA countries is that, despite their poor science and technology infrastructure, they could still attract FDI by promoting openness and political rights and civil liberties.

3. Discussion of the Data

Before presenting the methodological issues used in the paper we make a brief presentation of the variables included in that component and we will make an empirical analysis of trends observed over the period. In what concerns the FDI data, we use the inward FDI performance index provided by UNCTAD for the period 1999-2001 to 2004-2006, which ranks countries by the FDI they receive relative to their economic size. It is the ratio of a country's share in global FDI inflows to its share in global GDP, that is $IND_i = \dfrac{\frac{FDI_i}{FDI_w}}{\frac{GDP_i}{GDP_w}}$ [4]. Thus, a value greater than 1 indicates that the country receives more FDI than its relative economic size, a value below 1

means that it receives less. The index thus captures the influence on FDI of factors other than market size, assuming that, other things being equal, economic size is the base line for attracting investment.

In this study we apply the Index of Economic Freedom provided by Heritage Foundation for 162 countries, for measuring economic freedom, which included 50 independent variables which fall into 10 categories of economic freedom. Each country receives its overall economic freedom score based on the simple average of the 10 individual factor score. Each factor is graded according to a unique scale, which runs from 1 to 5, where a score of 1 indicates an economic environment that are most conducive to economic freedom and a score of 5 signifies the opposite. The 10 variables included in the overall index are the follows[5]:

- *Business freedom* is the ability to create, operate, and close an enterprise quickly and easily. Burdensome, redundant regulatory rules are the most harmful barriers to business freedom.
- *Trade freedom* is a composite measure of the absence of tariff and non-tariff barriers that affect imports and exports of goods and services.
- *Government size* is defined to include all government expenditures, including consumption and transfers. Ideally, the state will provide only true public goods, with an absolute minimum of expenditure.
- *Investment freedom* is an assessment of the free flow of capital. This factor scrutinizes each country's policies toward foreign investment, as well as its policies toward capital flows internally, in order to determine its overall investment climate.
- *Property rights* are an assessment of the ability of individuals to accumulate private property, secured by clear laws that are fully enforced by the state.
- *Freedom from corruption* is based on quantitative data that assess the perception of corruption in the business environment, including levels of governmental legal, judicial, and administrative corruption.
- *Labour freedom* is a composite measure of the ability of workers and businesses to interact without restriction by the state.
- *Financial freedom* is a measure of banking security as well as independence from government control;

state ownership of banks and other financial institutions such as insurer and capital markets is an inefficient burden, and political favouritism has no place in a free capital market.
- *Fiscal freedom* is a measure of the burden of government from the revenue side and it includes both the tax burden in terms of the top tax rate on income and the overall amount of tax revenue as a portion of GDP.
- *Monetary freedom* combines a measure of price stability with an assessment of price controls, because both inflation and price controls distort market activity.

Data on these variables are presented in annex and the brief analysis of its trends allows us to emphasize the following points:

- Regarding the Inward FDI *Performance Index* we note that the EU presented an atractiveness clearly superior to the MENA region, with the average values in the range of 7 triennia because the EU almost double the figure recorded by MENA countries (2.12 and 1.23, respectively). However, when comparing the evolution between 1999-01 to 2005-07, the average value of that indicator for the MENA countries ore than quintupled (from 0.39 to 1.99), while for the EU growth was only 9% (from 1.94 to 2.12). As a result of such trends over the last period (2005-07) the average values of the two groups were approximated, showing the two regions as very attractive in world terms of attracting FDI flows.
- For the Index of Economic Freedom we found that the average of the period (1999-05) in the EU was around 14% higher than the recorded value in the MENA region, meaning that this gap has widened over that period, rising from 7.8% to 17.4% between 1999 and 2005, respectively. Although changes in the values for the two regions are not very significant in this period, we note that the average of the EU improved slightly (4.2%) and, in an opposite trend, worsened in the MENA region (-1.1%). We also note that the performance of countries within each group was very different and, in particular in the MENA, the dispersion was very significant, indicating the existence of very different situations as far as promoting economic freedom.

In summary, we believe that the dispersion found in the variables within the two groups over the period reflects a high diversity of countries performances in order to attract FDI flows and in promoting an economic freedom environment[6]. This is especially evident within the MENA countries where very different economic and institutional realities coexist. In fact, a number of countries in the region have paid special attention to making themselves investor-friendly by making the business en-

---

[4]Where IND_i is the inward FDI performance index of the i-th country, FDI_i is the FDI inflows in the i-th country, FDI_w is the world FDI inflows, GDP_i is the GDP in the i-th country and GDP_w is the world GDP.

[5]For a detailed information, see the document Methodology: Measuring the 10 Economic Freedom, disponível no site da Heritage Foundation. http://www.heritage.org/research/features/index/chapters/pdf/Index2008_Chap4.pdf.

[6]For example, some countries in the Persian Gufl (Bahrain and United Arab Emirates), Jordan and Lebanon have been revealing in recent years a high capacity to attract FDI flows. Interestingly, these countries, with the exception of Lebanon, are in the group that presents a higher position in relation to index of economic freedom.

vironment more open and stepping up structural and institutional reforms, while others have been following other paths.

4. The Fuzzy Logic Approach

Given that, we think the purposes of the paper will be mainly achieved by the use of fuzzy clustering techniques, it is informative to start with a general discussion of this kind of approach.

Following the logic of crisp sets, the degree to which an element belongs to a set is either 1 or 0, by that meaning that the characteristic function discriminates respectively between members and non-members of the set in a crisp way. The generalisation to a fuzzy set is made by relaxing the strict separation between elements belonging or not to the set, allowing the degree of belonging/membership to take more than these two values, typically by allowing any value in the closed interval [0,1] (see, for instance, [29,30]).

The values then assigned by the membership function of a fuzzy set to the elements in the set indicate the membership grade or degree of adherence of each element in the set. Larger (smaller) values naturally indicate higher (lower) membership grades, degrees, or consistency between an element of the set and the full characteristics that the set describes. Hence, using fuzzy logic, one can deal with reasoning like: 'the observed value for the economic freedom index, say 5, can be considered high, normal or low with some degrees of membership'.

In terms of fuzzy logic, 'high', 'normal' or 'low' values (for the variable under question) can be considered to be subjective categories, as economic agents often evaluate those concepts differently. In what follows, it will be assumed that investors consider to be relevant their relative perception of economic freedom (in accordance to some subjective categories) for their willingness to invest, therefore assuming an approximate or qualitative reasoning.

In the particular case of this paper, we will use this kind of fuzzy logic reasoning to construct clusters in the space (FDI, Economic Freedom). This partition of the space can also be done in, say, a traditional/crisp way. The crisp/hard clusters algorithm tries to locate clusters in a multi-dimensional data space, U, such that each point or observation is assigned in that space to a particular cluster in accordance a given criterion. Considering c clusters, the hard cluster technique is then based on a c-partition of the data space U into a family of clus-

ters such that the set of clusters exhausts the whole universe, that a cluster can neither be empty nor contain all data samples, and that none of the clusters overlap.

Formally, the hard c-means algorithm finds a centre in each cluster, minimising an objective function of a distance measure. The objective function depends on the (Euclidean) distances between data vectors u_k ($k = 1, 2,…, K$) and cluster centres c_i. The partitioned clusters are typically defined by a $c \times K$ binary characteristic matrix $\mathbf{M}$, called the membership matrix, where each element m_{ik} is 1 if the kth data point u_k belongs to cluster i, and 0 otherwise. Since a data point can only belong to one cluster, the membership matrix $\mathbf{M}$ has the properties: (i) the sum of each column is one, and (ii) the sum of all elements is K.

The fuzzy c-means differs from hard c-means because it employs fuzzy partitioning, where a point can belong to several clusters with degrees of membership such that the membership matrix $\mathbf{M}$ is allowed to have elements in the range [0,1]. A point's total membership to all clusters, however, must always be equal to unity. In this sense, and despite that, in formal terms, none of the fuzzy clusters overlap, the fact is that, in general, each data point is assigned to every cluster, although with different degrees of membership. Generally speaking, in visual terms, each data point is then associated to the particular cluster to which its degree of membership is higher.

5. How Different are the MENA Countries from the EU in Terms of the Relationship between Economic Freedom and FDI

In this section we analyse a possible influence of economic freedom on FDI. Figures 1 and 2 plot the data and at the same time show the results from the fuzzy clustering technique[7].

Plainly, there are two well-defined clusters (identified in figures 1 and 2 by the dotted circles and empty circles, whose centres are given by the black crosses), one being associated with the higher level of perceived economic freedom countries and another associated with the lower level perceived economic freedom countries. In fact, the splitting of the countries clearly reflects the economic freedom values as it seems possible to separate the two groups of countries in accordance to a, say, critical level of perceived economic freedom around 52 for the first period under analysis and 61 in the second period under analysis[8].

Since the observed similarity between the results for the two periods, one can assert the robustness of the results. In fact, the clusters indicate a relationship between Economic Freedom and the Inward Performance of FDI, that is assumed to be of causal nature given the theoretical support of it. The results point to the fact that, in

---

[7]The data can be consulted in the annex. The source of the economic freedom data is the Heritage Foundation (http://www.heritage.org/research/features/index/downloads.cfm) and the source of the FDI performance index is the UNCTAD (http://www.unctad.org/Templates/WebFlyer.asp?intItemID=2471&lang=1).

[8]This result makes it quite easy to identify the countries in each cluster (see the annex).

Figure 1. The results for 1999/2001

Figure 2. The results for 2005/2007

6. Concluding Remarks

The results of the paper show that economic freedom and inward FDI are positively associated, in particular in the cluster of countries that present a higher economic freedom. Of particular interest is the result that some MENA countries belong to the same cluster of most of the EU-countries.

To conclude we would like to stress the main lesson from our paper as a policy implication. In order not to be considered less attractive for foreign investors and, therefore, be penalised by that, countries do indeed benefit from increased levels of transparency in order to escape from the cluster of countries where perceived levels of economic freedom are smaller. In other words, policy makers should make sure that their policies are transparent enough for potential foreign investors. After escaping from that cluster, the objective of attracting higher levels of FDI has to be crucially obtained by the use of other measures.

In the context of Dunning's framework, we could understand the results of our empirical research as supporting the inclusion of economic freedom in the set of the relevant elements for the location tier [14,15].

Given that (perceived) economic freedom reflects a variety of factors–which are clear even in the way the economic freedom data is obtained – an interesting issue to be further explored is the analysis of the specific factors or components that assume a more significant role on the attraction of FDI.

An analysis of the dynamics of the components of economic freedom or even of economic freedom itself seems to be a quite plausible improvement as the direction assumed by policy makers towards more transparent policies may have a marginal impact on the attraction of FDI much more evident than one may expect by the analysis of the absolute position of economic freedom. Straightforwardly, the more those measures are assumed to be credible by foreign investors, the more that can be the case.

Finally, we consider this paper as a promising starting point for the analysis of the factors that reveal to be essential for FDI, either in an inward perspective or in an outward perspective, both in performance and potential measures. The combination of all these perspectives, in a dynamic way, is to be considered in future studies. As a matter of fact, this kind of analysis can easily be extended to other set of countries that show some empirical support for the existence of a relationship between institutional factors and investment decisions. Furthermore, the inclusion of geographical factors, in what concerns the localization of the host countries and of investors, in those dynamics is also in our mind as relevant elements.

overall terms, there is a direct relationship between Economic Freedom and the Inward Performance of FDI. This relationship is apparently stronger in the cluster of countries with higher economic freedom. In fact, as the level of economic freedom is decisive in the clustering, the overall increase that could be observed in the economic freedom from 1999/2001 to 2005/07 – which can be noted at the centre of the clusters in the two periods – led to a more homogeneous, from the point of view of the number of countries, clustering in that last period. Consequently, whereas in the first period, 4 out of the 6 countries in cluster 1 were MENA countries, in the second period, 10 out of the 14 countries in cluster 1 were MENA countries, despite the general increase in economic freedom and FDI inward performance that these countries registered from 1999/2001 to 2005/07.

REFERENCES

[1]   O'Doriscoll, *et al.*, "Index of economic freedom," Heritage Foundation and Wall Street Journal, NewYork, 2001

[2]   Berggren, N., "The benefits of economic freedom—A survey," The Independent Review, Vol. 8, No. 2, pp. 193–211, Fall 2003, http://www.ratio.se/pdf/wp/nb_efi.pdf. 2003.

[3]   UNCTAD, "World investment report 2008: Transnational corporations and the infrastructure challenge," Geneva, UNCTAD-United Nations, 2008, http://www.unctad.org/Templates/WebFlyer.asp?intItemID=4629&lang=1.

[4]   Adkins, L. *et al.*, "Institutions, freedom, and technical efficiency," Southern Economic Journal, No. 69, pp. 92–108, July 2002.

[5]   Weede, E. and Kämpf, S., "The impact of intelligence and institutional improvements on economic growth," Kyklos, Vol. 55, No. 3, pp. 361–80, 2002.

[6]   Sekkat, K. and Veganzones-Varoudakis, M., "Openness, investment climate, and fdi in developing countries," Review of Development Economics, Vol. 11, No. 4, pp. 607–620, 2007.

[7]   Globerman, S. and Shapiro, D., "Governance infrastructure and U.S. foreign direct investment," Journal of International Business Studies, Vol. 3, pp. 19–39, 2003.

[8]   Gwartney, J., *et al.*, "Economic freedom of the world: 2003 annual report," The Fraser Institute, Calgary, 2003, http://www.freetheworld.com/release_2003.html

[9]   Dunning, J., "Determinants of foreign direct investment globalization induced changes and the roles of FDI policies," Annual World Bank Conference on Development Economics, Europe 2002–2003: Toward Pro-Poor Policies-Aid, World Bank Publications.

[10] Becchetti, L. and Hasan, I., "The effects of (within and with EU) regional integration: Impact on real effective exchange rate volatility, institutional quality and growth for MENA countries, World Institute for Development Economics Research of the United Nations University, Research Paper No. 2005/73, 2004, http://www.wider.unu.edu/publications/working-papers/research-papers/2005/en_GB/rp2005–73/

[11] Onyeiwu, S., "Does investment in knowledge and technology spur 'optimal' FDI in the MENA region? Evidence from logit and cross-country regressions," paper presented on 2008 African Economic Conference, promoted by The African Development Bank Group, November, Tunis, 2008, http://www.afdb.org/portal/page?_pageid=473,30752696&_dad=portal&_schema=PORTAL .

[12] Caetano, J. and Caleiro, A., "Corruption and foreign direct investment: What kind of relationship is there?" in R. Gupta, S. S. Mishra, The Causes and Combating Strategies, ICFAI Books, 2007.

[13] Drabek, Z. and Payne, W., "The impact of transparency on foreign direct investment," World Trade Organization —Economic Research and Analysis Division, Working Paper ERAD-99-02, 2001, http://www.wto.org/english/res_e/reser_e/wpaps_e.htm.

[14] Dunning, J., "The eclectic paradigm of international production: A restatement and some possible extensions," Journal of International Business Studies, Vol. 19, No. 1, pp. 1–31, 1988.

[15] Dunning, J., "Reappraising the eclectic paradigm in the age of alliance capitalism," Journal of International Business Studies, Vol. 26, No. 3, pp. 461–491, 1995.

[16] Voyer, P. and Beamish, P., "The effect of corruption on Japanese foreign direct investment," Journal of Business Ethics, Vol. 50, No. 3, pp. 211–224, 2004.

[17] Busse, M. and Carsten, H., "Political risk, institutions and foreign direct investment," HWWA Discussion Paper 315, 2005, http://www.hwwa.de/Forschung/Publikationen/Discussion_Paper/2005/315.pdf.

[18] Graeff, P. and Mehlkopb, G., "The impact of economic freedom on corruption: Different patterns for rich and poor countries," European Journal of Political Economy, Vol. 19, pp. 605–620, 2003.

[19] Dumludag, D., *et al*., "Determinants of foreign direct investment: An institutionalist approach," Seventh Conference of the European Historical Economics Society, Lund University, June 2007, http://www.ekh.lu.se/ehes/paper/devrim_dumludag_EHES2007_paper_new.pdf.

[20] Bénassy-Quéré, *et al.*, "Institutional determinants of foreign direct investment," The World Economy, Vol. 30, No. 5, pp. 764–782, 2007.

[21] Kamaly, "Evaluation of FDI flows into the MENA region," The Economic Research Forum, Working Paper Series, Cairo, 2002, http://www.erf.org.eg/CMS/getFile.php?id=694.

[22] Onyeiwu, S., "Analysis of FDI flows in developing countries: Is the MENA region different?" Selected Published Papers from the Tenth Annual Conference of the Economic Research Forum, Cairo, Egypt, pp. 165–182, 2004, http://www.mafhoum.com/press6/172E11.pdf.

[23] Chan, K. and Gemayel, E., "Risk instability and the pattern of foreign direct investment in the middle east and north africa region," International Monetary Fund, IMF Working Paper 04-139, 2004. http://www.imf.org/external/pubs/ft/wp/2004/wp04139.pdf.

[24] Méon, P. and Sekkat, K., "Does the quality of institutions limit the MENA's integration in the world economy?" The World Economy, Vol. 27, No. 9, pp. 1475–1498. 2004.

[25] Kaufman, D., *et al*., "Governance matters IV: Governance indicators for 1996-2004," World Bank Policy Research Working Paper Series No. 3630, Washington, D.C., 2005, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=718081.

[26] Daniele, V. and Marani, U., "Do institutions matter for FDI? A comparative analysis for the MENA countries," MPRA paper No. 2426, 2007, http://mpra.ub.uni-muenchen.de/2426/1/MPRA_paper_2426.pdf.

[27] Kobeissi, N., "Impact of governance, legal system and economic freedom on foreign investment in the MENA region," Journal of Comparative International Management, Vol. 8, No. 1, 2005.

[28] Ferragina, A. and Pastore, F., "FDI potential and short-falls in the med and ceec´s: determinants and Diversion Effects," paper presented at XVI Conferenza Scientifica Nazionale AISSEC, Parma, 2006, https://dspace-nipr.cilea.it/bitstream/1889/886/1/Ferragin a_Pastore_Aissec_07.pdf.

[29] Zimmermann, H. J., "Fuzzy set theory—and its applications," 2nd Edition, Kluwer Academic Publishers, Boston, 1991.

[30] Chen, C., "Fuzzy logic and neural network handbook," McGraw-Hill, New York, 1996.

Annex: The data

Country	Economic Freedom score		Inward performance FDI index	
	1999	**2005**	**1999/2001**	**2005/2007**
Algeria	57.2	52.7	0.168	0.466
Austria	64.0	67.5	0.362	1.706
Bahrain	75.2	71.2	0.302	4.178
Bulgaria	46.2	61.6	0.197	7.240
Cyprus	67.8	71.9	0.254	3.022
Czech Republic	69.7	64.9	0.259	2.169
Denmark	68.1	75.3	0.394	1.176
Egypt	58.0	56.4	0.184	2.948
Estonia	73.8	75.1	0.262	4.869
Finland	63.9	71.5	0.417	1.033
France	59.1	61.2	0.403	1.659
Germany	65.6	69.0	0.436	0.589
Greece	61.0	58.5	0.280	0.341
Hungary	59.6	63.2	0.258	1.991
Iran	36.8	48.6	0.206	0.098
Ireland	74.6	80.6	0.425	-0.353
Israel	68.3	62.4	0.361	2.441
Italy	61.6	64.9	0.342	0.624
Jordan	67.4	65.7	0.260	5.758
Kuwait	69.5	64.8	0.299	0.059
Latvia	64.2	66.4	0.210	2.585
Lebanon	59.1	58.0	0.209	4.386
Libya	32.3	32.8	0.267	1.372
Lithuania	61.5	70.5	0.203	1.838
Malta	59.3	68.9	0.282	6.372
Morocco	63.8	52.6	0.150	1.212
Netherlands, The	63.6	72.3	0.434	2.689
Oman	64.9	66.6	0.226	1.909
Poland	59.6	58.8	0.249	1.587
Portugal	65.6	62.9	0.286	1.241
Qatar	62.0	63.5	0.407	0.571
Romania	50.1	51.9	0.150	2.566
Saudi Arabia	65.5	63.8	0.309	1.877
Slovak Republic	54.2	65.6	0.232	1.903
Slovenia	61.3	60.1	0.309	0.797
Spain	65.1	67.7	0.349	0.996
Sweden	64.2	69.5	0.429	1.604
Syria	39.0	46.4	0.152	0.734
Tunisia	61.1	54.8	0.180	2.157
Turkey	59.2	51.6	0.155	1.117
United Arab Emirates	71.5	65.5	0.392	2.498
United Kingdom	76.2	79.3	0.472	2.661
Yemen	43.3	52.9	0.158	0.821

◆◆ Scientific
◆◆ Research

Food Safety and the Role of the Government: Implications for CSR Policies in China

Rosita DELLIOS[1], Xiaohua YANG[2*], Nadir Kemal YILMAZ[3]

[1]Associate Professor of International Relations, Bond University, Gold Coast, Australia; [2]School of Business and Management, University of San Francisco, San Francisco, United States; [3]Investment Professional, Fiskiye Sok, Ankara, Turkey; *The Corresponding Author.
Email: rosita_dellios@bond.edu.au, xyang14@usfca.edu, nadirkemal@mac.com

ABSTRACT

This study investigates food scandals and the role of government in corporate social responsibility (CSR) in the food industry and explores strategies for the Chinese government to tackle the food safety problems that abound in China. Based on the theoretical discussion of four types of CSR and the empirical evidence from four case studies, we argue that government influence on CSR in the food industry is determined by the intensity and salience of its own behavior and actions including regulations. We further believe that a balanced CSR strategy covering economic, legal, ethical and philanthropic considerations would work best for China. Our contributions include extending the CSR literature to the food industry and emerging economies like China and recognizing the distinctive role the government plays in the food industry. In addition, we provide a timely guide to establishing a food safety system in China.

Keywords: *Corporate Social Responsibility, China, Food Safety, Emerging Market*

1. Introduction

Food safety is an issue that remains unresolved in both the developed and developing world. Barely having recovered from the shock of "mad cow disease" emanating from the UK in the 1990s, the 2008 "tainted milk" crisis in China serves as a reminder that the problem of food safety has not been contained or adequately addressed. The World Health Organization [1] reports a rise of 30 percent in the number of people in developed countries who become ill from foodborne diseases each year. Smith and Riethmuller [2–4] provide numerous other examples that show foodborne diseases do not discriminate between rich and poor countries with many cases having occurred in Industrialized economies: these range from Japan's 1996 radish sprouts food poisoning incident that resulted in 10 deaths and 9,000 people being ill, to the "Arnotts Biscuits poisoning, the Australian peanut paste products affected by salmonella bacteria and the Jack in the Box contaminated beef incident in the USA". So why is food safety still a problem, and a growing problem at that, in the world today?

Apart from the food safety systems still being a "work in progress " irrespective of which country one wishes to consider [5], Riethmuller and Morison [6] have identified at least three reasons for the growing importance

of food safety issues. The first concerns changes in food consumption patterns. People are eating out more, resulting in greater consumer awareness of hygiene. Second, manufactured food products and prepared meals are available through supermarkets and other food outlets, so that the onus is on these food retailers to ensure hygiene standards are adhered to. Third, food safety has become a notable non-tariff barrier in international trade. There are numerous examples of developing nations accusing developed ones of using food safety as a protectionist measure for domestic industry rather than for genuine safety concerns (see, for example, [7,8]).

In view of these global developments the question of who is responsible for food safety arises. As consumers lack the scientific and infrastructural capacity to evaluate food risk, it is incumbent on the food industry to act with both integrity and within the legal guidelines, and for governments to provide those guidelines and enforce them for the consumer's protection [9].

This article will survey how governments in the United States, the European Union and Australasia regulate the food industry and influence the corporate social responsibility (CSR) behavior of companies—and even the official world of governmental authorities at home and abroad-by examining the relevant characteristics of

their food safety systems. We argue that government influence on CSR in the food industry is determined by the intensity and salience of its own behavior and actions including regulations. The case studies have been selected for the purposes of: a) providing exemplars of good and innovative international practice in prevention of food scandals and promoting good practices; and b) showing how national (USA), supranational-plus-international (EU), and bi-national (Australia and New Zealand) authorities handle food safety issues. This is of relevance to China in that it is a unitary state like the USA, but with a policy of strengthening regional cooperation in East and Central Asia (multilateral regionalism) as well as an internal system of provinces and autonomous regions whose collective population size more than doubles that of the EU. The PRC also functions as a "one country, two systems" entity with regard to the Special Administrative Regions of Hong Kong, Macao and potentially Taiwan that are different to the provinces and autonomous regions within China. The regulatory authority that covers Australia and New Zealand, Food Standards Australia New Zealand (FSANZ), provides a successfully functioning model for a regulatory function across two polities. While this "one system, two countries" is the reverse of China's "one country, two systems" formula, it does show that as the PRC and its Special Administrative Regions continue to converge in terms of a capitalistic system but emphasize politico-social differences within the One China concept, a common regulatory mechanism across various sectors could be an acceptable evolutionary move.

The final section of the article profiles China's "tainted milk" scandal and draws lessons from the theoretical discussion and case studies for the Chinese government and its food industry. The use of corporate social responsibility as a term applies in China to both private and public sectors as these are often combined, either from the transitional nature of China's economy (from command to market) or from an emerging trend demonstrated by the EU—the Public Private Partnerships (PPPs). As Howcroft [10] notes:

"Given the scale of the infrastructure and investment gap that the governments of Europe are facing and the constraints that they face in developing and financing their needs, an increased use of PPP approaches is inevitable… [Governments] need to invest in the public sector's understanding and capability to develop and procure such projects in ways which maximize the overall benefits to the public sector and the public at large."

In China the boundaries between the state and business are neither clear nor necessarily inevitable. CSR must therefore take a broader view in its purview of application when addressing recommendations for China. Section 2 (What is CSR?), however, will take a theoretical perspective and focus on business enterprises in order to set the stage for conceptual applications for diverse settings and actors (Section 3–Case Studies), from which lessons for China (Section 4) may be drawn.

2. Defining Corporate Social Responsibility

Corporate social responsibility varies in meaning and definitions depending on stakeholder perspectives, be they employees, consumers, unions, governments, local communities, shareholders, and executives. Unlike Friedman [11] who poses the conventional argument that an organization's only responsibility was generating profit, and that any activity that detracted from the goal of profit did not serve the shareholders' best interests, Carroll [12] goes beyond the economic limits of an organization's responsibility to add legal, ethical and philanthropic dimensions. Societal rules in the form of laws and regulations had to be followed, but not simply at the minimum required level. Organizations should seek to realize higher standards, thereby fulfilling an ethical responsibility. Moreover, this ethical responsebility feeds into an organization's philanthropic role of giving back to the community through donating part of its profits to the satisfaction of societal needs generally. Hence, CSR is an encompassing concept covering at least economic, legal, ethical and philanthropic considerations. We argue that a firm's disposition towards CSR is best understood in terms of whether it fulfils these four types of CSR obligations. For this reason, we will use this encompassing definition of CSR throughout the article.

Governmental influence on the CSR orientation of business firms is well recognized, and there is broad agreement that governments shape the attitude and behavior of company CSR through legislative measures [13–15]. Recently, scholars have noted broader roles that governments have played in promoting CSR [15–19]. Crane and Matten [20] argue that the role of government has changed from traditional regulator of dependent firms to that of multi-faceted player in the face of increased corporate power. We have adopted Fox *et al.*'s [15] identification of four key roles for governments in promoting CSR: mandating, facilitating, partnering and endorsing. Each of these roles may be expected to vary in intensity and salience in relation to company CSR depending on the types of CSR under study. We believe this definition captures the comprehensive nature of CSR in the food industry. The foremost responsibility of companies engaged in the food industry is not economic profit in preference to all else—for if only profits were at stake then the conesquences could be devastating, as China's "tainted milk" scandal revealed—but the need to be legally responsible and obey laws.
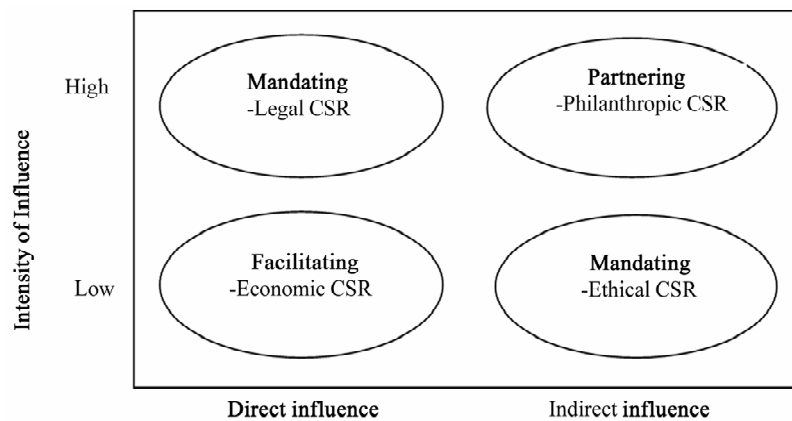
Source: Adopted from Yang and Casali, 2009
Figure 1. Government Influence on Four Types of CSR.

We agree with Yang and Casali [21] that government influence on CSR is determined by the intensity and salience of its own behavior and actions and have adopted their two-by-two matrix as a framework to illustrate that the interaction between the government's role and CSR is a function of (direct and indirect) government intervention (see Figure 1).

2.1 Government Influence and CSR

Institutional theories suggest that states develop formal institutions in the form of laws and regulations to effect order, reduce uncertainty, and influence social actor behavior in coordinating and promoting economic exchange [22]. Specifically, rational choice institutionalists argue the behavior and actions of the government are important to the extent that formal and informal rules, with their associated monitoring and sanctioning mechanisms, result in either enabling or constraining social actors [23]. On the other hand, firms may take corrective actions in response or in anticipation of government intervention by choosing to: a) legitimize them, b) avoid legislation, and/or c) avoid negative publicity. Governments, however, can deploy a preemptive strategy to create an institutional environment capable of fostering a CSR outlook in business.

2.2 Governmental Role with Four Types of CSR

2.2.1 Government as Mandator (Legal CSR)
Institutional theory holds that firms tend to comply with government legislation and regulation to legitimize their behavior in the marketplace [22]. Government can wield the power of formal institutions-such as legislation, the judicial system and regulatory agencies-to instill the attitude and shape behavior of company CSR [24]. Government as mandator influences company attitudes to CSR primarily through a carrot-and-stick strategy of: 1) providing tangible inducements for company resource allocation toward stakeholders and behavior that is socially

responsible; and 2) inflicting punishment through penalties if actions are not taken, or standards are contravened [21].

Government action in the form of legislation has been argued to be clearly influential in shaping company behavior because it is mandatory [14]. This form of government intervention is direct and its degree can be decisive from the company perspective. Coercion of this nature tends to result in CSR policies being internalized to reduce risks and search costs [25], as evidenced in a high profile US legislative framework-the Sarbanes-Oxley Act of 2002-which legislates financial reporting for publicly traded companies and their auditing firms [26]. A fine of US$5 million and a jail term of 20 years can be the penalty for CEOs and CFOs failing to certify statements or signing false statements [24].

Support is also available for the efficacy of regulation on the emergence of socially responsible behavior with regard to the environment. Stone, Joseph, and Blodgett [27] find that the higher the degree of regulation the greater the likelihood of businesses adopting socially responsible behaviors. The efficacy of government's role as Mandator is also seen in better compliance with the industry code of conduct [28]. It is argued that one of the motivating reasons for companies to follow codes of conduct appears to be a desire to avoid interference or legislation by government [29–31].

2.2.2 Government as Facilitator (Economic CSR)
Governments can play the role of facilitators to encourage and influence corporations toward being economically responsible. This can be done through government initiatives such as providing guidelines on content, fiscal and financial mechanisms, and creating framework conditions [13]. Notable actions toward facilitation include developing public policies for the training of skilled workers or establishing specialized government agencies to oversee these programs. The UK, for exam-

ple, implemented Industrial Training Boards and the Manpower Services Commission to encourage CSR in the areas of training and work experience opportunities. Moon [32] found that the Manpower Services Commission led to the Confederation of British Industry forming the Special Programs Unit to ensure large-scale training programs for businesses occurred.

Another mechanism for government as facilitator of economic CSR is provision of subsidies. Government subsidies allow firms to defray costs associated with employment and training schemes, thereby providing greater incentives to participate in new government employment programs [13]. Another example is the UK Department of Trade and Industry's subsidy of research, publication and a website for BITC reports. In this way, government as facilitator produces the kind of influence that is salient and of direct relevance to the resource capacity of companies to implement economic CSR.

It is common for governments worldwide to sponsor excellence in business awards as a mechanism to encourage economic CSR [32]. Another device is the tax credit designed to bring private enterprise into poorer areas, as shown by the case of the UK 2002 community invest tax credit scheme (ibid.). Businesses are demonstrably more likely to participate in economic CSR programs when the government is seen as Facilitator. For instance, one CSR business association, Business in the Community with its 700 members, accounts for 20% of private sector employment (ibid.).

Also of interest is the finding that governments can exert influence on economic CSR by owning financial institutions, as in the case of the UK [33]. In most cases, government influence on economic CSR is direct, but with low intensity compared to the higher intensity strategy of government as mandator.

2.2.3 Government as Partner (Philanthropic CSR)

Most governments pursue a less regulatory approach in relation to philanthropic responsibility. Rather, they seek to reward good behavior, as shown in the taxation laws of many countries that allow taxpayers to entirely or partially deduct philanthropic donations from their taxable income. For example, the Australian Income Tax Assessment Act of 1997 allows deductions for donations to recognized charities in order to encourage charitable behavior by enterprises. The Australian government benefits through its willingness to partner with business through philanthropy as this strategy spreads the economic burden of social responsibility across both the public and private sectors. Indeed, the popularity of this strategy can be seen through similar approaches taken by some EU countries [13]. According to a 2005 study conducted by the Australian government in collaboration with other organizations, $3.3 billion were given by

businesses in Australia between the 2003 and 2004 [34]. These businesses represented 67% of the total number of businesses in Australia (525,900) [34].

Tax deductibility, significant as it is, is not the most important reason for philanthropic behavior in firms. Other positive influences encouraging philanthropic responsibility have been found to embrace the following: a sense of reciprocation, respect for nonprofit organizations, the desire to strengthen the community, and improving the world [34]. In recent years, a phenomenon called "strategic philanthropy" has emerged. Its exponential growth is indicative of its economic value: strategic philanthropy is viewed as a new and innovative way to achieve a competitive advantage. It involves a company directly linking its core business-be it product or service-with charitable activity, for example, by donating one dollar for each purchase of the company's product, or a percentage of the sales profit for a particular day [35]. In this way, a business can simultaneously fulfill its philanthropic responsibilities, promote its own product or service, and obtain a tax deduction. This provides a strong argument for the efficacy of government as partner when it comes to influencing philanthropic CSR indirectly, albeit very strongly [21].

2.2.4 Government as Endorser (Ethical CSR)

As shown in the Figure 1, the government plays an Endorser role when it exercises its influence indirectly and at low intensity: in other words, when neither legal nor fiscal strategies can be used as influential means in ethical CSR.

Yang and Casali [21]demonstrate that there is a cross linkage between ethics and law, and this nexus reflects the reality of laws issuing from a societal process that identifies and validates collectively the perceived minimum standards in a society, that then become the formal responsibility of government to protect. Arguably, behaviors that have been converted into legislation and then executively reinforced contribute to the pool of societally agreed acceptable standards. These represent in a given era the minimum standards that are not negotiable and must be viewed as core principles [35,36]. The main purpose of the legislative process is to shift those principles from the very least influential government role (Endorser) to the areas where the government can have intensive and direct influence as Mandator, in order to impose those principles on firms in a powerful way. The law-ethics nexus thus represents a crucial space within which governments may manoeuvre to enhance the status and nature of CSR within society's ontological base.

Yet, government influence may lack power over all those actions that have not yet reached the grey area and that remain more a potentiality or "wish list" than an imperative or "must have list". Examples of principles in the wish list are: proactive action in environment pro-

tection, workplace safety, customer interest, and responsiveness to stakeholder concerns [21].

An example in 2008 of government as endorser was provided by the Prime Minister of Australia, Kevin Rudd, who asked the major Australian banks to pass on interest rate cuts to the consumer in full, emphasizing the fact that this would constitute a more ethical way to do business [37]. Despite political exhortation, there was no supporting legislation, or any tax benefits that would encourage the banks to behave in this way. The govern- ment was only endorsing a code of conduct that rested on a key principle: that the main purpose in cutting interest rates should be to reduce the burden of financial pressure on the consumer. It is not intended to increase bank pro- fits by reducing the price of resources (money)—that is, "profiteer at the expense of customers" (ibid.).

The above suggests that government as endorser is in a weak position to influence CSR in terms of intensity and salience as compared to the other three types of CSR. It may be hypothesized that government as mandator is the strongest, with the other two-government as facilitator and government as partner occupying a second tier of intensity and salience. Government as endorser occupies the bottom tier. All, however, are valuable when deployed in concert so as to produce a balanced outcome: too much of one, such as government as mandator, might lead to the "nanny state" syndrome for instance; or an over-emphasis on low-intensity indirect influence could result in an ineffectual CSR effort.

3. Case Studies: The US, the EU, Australia and New Zealand

To understand how China's food safety system may benefit from the theoretical-analytical discussion above, it is important to now turn to a number of empirical case studies in the developed world where international best practice may be expected to be found. The Chinese themselves have recognized that they lag behind in international norms and practices. At the Fifth China Food Safety Annual Meeting in 2007, China's Vice Minister of Health Chen Xiaohong admitted that food safety in China did not match that of the developed world. Among the problems he identified were: pollution; low quality of some food products; inadequate technology, equipment, and quality testing systems; as well as weaknesses in food safety management [38].

So what do the experiences of governments in industrialized nations reveal in relation to the hypothesized governmental role with the four types of CSR? What lessons do these findings have for China? The first is a brief case study from the United States where new comprehensive methods are used. The second derives from the European Union whose legislative strengths are especially pertinent to China's own regulatory instincts, and

the final investigation turns to Australia and New Zealand.

3.1 The United States: Comprehensive Strategies in a Unitary State

The Food Protection Plan (FPP), released by the US Food and Drug Administration in November 2007, represents an especially exemplary and up-to-date model in its dual features. These are its provision of: a) an inte- grated strategy that incorporates "both food safety and food defense for domestic and imported products"; and b) collaborative engagement "across the agency to address the three core elements of protection: prevention, inter- vention and response" [39]. It is in these collaborative engagements that the government's role as Mandator, Facilitator, Partner and Endorser in the specific area of food safety becomes evident—not only with the private sector but with all stakeholders.

Admittedly, it can only be judged by the short timeframe of its existence. Still, an overview of the first six months of its activities is available. In terms of prevention, outreach activities are prominent and these approximate the government as facilitator and partner models:

"This outreach has involved multiple meetings with various foreign countries, state and local organizations, and industry and consumer groups… Specific risk-based prevention activities include FDA working in collaboration with states, universities and industry on a Tomato Safety Initiative. In an effort to increase foreign capacity and FDA's presence beyond our borders, FDA has engaged with India and begun implementation of the China Memorandum of Agreement. The first bilateral meeting with China was held in Beijing in March 2008" [39].

The second core element of protection, intervention, has seen an increase in the number of state inspections and employees to conduct them. Here is a case of government as mandator. The legal regulatory element is evident but it is balanced by qualitative improvements in identifying "food safety threats at the border"; for example, the piloting of a new system called PREDICT. To coordinate developments such as these a research committee has been tasked with maintaining a "collaborative research agenda that supports activities under prevention, intervention and response, such as mitigation strategies and rapid detection systems" [39].

The third pillar of protection is response. Herein lies the government as endorser role, for the key group identified for improved response is that of stakeholders. It is they who are deemed to "be able to quickly identify where a contaminated product came from and where it has been distributed". Under development is so-called Incident Command System training and Rapid Response Teams "to enable rapid, localized response to incidents" [39].

3.2 The European Union: Codifier and Governance Coordinator

Europe provides a ready laboratory for recent food scandals and governmental responses. In June 1999, it was found that egg, pork, veal, beef, milk, cheese and butter products in Belgium were contaminated with dioxin. The owners of the Belgian company, where the problem was first traced, were suspected of knowingly fabricating or buying from Dutch suppliers feed grain mixed with cheap, second-hand oil or fat that turned out to be con- taminated with dioxin. The tainted feed was sold to 1,400 producers in Belgium, France and the Netherlands. Authorities of the European Union, based in Brussels, criticized the Belgian government for taking months to inform the EU about the problem once discovered [40]. This case not only highlights the public-private sector relationship in CSR but levels of government-to-government communication and influence. Codex Alimentarius Commission as the highest international body on global food standards represents a higher governance level than government as mandator within the unitary state. The Food and Agriculture Organization (FAO) and World Health Organization (WHO) Codex Alimentarius Commission met in Rome in mid-1999 to respond to the European crisis over dioxin-contaminated animal products. The Commission set up an intergovernmental task force to accelerate the adoption of a Draft Code of Practice on Good Animal Feeding. It also approved the establishment of an intergovernmental task force to speed up the elaboration of guidelines and standards for foods derived from biotechnology; and passed new international guidelines that clearly defined the nature of organic food production to prevent misleading claims. The new guidelines covered the production, processing, labeling, and marketing of organic food [41].

"Mad cow disease" was perhaps the most publicized of the European food scandals in recent time. It was related to the Bovine Spongiform Encephalopathy (BSE) and its human form, Creutzfeldt-Jakob (nvCJD) disease. The press release from the FAO on 26 January 2001 warned that the risk of BSE and its human form posed a risk worldwide and not only in Europe. The FAO also noted that all countries which imported cattle or meat and bone meal (MBM) from Western Europe, especially the UK, during and since 1980s, could be considered at risk from the disease [42].

The BSE and nvCJD issue once again showed that food scandals are closely associated with international trade and this is an area in which government as mandator has been less effective than the national level. In the aftermath of the issue, international actors such as FAO and WHO Codex Alimentarius engaged in a study on a "Code of Practice for Good Animal Feeding" to ensure that animal products do not pose health risks to consumers. Consultations were wide-ranging and included the European Union, Australia, Canada and the United States. This is a case suggestive of global governance as endorser to influence national governments to become Mandators of CSR to the food industry, to assist in its task; the FAO introduced an internet based information service that included a rapid alert system on food safety issues [43].

3.2.1 White Paper on Food Safety

In the aftermath of the BSE and dioxin food scandals, the EU published its "White Paper on Food Safety" (12 January 2000). It should be noted that at the time when the EU faced alarming food scandals, the EU was the world's largest producer of food and beverage products and this industry was the third largest industrial employer of the EU with over 2.6 million employees, of which 30% were in small and medium enterprises [44].

The white paper proposed a "radical new approach" for food safety in Europe. Like the recent US Food Protection Plan, food safety policy would be comprehensive and integrated in its conception. To this end, an independent European food authority was proposed.

3.2.2 Formation of a Food Safety Authority

The Commission stated that an independent European food authority would be entrusted with "scientific advice on all aspects relating to food safety, operation of rapid alert systems, communication and dialogue with consumers on food safety and health issues as well as networking with national agencies and scientific bodies" and it would serve an analytical function but only the European Commission would decide on what action to take. The food authority's fundamental principles would be independence, excellence and transparency [44]. A wide range of other legislative measures were proposed covering all aspects of food products from "farm to ta- ble". The legislation was aimed to be easily under- standable for all operators to put into effect. It gave "teeth" to an otherwise weak government as endorser function. The EU as a supranational government is showing the way forward in terms of governmental in- fluence combining Mandator with Endorser.

Like the American example above, stakeholder values were upheld by the white paper's proposed actions to keep consumers well informed about newly emerging food safety concerns and to involve them in food safety policy. The white paper also had implications for trade partners of the EU which, in its position as a massive importer and exporter of food products, must play an "active role" in international bodies and be effective in explaining the European position on food safety [44].

The outcomes of the action plan were the integration of food safety policies within the EU countries and—to an extent—the EU's trade partners, as well as a more coordinated system. Transparency at all levels of food safety

policy stands out as a key principle. Relating to BSE legislation, the white paper identified the problem of inconsistency in approach. In addition, the adoption of measures did not involve all EU institutions. In order to address the integration needs within the single market of the EU, a new approach was proposed for farming, food processing, handling and distribution.

As a result of the proposals, the European Food Safety Authority (see website EFSA, 2008a [45]) was set up based on Regulation (EC) No 178/2002 of the European Parliament and of the Council, on 28 January 2002, as an independent source of scientific advice and communication on risks associated with the food chain [46].

The lessons from the EU on government strategies for CSR in the food safety arena are that a combination of the top and bottom tier (legal Mandator and ethical Endorser) works best. The ethical (independence from government, transparency, and stakeholder consultation which are norms that are being entrenched) is in fact subsumed within the legal legislative framework through the EU's unique governance structure. This is a more codified, yet governance (not government)-based system which sets it apart from nation-states like the US. The EU, however, shares with the US and the bi-national case study below (Australia and New Zealand) the philosophy of a comprehensive and coordinated approach. This pertains to a systems approach where the whole system is examined and activated, rather than selective problem-solving.

3.3 Australia and New Zealand: One System, Two Countries

Integration and collaboration as twin themes of international best practice in food safety are also evident in the antipodes. Food Standards Australia New Zealand (FSANZ) is a bi-national government regulatory agency whose mission, according to its website is "to provide a safe food supply and have well-informed consumers". Its main responsibility is to develop and administer the Australia New Zealand Food Standards Code which is given legal force through these two countries' food legislation. FSANZ as well as other government agencies "monitor the food supply to ensure that it is safe, and that foods comply with standards for microbiological contaminants, pesticide residue limits and chemical contamination" [47]. At the time of the Chinese "tainted milk" scandal, FSANZ's website provided updates on Chinese imported food, including products withdrawn, product testing, consumer advice and maximum melamine levels in food. In a coordinated effort with other national and state food safety agencies, FSANZ engaged in the following actions [47]: working with importers and local food manufacturers to ascertain if products with Chinese dairy ingredients

are possibly contaminated with melamine; conducting precautionary testing of products on Australian shelves; monitoring of imports by the Australian Quarantine and Inspection Service; and working closely with food regulators around the world including the WHO.

The above is indicative of how the Australasian system responds to a food contamination problem when it involves imports. China turned to the WHO for help in its "tainted milk" crisis and more rigorous regulations on food safety were being introduced [48].

The final section of this article draws lessons for China from the foregoing two sections on CSR theory and empirical case studies.

4. Lessons for China

4.1 Background to China's "Tainted Milk" Scandal

The "tainted milk" scandal broke out in China in September 2008. The Sanlu brand of powdered milk formula was found to be tainted by the industrial chemical melamine, a binding agent used for plastics and glue but added to watered-down milk as it mimics protein. The contamination resulted in the deaths of at least four babies and some 54,000 infants needing medical treatment that month. The main symptom was kidney stones, for which 3,458 infants were hospitalized in Beijing alone; indeed, a survey of 308,000 households in Beijing indicated that a quarter had fed their children the contaminated milk prior to it being removed from the shelves [49]. Melamine was found not only in Sanlu baby for- mula but a total of 53 dairy brands in China, as well as foreign brands using Chinese dairy ingredients [50]. This was not the first food safety incident emanating from China. A range of goods exported from China, including toothpaste and pet food, have been found to contain melamine and other industrial chemicals.

Premier Wen Jiabao responded to the scandal by saying China had to strengthen monitoring at the production level, as well as instilling a stronger sense of social conscience and business ethics at the management level [51]. So, too, Chinese President Hu Jintao said lessons must be learned from the milk scandal to "ensure all dairy products sold to the market are qualified products" [52]. By late October 2008, a draft food safety law was being considered by the National People's Congress. The law would seek to a) "prevent any cover-ups by health authorities"—which was said to have occurred in the Sanlu case in order to avoid a scandal during the Beijing Olympics—and b) would confer on these same govern- ment health officials direct responsibility for approval of any additives in processed food [49]. Thus the authorities would be held responsible for what goes into processed food as well as for attempts to disguise the outcome.

4.2 Lessons

Besides a lack of proper governmental oversight and inadequate procedural mechanisms in quality testing, the "tainted milk" scandal (like other unsafe products in the past) showed the presence of corrupt business practices that bypassed China's quality controls. As noted by deGategno [53]: "China has not succeeded in building effective systems for monitoring and enforcing ethical standards among its officials. The central government has and is continuing to implement reforms that make officials increasingly accountable, but they have little control if no one reports corrupt acts." It is local officials, according to deGategno, who are key players in food safety and who need to abide by the rules. Here is a case of government needing to instill ethical CSR at the level of businesses and local officials. The strategy to do so requires not only an EU-style government as both Mandator and Endorser, but also more work on the government as Facilitator and Partner.

The corruption factor has an international dimension. Ironically, a company from New Zealand—an exemplary country in terms of food standards regulation and business ethics-was involved in the scandal. Owner of 43% of the Chinese company at the centre of the scandal, Sanlu, was New Zealand dairy co-operative Fonterra. It transpired that Fonterra had known of the melamine contamination six weeks before it "raised the alarm" (Sanlu allegedly had known for eight months) [54,55].

The involvement of Fonterra illustrates the global nature of food manufacturing and the wider governance responsibility this entails. The EU provides a quality model for the international dimension of how to codify, facilitate, communicate and develop a normative environment for food safety in cooperation with other governments and stakeholders. China's own white paper on food safety, published in August 2007, reflects a number of these lessons [56], even if they were to no avail for the victims of the "tainted milk" scandal within a year of its publication. Such was the impact of this scandal that the UN published its own report on food safety in China in October 2008. China needed to modernize its food safety legislation, overcome ambiguities in supervisory responsibilities; improve oversight and enforcement; better educate stakeholders—consumers, the food Industry and health authorities; and continue to pursue international standards of best practice. One of the problems in China was that there were too many small enterprises, many illegal, to monitor. It is these that are thought responsible for introducing illegal chemicals, with melamine having "apparently ended up in dairy products after middle men who collected milk from farmers and sold it to large dairy companies added the chemical" [57]. Approximately 350,000 of China's 450,000 registered businesses in food production and processing employ as few as 10 people or less. The UN report blamed these small enterprises for presenting "many of the greatest food safety challenges" (ibid.).

Despite the importance of government regulation in the government as mandator strategy, as discussed in this article, on its own it is inadequate and requires the other three types of CSR to combine for greater effectiveness. The Chinese system was found to be antiquated in that it was managed by different regulations and an ethos of government being expected to be responsible for the entire food system, whereas producers also needed to be responsible for food safety [58]. To induce greater responsibility on the part of food producers, the activation of government as facilitator, partner and endorser represents a more comprehensive strategy.

For China, the experiences of others allow it the advantage of being able to leapfrog in the construction of its own food safety system, so that the "workshop of the world" can simultaneously lift standards and glean the best practices the world has to offer on a comparative basis. That China is traveling this path is evident from the decision to publish its own white paper on food safety in 2007 to allay fears about the safety of China's exports, and other reforms that were underway after the "tainted milk" scandal in 2008. Already China's Food and Drug Administration has been placed under the Ministry of Health rather than having the responsibility divided among 16 organizations. Moreover, some companies—Sanlu included—which were previously allowed to conduct their own quality inspections are no longer permitted to do so [50]. An attempt to change the government culture of hiding problems to one of reporting them promptly is underway through legal measures but also needs to be strengthened through consumer protection mechanisms and an enhanced corporate social responsibility.

5. Conclusions

Our study demonstrates that food safety has to be bound with CSR and the government has a critical role to play by developing comprehensive strategies to make corporations in food industry behave in a socially responsible way. A number of contributions emerged from our exploratory study. First, we have extended the CSR literature to the food industry and emerging economies like China. Second, we have identified the distinctive role the government plays in the food industry and that government influence on CSR in the food industry is determined by the intensity and salience of its own behavior and actions. Third, we have provided a timely guide to establishing a food safety system in China based on empirical evidence that a balanced CSR strategy covering at least economic, legal, ethical and philanthropic considerations would work best for China.

REFERENCES

[1] WHO, WHO Initiative To Estimate The Global Burden Of Foodborne Diseases, 2007, http://www.who.int/foodsafety/publications/foodborne_disease/FERG_Nov07.pdf.

[2] Smith, D. and Riethmuller, P., "Consumer concerns about food safety in Australia and Japan," British Food Journal, Vol. 102, No. 11, pp. 838–855, 2000. (Previously published in the International Journal of Social Economics, Vol. 26, No. 6, 1999).

[3] New York Times, "Japan says radishes caused food," Poisoning, 1996, http://query.nytimes.com/gst/fullpage.html?sec=health&res=9C02E3D8153EF93BA3575BC0A960958260.

[4] English, Holly, "Arnott's doesn't crumble under pressure," St. James Ethics Centre, City Ethics, No. 28, 1997, http://www.ethics.org.au/about-ethics/ethics-centre-articles/ethics-subjects/business-ethics/article-0296.html/.

[5] ChinaCSR.com, UN Issues Paper on Food Safety in China, 2008, http://www.chinacsr.com/en/2008/10/24/3438-un-issues-paper-on-food-safety-in-china/.

[6] Riethmuller, P. and Morison, J., "A comparative survey of beef consumption behaviour in Australia, Japan and the United States," Agricultural and Resource Manage- ment Services Pty Ltd, 1995.

[7] Avila, John Lawrence, Non-Tarrif Barriers Facing Philippine Exporters, World Bank, 2005, http://siteresources.worldbank.org/INTEAPREGTOPINTECOTRA/Resources/579386-1145902243289/Philippines-John-6-27.pdf.

[8] Arun, S., Government Asks EU to Lift 'Paranoid' Health-Related Trade Barriers, The Financial Express India, 2008, http://www.financialexpress.com/news/Govt-asks-EU-to-lift-paranoid-healthrelated-trade-barriers/307540/0.

[9] Kennedy, D., "Humans in the chemical decision chain," In H. O. Cater and C. F. Nuckton, Ed., Chemicals in the human food chain: Sources, options and public policy. University of California, Davis, Agricultural Issues Center, 1988.

[10] Howcroft and Adrian, Public and Private Sectors: Who's Learning from Whom? European Business Forum, 2008, http://www.ebfonline.com/Article.aspx?ArticleID=246.

[11] Friedman, M., "The social responsibility of business is to increase its profits," New York Times Magazine, 1970.

[12] Carroll, A. B., "A three-dimensional conceptual model of corporate social performance," Academy of Management Review, Vol. 4, No. 4, pp. 497–505, 1979.

[13] Albareda, A., Lozano, J. M., Tencati, A., Midttun, A., and Perrini, F., "The changing role of governments in corporate social responsibility: drivers and responses," Business Ethics, Vol. 17, No. 4, pp. 347–363, 2008.

[14] Aguilera, R. V., Rupp, D. E., Williams, C. A., and Ganapathi, J., "Putting the S back in corporate social responsibility: A multi-level theory of social change in organiza-tions," Academy of Management Review, Vol. 32, No. 3, pp. 836–863, 2007.

[15] Fox, T., Ward, H., and Howard, B., "Public sector roles in strengthening corporate social responsibility: A baseline study," World Bank, 2002.

[16] Aaronson, S. A., "Corporate social responsibility in the global village: The british role model and the American laggard," Business and Society Review, Vol. 108, No. 3, pp. 309–38, 2003.

[17] Zappal, G., "Corporate citizenship and the role of government: The public policy case," Australian Research Paper 4, 2003–2004.

[18] Lepoutre, J., Dentchev, N., and Hene, A., "On the role of the government in the corporate social responsibility debate," paper presented at the 3rd Annual Colloquium of the European Academy of Business in Society, Ghent, 2004.

[19] Nidasio, C., "Implementing CSR on a large scale: The role of government," paper presented at the 3rd Annual Colloquium of the European Academy of Business in Society, Ghent, 2004.

[20] Crane, A. and Matten, D., "Business Ethics: A European Perspective—Managing Corporate Citizenship and Sustainability in the Age of Globalization," Oxford University Press, 2004.

[21] Yang, X. and Casali, L., "In search of the missing link: An analysis of government influence in CSR," Sinergie, 2009.

[22] North, D. C., "Institutions, institutional change, and economic preference," Norton, 1990.

[23] Campbell, L. J., "Why should corporations behave in socially responsible ways? An institutional theory of corporate social responsibility," Academy of Management Review, Vol. 32, No. 3, pp. 946–967, 2007.

[24] Yang, X. and Rivers, C., "Cross-border differences in company attitude to corporate social responsibility," Journal of Business Ethics, DOI 10.1007/s10551-009-0191-0, September 2009.

[25] Husted, B. W. and Allen, D. B., "Corporate social responsibility in the multinational enterprise: strategic and institutional approaches," Journal of International Business Studies, Vol. 37, No. 6, pp. 838–849, 2006.

[26] Rockness, H. and Rockness, J., "Legislated ethics: From enron to sarbanes-oxley, the impact on corporate America," Journal of Business Ethics, Vol. 57, pp. 31–54, 2005.

[27] Stone, G., Joseph, M., and Blodgett, J., "Toward the creation of an eco-oriented corporate culture: A proposed model of internal and external antecedents leading to industrial firm eco-orientation," Journal of Business and Industrial Marketing, Vol. 19, No. 1, pp. 68–84, 2004.

[28] Aguilera, R. V. and Cuervo-Cazurra, A., "Codes of good governance worldwide: What is the trigger?" Organization Studies, Vol. 25, No. 3, pp. 415–43, 2004.

[29] Brereton, D., "The role of self-regulation in improving corporate social performance: The case of the mining industry," in Proceedings the Current issues in regulation: Enforcement and compliance conference, Australian In-

stitute of Criminology, Melbourne, 2002.

[30] Diller, J., "A social conscience in the global marketplace? Labour dimensions of codes of conduct, social labelling and investor initiatives," International Labour Review, Vol. 138, No. 2, pp. 99–129, 1999.

[31] McInnes, D., "Can self-regulation succeed?" Canadian Banker, Vol. 103, No. 2, pp. 30–36, 1996.

[32] Moon, J., "Government as a driver of corporate social responsibility: The UK in comparative perspective," ICCSR Research Paper Series 20-2004, ICCSR, University of Nottingham, pp. 11–27, 2004.

[33] Yamak, S. and Süer, Ö., "State as a stakeholder," Corporate Governance, Vol. 5, No. 2, pp. 111-120, 2005.

[34] Australian Government, Australian Giving: Research on Philanthropy in Australia, 2005, http://www.ag.au/cca.

[35] Ferrell, O. C, Fraedrick, J., and Ferrell, L., "Business ethics: Ethical decision making and cases, 6th," Houghton Mifflin Company, 2005.

[36] Trevino, L. K. and Nelson, K. A., "Managing business ethics: Straight talk about how to do it right, 4th," John Wiley and Sons, 2007.

[37] Uren, D. and Richard, G., RBA Tells Banks: Pass on Rate Cut, The Australian, 2008, http://www.theaustralian.news.com.au/story/0,25197,241 77884-601,00.html.

[38] ChinaCSR.com, 300 Chinese Food Companies Pledge To Ensure Food Safety, 2007, http://www.chinacsr.com/en/2007/09/25/1715-300-chines e-food-companies-pledge-to-ensure-food-safety/.

[39] FDA—US Food and Drug Administration, FDA Food Protection Plan Six-Month Progress Summary, 2008, http://www.fda.gov/oc/initiatives/advance/food/progressre port.html.

[40] New York Times, Brussels Journal; Food Scandal Adds to Belgium's Image of Disarray, 1999, http://query.nytimes.com/gst/fullpage.html?res=9F06E0D C1139F93AA35755C0A96F958260&sec=health&spon= &pagewanted=all.

[41] FAO, Codex meeting passes new international standards for food safety, 1999, http://www.fao.org/News/990701-e.htm.

[42] FAO, FAO: Countries Around the World Should be Concerned About 'Mad Cow Disease' and should Take Action to Reduce and Prevent Risks, 2001a, http://www.fao.org/waicent/ois/press_ne/PRESSENG/200 1/pren0103.htm.

[43] FAO, FAO/WHO Call for More International Collaboration to Solve Food Safety and Quality Problems, 2001b, http://www.fao.org/WAICENT/OIS/PRESS_NE/PRESSE NG/2001/pren0143.htm.

[44] European Commission (Commission of the European Communities), White Paper on Food Safety, 2000, http://ec.europa.eu/dgs/health_consumer/library/pub/pub 06_en.pdf.

[45] EFSA—European Food Safety Authority, 2008a,

http://www.efsa.europa.eu.

[46] European Commission (Commission of the European Communities), Regulation (EC) No 178/2002 of the European Parliament and of the Council of 28 January 2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety, 2002, http://eur-lex.europa.eu/smartapi/cgi/sga_doc?smartapi!ce lexapi!prod!CELEXnumdoc&lg=EN&numdoc=32002R0 178&model=guichett.

[47] FSANZ—Food Standards Australia New Zealand, 2008, http://www.foodstandards.gov.au.

[48] ABC Radio Australia, Head of China's product quality agency resigns, 2008a, http://www.radioaustralia.net.au/news/stories/200809/s23 71425.htm.

[49] AFP, Quarter of Beijing Kids Fed Bad Milk, The Australian, pp. 10, 2008a.

[50] Magistad and Kay, M., China's Image Sullied by Tainted Milk, Yale Global Online, Yale Center for the Study of Globalization, 2008, http://yaleglobal.yale.edu/display.article?id=11403.

[51] China Economic Net, "Premier vows to provide safe products for people and world," 2008, http://en.ce.cn/National/Politics/200809/28/t20080928_16 950835.shtml.

[52] Xinhua News Agency, Domestic Service, Beijing, in Chinese, translated by BBC Monitoring Asia Pacific, China: Hu Jintao Inspects Anhui Farms; CCP Third Plenum to Discuss Rural Reform, 2008, http://www.redorbit.com/news/science/1575336/china_hu _jintao_inspects_anhui_farms_ccp_third_plenum_to/inde x.html?source=r_science.

[53] deGategno and Patrick, China, Land of Tainted Milk and Honey, New Atlanticist, 2008, http://acus.org/new_atlanticist/china-land-tainted-milk-an d-honey.

[54] ABC Radio Australia, Tainted milk scandal costs NZ giant $95m, 2008b, http://www.radioaustralia.net.au/news/stories/200809/s23 73173.htm.

[55] ABC Radio Australia, "NZ milk firm 'sorry' for contamination crisis," 2008c, http://www.radioaustralia.net.au/news/stories/200809/s23 73569.htm.

[56] State Council Information Office, The Quality and Safety of Food in China (White Paper), 2007, http://www.chinadaily.com.cn/bizchina/2007-08/18/conte nt_6032837.htm.

[57] Ang, Audra, UN Urges Improved Food Safety for China, Yahoo News, 2008, http://news.yahoo.com/s/ap/20081022/ap_on_re_as/as_ch ina_tainted_products.

[58] AFP, UN Urges China to Revamp Food Safety after Milk Crisis, 2008b, http://findarticles.com/p/articles/mi_kmafp/is_200810/ai_ n30920744.

Scientific
Research

A Multivariate Poisson Model of Consumer Choice in a Multi-Airport Region

Andrew J. BUCK*, Erwin A. BLACKSTONE, Simon HAKIM

Department of Economics, Temple University, Philadelphia, USA;
*Corresponding Author.
Email: buck@temple.edu

ABSTRACT

Using the results of a unique telephone survey the frequency of consumer flights from airports in a multi-airport region are modeled using a multivariate Poisson framework, the parameters of which were estimated using a latent variable application of the expectation maximization algorithm. This offers a different perspective since other work on airport choice uses the results of airport intercept surveys that capture only a single choice per respondent, whereas the data from the phone survey is count data for the airports in the study. An airport's own-distance had the expected negative impact on mean usage of the airport, although the cross effects were somewhat mixed. Ticket price differences between airports were not always statistically significant. Mean usage was found to be increasing in income for PHL, but was decreasing for the other airports, reflecting the increasing value of respondents' time as their income rises. If the destination of flights is domestic (international) then the result is to increase usage of PHL, BWI and EWR (JFK). Except for JFK, if the purpose of travel is mostly pleasure then it results in more travel from JFK and less from the other three airports. The availability of a low cost carrier would result in more frequent travel.

1. Introduction

Using the results of a unique telephone survey the frequency of consumer flights from airports in a multi-airport region are modeled using a multivariate Poisson framework. This offers a different perspective from previous research in two important ways. First, other work on airport choice uses the results of airport intercept surveys that capture only a single choice per respondent, whereas the data from the phone survey used in this paper is count data for the four airports in the study. Second, models based on intercept surveys uniformly use binary choice models such as either probit or logit methods to estimate the model parameters of the mutually exclusive choices [1–3]. The consumers in the present study are observed to choose from among four airports on a repeated basis, resulting in a n-tuple of count data.

Modeling count data requires use of Poisson or negative binomial specifications. The present study expands the usual statistical count model to the appropriate n-tuple count model in the form of the multivariate Poisson so that the counts can have non-zero covariances. The fundamental difference between earlier work and that presented here is the difference between allocation modeling and modeling at the extensive margin. Until recently the

use of multivariate Poisson regression was not an option [4]. An expectation maximization algorithm is used to estimate the parameters of a multivariate Poisson model of consumer decisions.

Until 1983 the Civil Aeronautics Board (CAB) was responsible for regulating airfares in the United States. As a consequence of that regulation commercial passenger carriers competed on many dimensions other than price. Such behavior was recognized as being economically inefficient: the price system was not being allowed to direct resources to their greatest value in use. The CAB was dismantled on the premise that price competition among carriers would benefit consumers and direct productive resources to their greatest value in use. It was felt that, inter alia, the threat of entry would be sufficient to prevent airlines from being able to exploit apparent monopoly power. That premise ignores the fact that consumers are an essential element in the exercise of market power. If consumers do not search for low fares or fare differences are unimportant, then it is unlikely that the threat of entry will have much impact on the fare structure: The effect of the entry of a low fare carrier will only be the reallocation of fliers among carriers at an

airport, with little impact on the allocation of passengers among airports. Indeed, one of the current stylized facts about air travel is that there is more variation in price among carriers at an airport than among airports. It is possible to evaluate the effect of low fares on consumer behavior, and by implication the likely success of the threat of entry as a disciplinary device, by examining multi-airport markets. The unwillingness of flyers to travel to other airports to obtain lower fares increases the ability of carriers to exploit monopoly power and discriminate in prices[1]. Since broad geographic markets are often used in merger cases[2] our analysis may shed some light on such markets.

Heretofore airport choice studies have focused on the choice of airport for a particular trip using intercept surveys of travelers in the chosen airport. Ashford and Benchemam [5] studied airport choice in central England for the period 1975–1978. Among business travelers distance to the airport was the most important variable, followed by frequency of service. Fare was found to be most important among those traveling for pleasure. Caves *et al*. [6] found that access time, frequency and fare to be significant variables in a model of choice between mature and emerging airports in England. Thompson and Caves [7] used data for 1983 to study airport choice in northern England. For both business and leisure travelers distance to the airport and number of available seats were important. Frequency of service was also important for business travelers. In the San Francisco market Harvey [8] found access time and frequency of service to be determinative. None of these earlier efforts would lead one to believe that the difference in fares from different airports would lead to more competition among carriers, or that fare differences could lead to the reallocation of market share among airports. More recent studies, using various modifications of the multinomial logit model also confirm the importance of access time and frequency of flights in airport choice [9–13]. Interestingly, cost was also of secondary interest in the choice of airport by air freight carriers [14]. Gosling [15] offers a comprehensive review of the literature. The lack of searching for the best fare among airports is perhaps understandable given the time cost of travel to a lower fare airport may swamp any differences in fares.

In spring of 2000 a phone survey was conducted of

residents of the market area of Philadelphia International Airport (PHL). The eventual goal of PHL was to learn about its customer base with an eye to increasing its market share in a multi-airport region. PHL management considers its facility in competition with its large neighbors to the north and the south: JFK International, Newark International (EWR) and Baltimore-Washington International (BWI). The relevant market was defined by PHL's management; see Fgure 1 for a map of the market. Newark is the largest of the four and Baltimore-Washington is the smallest.

The 1100 respondents in the final sample[3] were asked a wide variety of questions about their travel and airport usage. From the survey data both univariate and multivariate Poisson models of airport usage were estimated. A preference for using a low fare airport was expressed by survey participants. A rising fare premium for using PHL resulted in higher mean use for Newark (EWR), Baltimore (BWI) and New York (JFK). The fare premium was also positive for use of PHL, reflecting that market power of PHL's dominant carrier at the time of the survey. The fare coefficients were not always statistically significant. Apparently respondents liked the idea of using a low fare airport but did not base their eventual choice on fare differences. As a new entrant in a multi-airport region, a discount airline should enter at that airport where there is the greatest opportunity for winning market share from incumbents without relying on attracting new passengers from other airports.

Income was a significant variable in the use of the three distant airports: BWI, JFK and EWR. Higher income increased the likelihood of flying from either JFK or BWI in the previous year, but the sign is reversed for BWI. If distance from the respondent's residence to the airport was an important consideration then it increased their likelihood of using any of the airports. The actual distance had the expected own airport effects and cross effects. If the purpose of the trips was predominantly business than respondents were more likely to fly from PHL, BWI, and EWR, but not JFK.

2. The Model

The phone survey used to assemble the data asked respondents to think about all of their travel in the prior year. This precluded directly asking about choice of airline as could be done in an intercept interview in an airport. Consequently the model used here addresses only the frequency of having chosen an airport in the prior year, although the respondents were asked about the importance of being able to use their carrier of choice in their selecting an airport.

Over a very short interval of time the decision about which airport to use can be cast as either an index function model or a random utility model [18,19]. In the index function approach the agent makes a marginal bene-

---

[1]At the time of our study US Airways garnered at least 60 percent of the business at Philadelphia International Airport, and in 2005 after the entry of low fare carriers they still had 63% [16]. At sixteen large airports the leading carrier had at least 50 percent of airline departures in 2000 [17].

[2]For example, in hospital merger cases the geographic market has been considered to be as large as 100 miles.

[3]In the sample 827 respondents had traveled outside the region, not necessarily by air, and only those respondents were included in the estimation. A survey research firm conducted the phone interviews. Calls were made, nearly 5000, until there were 1100 complete responses.

fit—marginal cost calculation based on the utility achieved by choosing to fly from a particular airport between one origin-destination pair instead of another. The difference between benefit and cost is modeled as an unobservable variable y* such that

$$y* = x'\beta + \varepsilon \tag{1}$$

The error term is assumed to have a particular known distribution. The net benefit of the choice is never observed, only the choice itself. Therefore the observation is

$$y = \begin{cases} 1 & if \ y* > 0 \\ 0 & if \ y* \le 0 \end{cases} \tag{1'}$$

and x'β is known as the index function.

The preponderance of airport choice studies rely on intercept interviews in the airports. Consequently the respondent has made an airline and airport choice from among mutually exclusive alternatives in a short interval of time. In this context a multinomial logit or multinomial probit model is appropriate (see the earlier citations).

The individual studies and the methodological approach reviewed above all suppose that in a short time interval the economic agent is choosing from among mutually exclusive alternatives. In the phone survey conducted for the Philadelphia International Airport the respondents were not at a particular airport, having made a travel mode decision. Rather, they were at home and were asked to reflect on all the choices that they had made in the previous year. If the decision to fly from an airport is made a large number of times during the year, with a small probability of flying in each interval then in the limit the observed Bernoulli process

---

[4]Alternative methods for modeling count data are References [21–23] Aitchison and Ho propose the use of a Poisson and log normal mixture to model multivariate count data. The mixture involves a Possion specification of the counts with a multivariate log normal distribution over the Poisson rate parameters. This approach permits negative correlations between the counts, which does no occur in the data used here. Further, their model is more flexible with regard to over dispersion in the marginal distributions. In the data used here the over dispersion is not observed in the joint distribution. Finally, they state that their model cannot describe the variability of multivariate counts with small means and little over dispersion, the case here. Terza and Wilson use a mixed multinomial Poisson process to model event frequencies. Built into their approach is the problem of the independence of irrelevant alternatives and no covariance between choices. Shonkwiler and Englin use a multinomial Dirichlet negative binomial process to model a system of incomplete demands. In their approach the covariance between trip choices must be negative. The procedure used here does not suffer from the independence of irrelevant choices problem but restricts the (Y_i, Y_j) gross covariances to be positive, although covariates can have negative coefficients. All three alternatives to the multivariate Poisson are mixtures. As such, they are in the spirit of Bayesian modeling since one must make a specific assumption about the mixing distribution.

of (1') is a Poisson random variable [20]. Having flown from, say, Newark Airport at least once in the year does not preclude having flown from another airport, perhaps several times, during the same year. Hence, the cost-benefit calculation of (1) is made many times during the year for each of the airports in the region. Since the net benefits of flying from a particular airport a given number of times is unobserved, the observed data on the dependent variable is the quadruplet $y_1 \ge 0$, $y_2 \ge 0$, $y_3 \ge 0$, $y_4 \ge 0$. The count data in y_1, y_2, y_3, and y_4 are not independent of one another.

Since the frequency of flying from any one of a choice of airports is by its nature an n-tuple of counts, the appropriate statistical model must be multivariate with non-zero correlations. With this in mind the choice model for the four airports included in the Philadelphia International Airport study of (1) becomes a multivariate Poisson model[4] derived by Mahamunulu, 1967 and is of the form

$$P(Y) = \left[\prod_i^4 \varphi(y_i) \right] \sum_{r=0}^{\infty} \left[\frac{1}{r!} \left\{ \sum_{i<j}^4 \mu_{ij} K_r(y_i) K_r(y_j) \right\}^r \right] \tag{2}$$

where $K_r(y_r) = (-1)^r \Delta^r \varphi(y_i - r)/\varphi(y_i)$ is the Charlier polynomial and $\phi(y)$ is the Poisson probability density function. The problem with the representation in (2) is that it is an infinite series and is therefore not directly empirically implementable.

Fortunately there is a much simpler representation of the multivariate Poisson using unobserved, or latent, variables. With specific reference to the frequency of choosing from among the four airports, consider a vector $X = (X_1, X_2, X_3, X_4, X_{12}, X_{13}, X_{14}, X_{23}, X_{24}, X_{34})^T$ where the X_{ij} are independent latent random variables and each follows a Poisson distribution. The mean of this vector is then $\theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_{12}, \theta_{13}, \theta_{14}, \theta_{23}, \theta_{24}, \theta_{34})^T$. Now define the four element vector of observable frequency of flights from each of the four airports as $Y = AX$ where A is defined as

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \tag{3}$$

Under this specification of the problem each of the y_i is the sum of a specific four member subset of ten independent Poisson random variables. That is, the marginal probability function for the random vector Y can be written as

$$\Pr\begin{bmatrix} Y_1 = y_1 \\ Y_2 = y_2 \\ Y_3 = y_3 \\ Y_4 = y_4 \end{bmatrix} = \begin{bmatrix} \dfrac{\exp(-\theta_1 - \theta_{12} - \theta_{13} - \theta_{14})(\theta_1 + \theta_{12} + \theta_{13} + \theta_{14})^{y_1}}{y_1!} \\[2mm] \dfrac{\exp(-\theta_2 - \theta_{12} - \theta_{23} - \theta_{24})(\theta_2 + \theta_{12} + \theta_{23} + \theta_{24})^{y_2}}{y_2!} \\[2mm] \dfrac{\exp(-\theta_3 - \theta_{13} - \theta_{23} - \theta_{34})(\theta_3 + \theta_{13} + \theta_{23} + \theta_{34})^{y_3}}{y_3!} \\[2mm] \dfrac{\exp(-\theta_4 - \theta_{14} - \theta_{24} - \theta_{34})(\theta_4 + \theta_{14} + \theta_{24} + \theta_{34})^{y_4}}{y_4!} \end{bmatrix}$$

(4)

The mean vector for Y, the frequencies for flying from the four different airports, is given by

$$A\theta = [\theta_1 + \theta_{12} + \theta_{13} + \theta_{14} \quad \theta_2 + \theta_{12} + \theta_{23} + \theta_{24}$$
$$\theta_3 + \theta_{13} + \theta_{23} + \theta_{34} \quad \theta_4 + \theta_{14} + \theta_{24} + \theta_{34}]^T$$

(5)

The frequencies with which an individual flies from the airports are pair-wise correlated and the covariance matrix for Y is

$$A\Sigma A^T = \begin{bmatrix} \theta_1 + \theta_{12} + \theta_{13} + \theta_{14} & \theta_{12} \\ \theta_{12} & \theta_2 + \theta_{12} + \theta_{23} + \theta_{24} \\ \theta_{13} & \theta_{23} \\ \theta_{14} & \theta_{24} \end{bmatrix}$$
$$\begin{bmatrix} \theta_{13} & \theta_{14} \\ \theta_{23} & \theta_{24} \\ \theta_3 + \theta_{13} + \theta_{23} + \theta_{34} & \theta_{34} \\ \theta_{34} & \theta_4 + \theta_{14} + \theta_{24} + \theta_{34} \end{bmatrix}$$

(6)

For estimation of the rate parameters, θ, let the vector $y_i = (y_{i1}, y_{i2}, y_{i3}, y_{i4})'$, $i = 1, 2, \ldots, n$ denote the observations on the frequency of flights from the four airports. To ease the notational burden define the set $S = R_1 \cup R_2$, where $R_1 = (1,2,3,4)$ is an index set over the means of the latent Poisson variates unique to each airport and $R_2 = (rs,$ with $r,s = 1,2,3,4$ and $r < s$) is an index set over the latent Poisson variates that create the covariances between the observed count data for the airports. The observable data is characterized as a 4-variate Poisson denoted $Y_i \sim$

---

[5]The survey instrument is in an appendix available from the authors.
[6]PHL management and consumers may not have the same definition of the relevant market. Unfortunately we were compelled to accept management's definition of the market for the purpose of generating the phone call database. Their definition was based on drive time and the sense that they could effectively market their product to those households within one hour of the airport. Basically they had a market retention mentality.
[7]In effect the data set is a general stratified sample. Ben-Akiva and Lerman [24] address this issue and the estimators for slopes in a choice model. The punch line is that in choice models the estimators are consistent for all except the constant term.

$MP(\theta_i)$ for the $i = 1, 2, \ldots, n$ observations and $\theta_i = \{\theta_{ij}; j \in S\}$ is the vector of parameters for the ith observation. The parameters for the ith observation in turn depend on a vector of independent variables z_{ij}, $j = 1, 2, \ldots, p_j$ through a univariate Poisson regression structure

$$\log(\theta_{ij}) = z_{ij}^{T'}\beta_j \quad i = 1, 2, \ldots, n \text{ and } j \in S \quad (7)$$

and $\beta_j^T = (\beta_{j1}, \beta_{j2}, \ldots, \beta_{jp_j})$ is a p_j vector of regression coefficients.

The unknown parameters are estimated by an expectation maximization (EM) algorithm [4]. The EM algorithm is used for finding maximum likelihood estimates of probabilistic model parameters where the underlying data are unobservable. EM alternates between performing an expectation step and a maximization step. In the expectation step an empirical expectation of the likelihood is computed as though, based on current estimates of the parameters, the latent variables had been observed. In this step the current values for the $\hat{\theta}_i(z, \beta) = \{\hat{\theta}_{ij}; j \in S\}$ are used to construct expected values for the $x_i = \{x_{ij}; j \in S\}$, given the current guess for the parameters what must have been the values taken by the latent variables contingent on the Y observations, and the empirical likelihood is computed. In the maximization step the maximum likelihood parameter estimates $\hat{\theta}_i(z, \beta) = \{\hat{\theta}_{ij}; j \in S\}$ are recalculated on the basis of the expected likelihood computed in the expectation step; given the guesses for the elements of the latent variables in the previous step, how should the parameters by revised in order to maximize the empirical likelihood. In the present context this amounts to fitting univariate Poisson models using the conditional expectations of the estimation step. The open question is the modeling of the rate parameters.

3. The Data

In April and May 2000 a phone survey[5] was conducted on behalf of the management of the Philadelphia International Airport. Approximately 5000 households in a market region defined by the management of the Philadelphia International Airport[6] (shown in Figure 1) were contacted regarding their participation in the survey about travel outside the region and modal choice. The phone contacts were selected from one of two sub-populations; those who had previously expressed an interest in travel and those from the general population[7]. Those who had flown out of Philadelphia International Airport are over-represented in the sample. The resulting final sample had 1100 usable responses, of which 827 had traveled out of the region and 627 had flown out of

Table 1. (a) Airport size 2000; (b) Frequency of usage correlations 2000

	Airlines	Nonstop destinations	Plane movements	Total passengers (enplanements + deplanements)	Automobile Parking Spaces			BWI	JFK	Newark
				(a)					(b)	
BWI	22	61	275,000	19,500,000	12,000		PHL	.2779	.2471	.3537
JFK	57	387	339,597	31,000,000	12,300		BWI		.1347	.0481
EWR	37	543	455,000	33,000,000	17,000		JFK			.0806
PHL	26	111	484,000	24,900,000	6,500		All correlations are significantly greater than zero at the 1% level.			

Table 2. Descriptive statistics

		Variable	Coding	Mean or Frequency
Dependent Variables		Flown from PHL	Counts for flights from airport in prior year.	3.2648
		Flown from BWI		.4812
		Flown from JFK		.2140
		Flown from EWR		.5574
Indirect Utility Arguments		Income	Continuous, dollars.	$67,308.59
		Distance to PHL	Continuous, miles.	26.38
		Distance to BWI		99.41
		Distance to JFK		90.98
		Distance to EWR		76.63
		PHL Premium	Continuous, Cost of flight from PHL over flight from other airport, dollars.	$546.34
Demographics		Age	Years	48.98
		Gender	Female = 1 Male = 0	307 Male
Tastes and Preferences	Purpose of Trips is Mostly Pleasure	PHL	Pleasure = 1 Otherwise = 0	346
		BWI		73
		JFK		61
		EWR		98
	Destination	PHL	Destination is domestic = 1 Otherwise = 0	443
		BWI		132
		JFK		62
		EWR		90
	Will consider use of PHL in Future		Yes = 1 No = 0	527
	Will consider use of BWI, JFK, EWR in future			155
	Importance of airport attribute in choice	Choice of Carrier	Important or Very Important = 1 Otherwise = 0	492
		Distance from home to airport		468
		International flights		360
		Non-stop flights		508
		Low ticket prices		546

one or more of the major airports in the region[8].

Travelers in the Philadelphia region have an abundance of commercial airports from which to choose. At the southern edge of the city is Philadelphia International Airport (PHL). Further to the south are Wilmington and Baltimore-Washington International (BWI). To the northwest is Lehigh Valley International Airport. To the west is Reading Airport. To the east is Atlantic City Airport. To the north are Newark Airport (EWR) and

John F. Kennedy International Airport (JFK). For the purposes of this paper we have modeled only the intensity of usage of the four major airports: BWI, JFK, EWR, and PHL[9]. The sizes of the four airports are indicated by the data in Table 1, Part A. The size rank order depends on the variable in question, although BWI is the smallest of the four by every standard except available parking spaces.

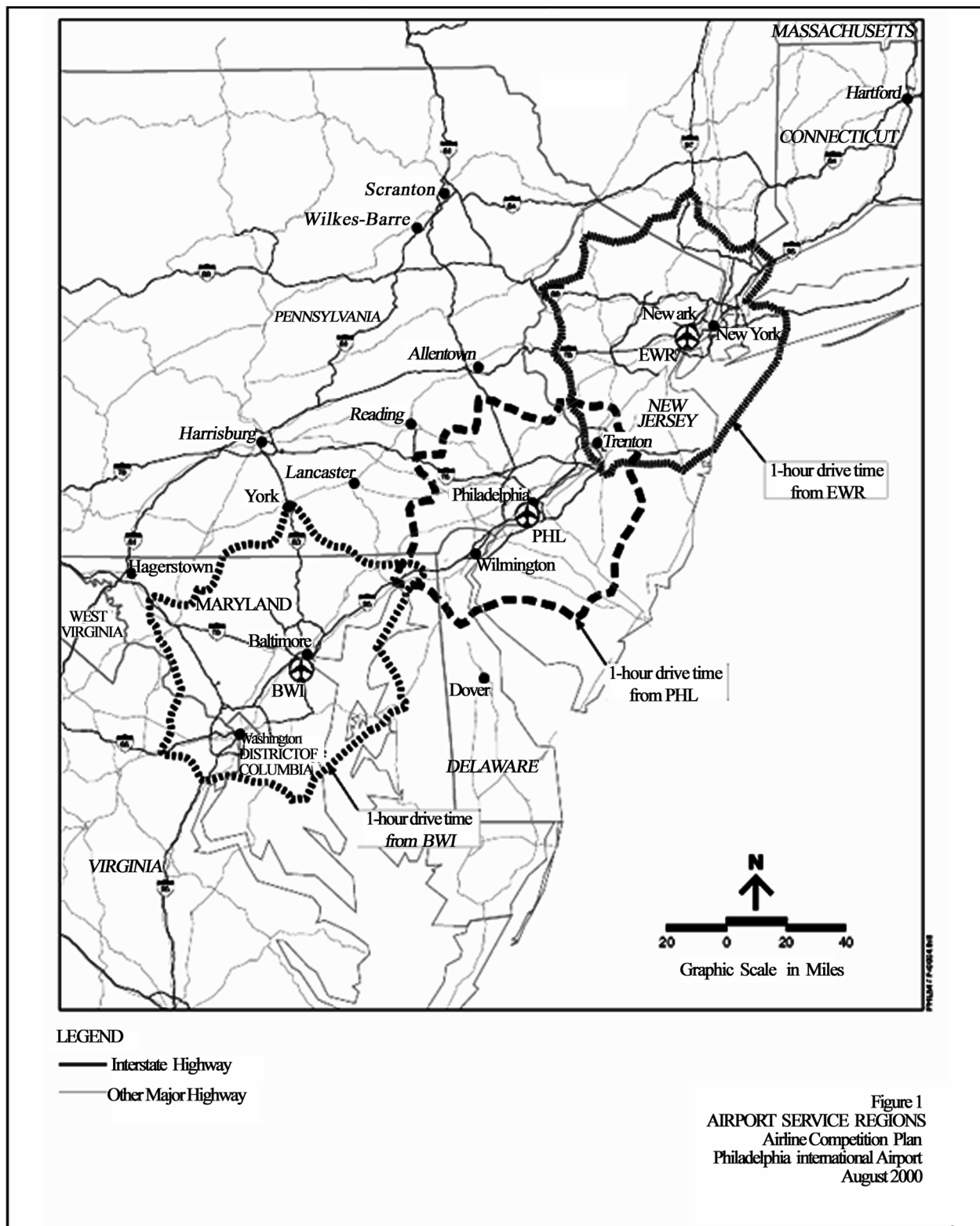Based on the sample data, and relying on the simple

Figure 1. Philadelphia international airport market area

Table 3. Univariate poisson[1,2]

Variable	PHL	BWI	JFK	EWR
Intercept	–4.6502* (30.24)	–2.7635 (1.60)	1.7209 (0.27)	–0.7667 (0.15)
Income	0.0959* (90.09)	–0.0186 (0.37)	–0.0154 (0.18)	0.1589* (41.36)
Distance to PHL	–0.0048* (3.39)	.0030 (0.17)	.0102 (1.01)	.0344* (24.29)
Distance to BWI	.0126* (7.73)	–0.0192* (3.02)	–0.0255 (1.95)	–0.0173 (2.26)
Distance to JFK	.0010 (0.02)	.0209 (0.87)	–0.0297 (1.04)	–0.0288* (3.08)
Distance to EWR	.0171* (8.05)	–0.0277 (2.54)	.0086 (0.14)	–0.0032 (0.05)
PHL Cost Premium	.0004* (52.29)	.0003* (6.20)	.0002 (0.84)	.0005* (23.02)
Purpose of trips	–0.8375* (352.93)	–0.0555 (0.24)	.8321* (11.62)	.3719* (10.92)
Age	.0756* (67.94)	.0894* (9.50)	.0215 (0.42)	.0937* (15.64)
Age2	–0.0009* (85.24)	–0.0010* (9.01)	–0.0004 (1.08)	–0.0011* (17.91)
Gender	–0.2045* (25.56)	–0.0192 (0.03)	–0.2715* (2.78)	–0.2753* (7.95)
Carrier of Choice	.0385 (.68)	–0.0599 (0.28)	–0.1538 (0.72)	.0088 (0.01)
Distance to Airport	–0.2617* (37.06)	–0.2917* (7.68)	–0.0189 (0.01)	–0.1745 (2.50)
International Flight Available	.2899* (48.90)	.2000* (3.48)	.6214* (9.64)	.5235* (30.15)
Non-stop flights available	.1625* (8.86)	–0.3894* (10.18)	.1561 (0.60)	–0.4021* (11.83)
Low ticket prices	.4746* (57.05)	.2682 (1.88)	.2268 (0.86)	.1414 (0.93)
Will consider airport in future	.7042* (118.90)	.4221* (12.33)	.0403 (.05)	.9112* (73.47)
Domestic Destination[3]	.8282* (220.12)	3.7195* (364.78)	2.0284* (66.68)	1.2101* (106.00)
Goodness of Fit[4]	3.2074	0.5975	0.6053	1.4583
Overdispersion[5]	44173.58	384.99	158.20	305.79

[1]Numbers in parentheses are chi-square statistics.
[2]*denotes statistical significance at 10% or better.
[3]Destination for JFK is coded as 1 = International, the reverse of the other airports, for computational reasons.
[4]Goodness of Fit is the scaled deviance. It is a chi-square divided by the degrees of freedom and with an expected value of one.
[5]The overdispersion statistic is computed from Greene [25] and is distributed as Chi-square with one degree of freedom. The 1% critical value is 6.635.

proportions shown in Table 2, EWR and BWI were the most significant competitors for PHL. EWR and JFK are significant competitors only for international travel. Business travelers are much more likely to shift among

the regional airports than are those traveling for pleasure[10]. This is corroborated by the simple frequency of use correlations between airports in Part B of Table 1.

Although the survey was quite comprehensive in its topical coverage, only demographic data, frequency of travel from other airports, preferences regarding airport attributes, and comparison price shopping were used in the empirical model[11]. Descriptive statistics for these variables appear in Table 2. The dependent variables for the model are the frequencies with which individuals in the respondent's household had flown from one of the major airports in the previous year. Of PHL's three rival airports, the greatest proportion reported having flown out of EWR. Given its relative inaccessibility it is not

[8]The model is fit to the 827 households that traveled outside the region; the 273 households that did not travel were excluded from the sample. Excluding households that did not fly because they did not travel may introduce overdispersion. Over dispersion test were performed and the null was not rejected. Any attempt to include these households would have resulted in missing observations excluding them anyway.
[9]BWI is south of PHL on US I-95. EWR is north of PHL on US I-95, and JFK is located on the south shore of Long Island, about forty minutes east of EWR.
[10]Convenience for the business traveler goes beyond access to the airport to include considerations of departure time, connections, etc.

Table 4. (a) Multivariate poisson estimates[1]: Own parameters; (b) Multivariate poisson estimates[1]: Cross parameters

(a)

Variable	PHL			BWI			JFK			EWR		
	Intercept Only	Only Own Covariates	Own and Cross-covariates	Intercept Only	Only Own Covariates	Own and Cross-covariates	Intercept Only	Only Own Covariates	Own and Cross-covariates	Intercept Only	Only Own Covariates	Own and Cross-covariates
Intercept	-0.9926 (0.0707)	-4.8811 (0.8761)	-4.1653 (0.8737)	-1.4863 (0.1034)	-2.6713 (2.2946)	-0.4840 (2.3720)	-2.8730 (0.0755)	1.2358 (3.4014)	-1.1430 (4.2394)	-1.1690 (0.3187)	-4.2767 (2.9462)	6.1828 (2.3599)
Income		0.1009 (0.0420)	0.0873 (0.0103)		-0.0397 (0.0323)	-0.1101 (0.0338)		-0.0175 (0.0373)	0.0277 (0.0424)		0.1682 (0.0317)	0.1630 (0.0284)
Distance to PHL		-0.0051 (0.0027)	-0.0053 (0.0027)		0.0040 (.0076)	0.0056 (0.0080)		0.0103 (0.0104)	-0.0049 (0.0128)		0.0454 (0.0115)	0.0569 (0.0079)
Distance to BWI		0.0129 (0.0047)	0.0104 (0.0047)		-0.0230 (0.0115)	-0.0269 (0.0122)		-0.0227 (0.0186)	-0.0078 (0.0234)		0.0017 (0.0187)	-0.0608 (0.0142)
Distance to JFK		-0.0048 (0.0083)	0.0039 (0.0083)		0.0083 (.0232)	0.0241 (0.0240)		-0.0265 (0..0294)	0.0023 (0.0379)		-0.0346 (0.0218)	-0.0281 (0.0203)
Distance to EWR		0.0232 (0.0062)	0.0122 (0.0063)		-0.0202 (0.0181)	-0.0359 (0.0188)		0.0074 (0.0234)	-0.0108 (0.0328)		0.0032 (0.0174)	-0.0479 (0.0220)
PHLCost Prmium		0.0004 (0.0104)	0.0004 (0.0001)		0.0004 (0.0001)	0.0005 (0.0001)		0.0002 (0.0002)	0.0000 (0.0001)		0.0004 (0.0001)	0.0006 (0.0001)
Purpose of trips		-0.8264 (0.0096)	-0.8274 (0.0455)		-0.1651 (0.1159)	-0.2360 (0.1221)		0.7357 (0.2444)	3.2784 (0.2804)		0.1482 (0.1399)	0.5176 (0.1218)
Age		0.0785 (0.0001)	0.0701 (0.0093)		0.0801 (0.0308)	0.0363 (0.0286)		0.0184 (0.0336)	0.0102 (0.0397)		0.0871 (0.0335)	0.0720 (0.0270)
Age2		-0.0009 (0.0001)	-0.0008 (0.0001)		-0.0008 (0.0003)	-0.0004 (0.0003)		-0.0003 (0.0003)	-0.0003 (0.0004)		-0.0009 (0.0003)	-0.0008 (0.0003)
Gender		-0.1899 (0.0412)	-0.2313 (0.0413)		-0.0056 (0.1096)	0.0535 (0.1145)		-0.2484 (0.1656)	-0.4986 (0.2057)		-0.2114 (0.1303)	-0.2178 (0.1104)
Airline of Choice		0.0226 (0.0448)	0.0491 (0.0478)		-0.1131 (0.1176)	0.1133 (0.1279)		-0.1296 (0.1846)	0.1674 (0.2436)		-0.0146 (0.1366)	0.2332 (0.1239)
Distance to Airport		-0.2390 (0.0429)	-0.2308 (0.0440)		-0.3580 (0.1096)	-0.4093 (0.1154)		-0.0528 (0.1843)	0.2641 (0.2223)		-0.0890 (0.1582)	-0.1549 (0.1239)
Intern. Flight		0.3071 (0.0583)	0.2617 (0.0424)		0.1345 (0.1120)	0.1085 (0.1191)		0.6645 (0.2057)	0.5913 (0.2338)		1.0758 (0.1579)	0.4809 (0.1287)
Non-stop flights		0.2567 (0.0642)	0.1804 (0.0560)		-0.5003 (0.1265)	-0.3728 (0.1326)		0.1326 (0.2035)	-0.4717 (0.2378)		-0.1917 (0.1599)	-0.6086 (0.1256)
Low Price		0.4132 (0.0669)	0.5078 (0.0651)		0.2530 (0.2036)	0.0289 (0.2065)		0.2009 (0.2461)	0.4567 (0.3392)		-0.5590 (0.1793)	0.3051 (0.1683)
Airport in Future		0.6895 (0.0462)	0.7072 (0.0666)		0.4571 (0.1281)	0.7316 (0.1334)		0.0375 (0.1880)	0.0884 (0.2089)		1.4036 (0.1499)	1.2993 (0.1195)
Domestic Destination		.8867 (0.0587)	0.7911 (0.0569)		4.9473 (0.3146)	4.0638 (0.2289)		2.1647 (0.2497)	-0.2652 (0.2765)		1.5270 (0.1448)	0.6469 (0.1323)
Goodness of Fit	7.68	3.81	3.26	2.81	0.67	0.47	1.22	0.95	0.78	2.98	1.49	1.19
Over-dispersion	72.8494	17.8356	12.3120	16.4139	0.2512	0.0517	1.7510	1.4743	0.8327	51.0083	6.7871	8.1977

surprising that JFK was used the least by those participating in the study.

The phone survey was conducted during both daytime and evening hours, still women appear to be over-represented in the sample and the average age of respondents. The respondents' age seems to be somewhat higher than the general population[12]. Any biases introduced by this are ameliorated in part because the questions referred not just to the individual but also to other members of the household. A second ameliorating reason is that those making travel decisions are older than the general population.

To capture the respondent's preferences we questioned them about the importance of different attributes of the airports they choose for their departures: choice of carrier, distance to the airport, availability of international flights, availability of non-stop flights, and presence of a low fare carrier. In response to each named attribute the respondent had to rate the importance of the attribute on a scale from 0 to 5. A 0 meant that the attribute was not at all important in the choice of airport, while a 5 meant that the attribute was extremely important. The categorical variables were recoded as dummy variables in which the dummy took a value of one if the attribute or characteristic was important or extremely important, and zero otherwise. Even with 827 observations this was necessary in order to preserve degrees of freedom since each of five categorical variables would have needed five dummies in each of four equations for a total of 100 coefficients to be estimated in the 'own' latent variable parameters and thirty more in the 'cross' latent variable parameters.

(b)						
	β_{12} PHL-BWI	β_{13} PHL-JFK	β_{14} PHL-EWR	β_{23} BWI-JFK	β_{24} BWI-EWR	β_{34} JFK-EWR
	Intercept Only					
Constant	−1.4203 (0.8889)	−2.1834 (0.1913)	−1.5498 (0.0200)	−6.4847 (0.0728)	−4.4334 (0.1463)	−3.4124 (0.0624)
	Own Covariates Only					
Constant	−3.3126 (0.1819)	−11.1798 (8.2713)	−1.6524 (0.0794)	−5.6071 (10.3666)	−11.9811 (0.5727)	−4.8901 (0.4009)
	Own- and Cross-covariates					
Constant	−158.9580 (70.0804)	−186.1933 (65.0043)	−35.7531 (11.0068)	−168.7721 (57.3307)	−98.3382 (116.3201)	−48.4551 (20.3832)
Gender	0.6821 (3.2583)	9.0643 (3.7696)	−0.1828 (0.4854)	−12.1547 (4.6281)	−5.7411 (8.1581)	3.3350 (2.2894)
Income	2.4149 (0.9592)	0.7534 (1.1107)	0.5740 (0.1370)	3.2353 (0.9174)	−1.6019 (1.8122)	−1.1716 (0.5320)
Age	3.4240 (2.1153)	4.3864 (1.7690)	0.4212 (0.1264)	2.5693 (1.7027)	−0.4105 (3.6229)	1.7731 (0.7870)
Age2	−0.0342 (0.0210)	−0.0291 (0.0143)	−0.0038 (0.0012)	−0.0339 (0.0177)	0.0066 (0.0427)	−0.0296 (0.0115)
Carrier	17.2770 (9.7328)	−0.5263 (2.6063)	−0.3163 (0.4169)	−5.8861 (11.2654)	−30.2447 (5.4843)	2.5751 (1.9926)
Distance	7.5194 (4.7410)	−5.3075 (2.3233)	−1.3789 (0.4175)	6.8751 (3.5287)	9.8577 (11.0292)	1.1548 (2.1540)
International	0.1996 (2.9954)	−9.0871 (3.8296)	0.3889 (0.4060)	3.9001 (4.0224)	11.9829 (3.1542)	3.8812 (3.3086)
Non-stop	−10.8234 (5.9471)	10.4281 (5.2275)	2.4725 (0.7993)	4.3507 (7.0232)	−1.6884 (8.1904)	4.0210 (2.9817)
Pricing	5.6050 (3.7408)	−10.6328 (5.6392)	−3.3539 (0.9096)	−9.417 (12.9732)	12.5169 (8.0006)	3.7280 (3.6626)
Will Use PHL	7.6172 (4.4773)	−4.2938 (5.9367)	−4.3224 (0.9744)	--	--	--
Will Use Other	−16.3339 (6.7274)	−13.4868 (7.9894)	−6.5853 (1.3802)	−8.8263 (5.4286)	−10.7623 (4.7005)	−5.1143 (3.7945)
PHL Premium	−0.0067 (0.0006)	0.0004 (0.0049)	0.0020 (0.0006)	--	--	--
Distance to i	−0.2101 (0.3202)	0.3084 (0.1157)	−0.0014 (0.0160)	0.2039 (0.2592)	0.1501 (0.3319)	0.0499 (0.1969)
Destination from i	16.2027 (10.2507)	−1.6024 (3.6469)	−0.7071 (0.9399)	7.1347 (15.5945)	35.4171 (4.6455)	15.7171 (4.9331)
Distance to j	0.0583 (0.1192)	−0.2442 (0.1237)	0.1003 (0.0222)	0.7135 (0.1322)	0.71345 (0.3797)	−0.0265 (0.1809)
Destination from j	8.5510 (4.0695)	44.2165 (14.4379)	17.2567 (9.7092)	20.4702 (6.4014)	−14.7535 (8.6005)	6.5796 (2.8247)
Purpose i	−9.4457 (5.2732)	13.7023 (4.9055)	3.2956 (0.9270)	2.7957 (5.0255)	−3.0041 (6.8273)	−11.0594 (4.3178)
Purpose j	5.5829 (2.3648)	−24.4502 (9.4837)	−1.8782 (0.6858)	−11.7110 (6.6219)	14.3912 (6.0686)	−1.4067 (2.4231)
[1]Numbers in parentheses are standard errors.						

Only the presence of international flights was of little or no importance to PHL users. This is somewhat surprising given PHL's notoriously poor international service at that time. A surprising 20% of respondents reported that they had compared a fare out of PHL with fares available at other airports. As a follow-up they were also asked about the fare difference in that comparison. For the 165 travelers that made the comparison the average fare premium was $546.34.[13]

4. Empirical Results

The index function that is used here is a mix of indirect utility arguments, such as price premium for flying from PHL, actual distance to the airport and income[14], and tastes and preferences, such as the assessment that using the carrier of choice is important. The survey results included data on the respondents' age and gender[15].

The signs on age and gender are indeterminate a priori, although it is reasonable to expect that frequency of flying and age is a nonlinear relationship. The marginal effect of an increase in income on the probability of using a more distant airport could be negative or positive. As an individual's income rises she finds the opportunity cost of increased travel time to a more distant airport to be a disincentive to using that airport[16]. On the other hand service and fare might overcome that (dis)incentive.

iB

<div align="center">Table 5. Marginal effects on mean number of trips</div>

	Univariate				Multivariate			
	PHL	BWI	JFK	EWR	PHL	BWI	JFK	EWR
DPHL	–0.012	0.0003	0.0011	0.0095	–0.0155	0.0002	–0.0002	0.0114
DBWI	0.0314	–0.0018	–0.0027	–0.0048	0.0303	–0.0010	–0.0003	–0.0122
DJFK	0.0025	0.0020	–0.0032	–0.0079	0.0114	0.0009	0.0001	–0.0056
DEWR	0.0426	–0.0027	0.0009	–0.0009	0.0356	–0.0013	–0.0004	–0.0010
Income	0.2388	–0.0073	–0.0016	0.0438	0.2545	–0.0039	0.0001	0.0328
Cost Premium	0.0010	0.0001	0.00002	0.0001	0.0012	0.00002	0.0000	.0001
Age	–0.0313	–0.0033	–0.0019	–0.0039	–0.0241	–0.0001	–0.0007	–0.0013
Carrier	0.0955	–0.0058	–0.0167	0.0002	0.9183	0.0041	0.0062	0.0460
International	0.7382	0.0196	0.0705	0.1510	0.5652	0.0039	0.0236	0.1007
Non-stop	0.3979	–0.0396	0.0165	–0.1166	0.3749	–0.0142	–0.0190	–0.1329
Low cost	1.1057	0.0249	0.0235	0.0381	1.019	0.0010	0.0162	0.0586
Distance	–0.6642	–0.0294	–0.0021	–0.0487	–0.0674	–0.0287	–0.0127	–0.0311
Purpose	–2.0045	–0.0052	0.1310	0.1188	–2.3191	–0.0146	0.9475	–0.0000
Domestic	2.0526	2.1500	0.6083	0.5686	1.2798	2.0302	0.286	1.7834
Will Use	1.622	0.0469	0.0494	0.3455	1.9068	0.0638	.0043	0.4159
Gender	–0.5239	–0.0018	–0.0397	–0.1039	–0.6965	0.0036	–0.0079	–0.0447

The indirect utility arguments include whether the respondent had obtained the price of a comparable flight from an airport other than Philadelphia and what the price difference turned out to be. One would expect that a consumer's price research would induce them to use the flight departing from the cheaper airport.

Tastes and preferences are modeled from a sequence of questions regarding factors that the traveler finds important in choice of airport as well as the purpose and destinations of trips taken. The survey[17] asked for an ordered response to eight questions regarding airport attributes, although only five are used here[18]. Survey participants could rank an attribute of an airport and its services from 0 to 5; a response of 0 indicated that the factor was not at all important, a response of 5 indicated that the factor was extremely important in the decision making process. Table 2 provides the variables and cor-

responding descriptive statistics.

If ability to choose a particular airline or fly an international carrier is important then one would expect that the respondent would be more likely to have flown out of JFK, all other things equal, given its much wider choice of carriers (See Table 1). People for whom distance to the airport is an important consideration would be less likely to have flown out of JFK. If finding a nonstop flight is extremely important then the respondent should be more likely to have flown out of EWR. The folk wisdom at the time of the survey was that because USAir had dominated PHL for so long it had the ability to charge higher fares. There was no similar carrier dominance in the other three airports. Therefore, if price is an extremely important consideration then a respondent should be less likely to have flown out of PHL in the preceding year.

Both univariate, Table 3, and multivariate, Table 4, Poisson models were fit to the data[19]. For both sets of results measures of goodness of fit and over dispersion are included. Three specifications of the multivariate model for each airport are reported in Table 4. The first specification assumed homogeneity across all respondents and involved estimating the 10x1 vector θ of Equation (5) as though all coefficients except the intercept on the covariates of Equation (7) were zero. The second specification assumed heterogeneity in the θ_i (i=1, 2, 3, 4) but homogeneity in the covariance terms, θ_{ij} (i,j = 1,2,3,4 and i<j). In the third specification all of the θ were treated as heterogeneous across the respondents.

In comparing the univariate and multivariate results there is essentially no change in the sign pattern on the covariate coefficients or which coefficients are significant[20]. The goodness of fit statistics[21] are roughly comparable for the two models. The biggest difference arises in the over dispersion statistics[22]. For the unvaried model the null hypothesis of no over dispersion is rejected for

---

[11]Household size was included in the survey, but was not significant in any of the model specifications. If the dependent variable had been the number of tickets purchased for flights from each airport then household size would have been essential. If the trip or journey is the dependent variable then the number of individuals making the trip is irrelevant. If, say BWI, is the cheaper and closer airport for one member of the family then it is still cheaper and closer when they travel as a group.]

[12]Only respondents indicating that they were over 18 were included in the survey. The survey was conducted only among landline telephone subscribers. In 2004 the percent of the population with 'cell phone only' service was 6.3 percent in metropolitan areas [26].

[13]Or $109 when averaged over the entire sample.

[14]At the time of the phone survey the respondent's 3 digit telephone exchange was captured. Using the airport phone exchanges it was then possible to retrieve the distance from the respondent to each of the airports from a commercially available database [27]. The same database was used to code income as the median for residents of the particular telephone exchange. As part of the survey participants were asked to respond categorically to a question about their household income. Of course not all respondents answered that question. The correlation between our income construct and the categorical responses was 0.87.

each of the four airports[23]. With the exception of the intercepts only specification for PHL the null of no over dispersion is never rejected for the multivariate model. It would appear that the latent variable specification allowing for covariance between airport usage eliminates the over dispersion problem apparent in the univariate models. To put it somewhat differently, the univariate model is not correctly specified. Finally, the θ_{ij} terms, the count covariance latent variables, are statistically different from zero in nine out of twelve instances in the intercepts only and own covariates only versions of the multivariate Poisson models.

Since particular covariates appear in both the coefficient vector of the own-latent variables and the cross-latent variables it is more useful to consider the incremental effects of the covariates on the mean response. The results for both the univariate and multivariate models are summarized in Table 5: Marginal effects on mean number of trips[24]. In the case of continuous covariates the marginal effects are derivatives. In the case of the discrete covariates the model is evaluated for the two values of the dummy variable and the difference computed. All derivatives and differences are evaluated at the means of the covariates.

The effect of distance from a given airport on the frequency of choice of that airport has the expected negative sign for PHL, BWI and EWR. The sign for JFK is positive due to the dominant cross effect between JFK and BWI in part B or Table 4; as one gets further from either one of them one uses one or the other more often. A greater distance from any of the other three airports will increase the frequency of flights from PHL. As a respondent gets further from PHL or JFK, her mean usage of BWI increases. However, as they become more distant from EWR their mean usage of BWI decreases. This is attributable to the geography of the region and cross effects. If one is on the north side of PHL and moves further from EWR then one must be getting closer to PHL, hence there is a shift from BWI to PHL. If one is to the south of PHL and one gets further from EWR then one must be getting closer to BWI and further from PHL. It may be that the attributes other than distance overwhelm the distance effect for BWI. Mean usage of JFK is decreasing in distance from any of the other three airports. This is easy to understand for EWR since a greater distance from EWR means that one is more distant from JFK. If one is more distant from BWI than one must be closer to JFK, but the total distance remains great and PHL is relatively more attractive as a choice. The relative attraction of PHL overwhelms any gain that might be attributed to being further away from PHL, hence the negative sign. The negative sign on distance from JFK in the EWR mean is explained by the fact that being further from JFK means being further from EWR and closer to PHL. Similarly, being further from BWI moves one closer to EWR, but the proximity effect of PHL is overwhelming.

Higher income results in an increase in the mean use of PHL, JFK and EWR, although the effect on use of JFK is numerically very much smaller than that for either PHL or EWR. The sign on income is negative for BWI. As it happens, mean income increases with distance from BWI so there is a confounding income-distance effect for the use of BWI.

At the time of the survey the folk wisdom was that as a consequence of USAir's dominance of PHL that fares out of PHL were higher than the other airports and that travelers would use the other airports to get lower fares. In Table 5 the effect of a greater PHL premium is to increase usage of the other airports. Unfortunately, there is also a positive effect on the mean usage of PHL. This may be due in part to the fact that the effects of distance overwhelm any cost advantage to flying from another airport [28]. This could be thought of as a barrier to customer mobility that results in limit pricing by the carrier: US Airways might charge a premium with the expectation that customers will not defect to another airport.

The coefficients on covariates age and its square are respectively positive and negative, although their aggregate effect on mean use is negative for all four airports. The gender effect is that women fly less often than men from all airports but BWI. If the purpose of one's trips was mostly for pleasure then one would use JFK more frequently and the others less frequently, on average. At the time of the survey PHL's choice of carrier, international, and non-stop service was poor. When traveling for

---

[15]Gender of respondent may be serving as a proxy for many different aspects of the airport choice process. Including it in the models has a small effect on statistical efficiency, but excluding a relevant variable introduces bias.

[16]The geographically more distant airport does not always mean greater travel time. Traffic congestion, high speed rail links, etc. may result in less travel time to the more distant airport. For the airports in the region under study greater distance translates to greater travel time.

17The survey instrument is available from the authors.

[18]The omitted questions include ease of parking, ease of check-in, and presence of public transit. For any given airport the variability in categorical rating was quite narrow so the variables were omitted from the analysis.

[19]The univariate model was fit using PROC GENMOD in SAS. The multivariate EM estimation algorithm was programmed in MATLAB. The starting values for the MATLAB program were taken from the univariate results. The convergence criterion for the EM algorithm was a percent change in the empirical log likelihood of less than 1×10^{-12}.

[20]Similar sign pattern, statistical significance and coefficient magnitude should not be confused with goodness of fit. The goodness of fit statistics are higher in the multivariate specification. Even if the covariates of the covariance structure were not significant for a specific sample in the multivariate model it would not reduce the importance of the approach.

[21]The goodness of fit statistic is the scaled deviance (SAS 9.0).

[22]The over dispersion statistic is the lagrange multiplier statistic from Greene [25].

[23]It is worth noting that the over dispersion in the univariate models leads to overstating the significance of the individual coefficients.

pleasure, and time in transit has a lower opportunity cost, one might be more inclined to use a more inconvenient airport in order to get the desired service attributes. If destination of the trips was domestic then one was more likely to use PHL, BWI and EWR. Since domestic destination for JFK was coded as the reverse of the other three airports the sign must be switched[25]. Thus, if the destination of the trips was international then travelers increased their mean use of JFK.

Six taste and preference questions were included in the specifications: Importance of choice of carrier, importance of international flights, importance of availability of non-stop flights, importance of low fare carriers, importance of distance to the airport, and willingness to use the airport again[26]. In magnitude, the marginal effect of international flight availability on mean usage of PHL is much greater than that for the other airports due to the size of the cross covariates in part B of Table 5; at the time of the survey PHL had the reputation of being very inconvenient for international travelers. Apparently, if an international flight is available at all three airports then a consumer in the PHL market area will be more likely to travel from PHL. Apparently the management at PHL had at least a visceral understanding of this. Since the time of the survey PHL has constructed a new international terminal in order to address the needs of overseas travelers in its market. When the availability of non stop flights is an important consideration travelers use PHL more often and are less likely to use the other airports as often. Table 1 shows that two airports had better non-stop service than PHL, and at that time PHL was not a hub for any of its carri-

---

[24]There are no significance tests indicated in Table 5 since they are unnecessary. One or more of the coefficients on each variable for a given airport is significant so the corresponding effect on the rate parameter will also be significant. The Poisson rate parameter is recovered from the tabled numbers by Equation (7). The mapping from the estimated coefficients to the rate parameter is an affine transformation. Affine transformations preserve ordering and distance. Also, the usual test statistics are scale invariant. Hence, if a significant relationship exists before the transformation it will be significant after the transformation. Greene [25] addresses the same sort of question.

[25]Among those in the sample who had flown from JFK the proportion using that airport to get international service was much greater than those using the airport for domestic service, the reverse of the other airports. As it happened, this switch also resulted in the EM algorithm converging more rapidly.

[26]Willingness to use the airport again is a taste and preference variable to the extent that it reflects changing proclivity on the basis of prior experience. The neoclassical model assumes stable preferences, but in reality preferences do change in the aftermath of experience.

[27]Although USAir dominated the airport by any measure, PHL was not its east coast hub. Its hub remained in Pittsburgh even though it had more traffic in and out of PHL.

[28]Southwest Airlines. The addition of new terminals, another discount carrier and more international service has resulted in PHL being the second fastest growing airport in the world, behind only Beijing.

[29]A response rate of 22% is typical for a phone survey that employs no special devices to increase response rates and rates of cooperation [29].

ers[27]. The importance of the presence of a low cost carrier also had its greatest impact on PHL. Again, this is not surprising since at the time of the survey Airtran, a low cost carrier, had only recently come to PHL. Since the time of the survey PHL has built a short commuter runway, built another domestic service terminal, and added a second low cost carrier[28]. When distance to the airport is an important consideration the effect for all four airports is to reduce the mean number of trips, consistent with the findings for actual distance. Finally, a willingness to use the given airport again will increase the mean use of any of the airports.

5. Conclusions

A multivariate Poisson specification was used to analyze data on the choice of airport from a phone survey of the Philadelphia International Airport (PHL) market. The survey polled nearly 5000 homes to generate a usable sample of 827 respondents that had traveled outside the region[29]. In airport choice studies the respondents are intercepted in an airport and queried about the choice that has brought them to that location instead of others in the choice set. The corresponding appropriate analytical methodology is multinomial logit or probit. The phone survey used here asked respondents to report on all of their air travel in the prior year. Hence, for each respondent there was a count of the number of times she had flown from each of the four airports in the region. Since the count data represents the results of choices made repeatedly over many short time periods it is in principle Poisson distributed.

Since each respondent was flying from among four major airports the correct specification is multivariate Poisson. The multivariate Poisson, which does not have a closed form, can be recast as a latent variables problem that results in marginal distributions for correlated Poisson variates. The parameters in the multivariate Poisson model were estimated using an expectation maximization algorithm.

An airport's own-distance had the expected negative impact on mean usage of the airport, although the cross effects were somewhat mixed. Mean usage was found to be increasing in income for PHL, but was decreasing for the other airports, reflecting the increasing value of respondents' time as their income rises. On balance the quadratic form in respondent's age resulted in less frequent flights among older respondents. A rising fare premium for using PHL resulted in higher mean use for Newark (EWR), Baltimore (BWI) and New York (JFK). The fare premium was also positive for use of PHL, reflecting that market power of PHL's dominant carrier at the time of the survey. If the destination of flights is domestic (international) then the result is to increase usage of PHL, BWI and EWR (JFK). Except for JFK, if the

iB

purpose of travel is mostly pleasure then it results in more travel from JFK and less from the other three airports. The availability of a low cost carrier would result in more frequent travel.

Since the time of the survey the entry of a new low cost carrier and the construction of a new international arrivals terminal have caused PHL usage to increase dramatically. It experienced a 10.5 percent increase in passengers in 2005 alone and a 28 percent increase since 2003. In terms of aircraft activity PHL is now the ninth largest in the world [30].

In summary, given the results of the model, it appears that at the time of the study airlines at Philadelphia International Airport made a profit maximizing decision to take advantage of their regional monopoly. Their prices were high enough to extract monopoly rent while losing only small numbers of passengers to lower cost carriers at other airports. Hence outmigration of potential passengers is not a significant constraint on monopoly power at airports. These results also tend to support smaller geographic market definitions and perhaps even the practice of price discrimination. The entry of Southwest into Philadelphia International Airport may have reclaimed some marginal travelers that had been going to the competing airports, but the biggest impact will be on fare competition among airlines already serving PHL.

REFERENCES

[1] Hess, Stephane, and Polak, J., "Mixed logit modeling of airport choice in multi-airport regions," Journal of Air Transport Management, Vol. 11, No. 2, pp. 59–68, 2005.

[2] Hess, Stephane, and Polak, J., "Airport, airline and access-mode choice in the san francisco bay area airports," Papers in Regional Science, Forthcoming, 2006.

[3] Blackstone, E., Buck, A. J., and Hakim, S., "Determinants of airport choice in a multi-airport region," Atlantic Economic Journal, Vol. 34, No. 3, pp. 3–326, September 2006.

[4] Karlis, D., and Meligkotsidou, L., "Multivariate poisson regression with covariance structure," Statistics and Computing, Vol. 15, pp. 255–265, 2005.

[5] Ashford, N. and Benchemam, M., "Passengers' choice of airport: An application of the multinomial logit model," Transportation Research Record 1147, National Research Council, Washington, D.C., 1987.

[6] Caves, R. E., Ndoh, N. N., and Pitfield, D. E., "Route choice modeling applied to the choice between mature airports and emergent airports in their shadow," Paper presented at the 31st RSA European Congress, Lisbon, Portugal, August 27–30, 1991.

[7] Thompson, A. and Caves, R. E., "The projected market share for a new small airport in the North of England," Regional Studies, Vol. 27, No. 2, pp. 137–147, 1993.

[8] Harvey, G., "Airport choice in a multiple airport region," Transportation Research, Vol. 21A, No. 6, pp. 439–449, 1987.

[9] Basar, G. and Bhat, C., "A parameterized consideration set model for airport choice: An application to the san francisco bay area," Transportation Research Part B: Methodological, Vol. 38, No. 10, pp. 9–904, December 2004.

[10] Hess, S., "An analysis of airport-choice behavior using the mixed multinomial logit model," Center for Transport Studies, Imperial College, London, 2004, http://www.ersa.org/ersaconfs/ersa04/PDF/61.pdf.

[11] Ishii, J., Jun, S., and VanDender, K., "Air travel choices in multi-airport markets," 2005.

[12] Pels, E., Nijkamp, P., and Rietveld, P., "Access to and competition between airports: A case study for the san francisco bay area," Transportation Research Part A: Policy and Practice, Vol. 37, No. 1, pp. 71–83, January 2003.

[13] Windle, R. and Dresner, M., "Airport choice in a multi-airport region," Journal of Transportation Engineering, Vol. 121, No. 4, pp. 332–338, July–August, 1995.

[14] Gardiner, J., Humphreys, I., and Ison, S., "Freighter operators' choice of airport: A three-stage process," Transport Reviews, Vol. 25, No. 1, pp. 85–102, January 2005.

[15] Gosling, G. D., "Southern California association of governments regional airport demand model: Literature review," Project #5437, June 2003.

[16] Belden, T. and Ginsberg, T., "A sign of US Airways' health," Philadelphia Inquirer, pp. A1 and A4, November 16, 2006.

[17] Morrison, S. A., "Airline service," in L. L. Deutsch, Industry Studies, 3rd Edition, M. E. Sharpe, Armonk, NY, pp. 151–152, 2002.

[18] McFadden, D., "Econometric analysis of qualitative response models," In Z. Griliches and M. Intriligator, eds., Handbook of Econometrics, North Holland, Amsterdam, Vol. 2, 1984.

[19] Maddala, G. and Flores-Lagunes, A., "Qualitative response models," in B. Baltagi, ed., A Companion to Theoretical Econometrics, Blackwell, Oxford, 2001.

[20] Ross, S. M., "Introduction to probability models," Academic Press, New York, pp. 174–175, 1980.

[21] Aitchison, J. and Ho, C. H., "The multivariate poisson-log normal distribution," Biometrika, Vol. 76, No. 4, pp. 643–653, 1989.

[22] Terza, J. V. and Wilson, P. W., "Analyzing frequencies of several types of events: A mixed multinomial poisson approach," The Review of Economics and Statistics, Vol. 72, No. 1, 108–115, 1990.

[23] Shonkwiler, J. S. and Englin, J., "Welfare losses due to livestock grazing on public lands: A count data system-wide treatment," American Journal of Agricultural Economics, Vol. 87, No. 2, pp. 302–313, 2005.

[24] Ben-Akiva, M. and Lerman, S. R., Discrete Choice Analysis: Theory and Application to Travel Demand, pp. 235–239, the MIT Press, Cambridge, MA, 1985.

[25] Greene, W., Econometric Analysis, 5th Edition, Prentice

Hall, New York, pp. 743–744, 2003.

[26] Blumberg, S. J., Luke, J. V., and Cynamon, M. L., "Telephone coverage and health survey estimates: evaluating the need for concern about wireless substitution," American Journal of Public Health, Vol. 96, No. 5, pp. 926–931, May 2006.

[27] Jay Computer Services, 9594 208th Street West, Lakeville, MN 55044, http://nt.jcsm.com or whttp://www.jcsm.com, last accessed on 11/09/2006.

[28] Shapiro, H., "The frugal flier: Getting the best buy on an airline ticket means considering much more than the fare," Philadelphia Inquirer, pp. N06, March 28, 2004.

[29] deLeeuw, E. and Hox, J. J., "I am not selling anything: 29 experiments in telephone introductions," International Journal of Public Opinion Research, Vol. 16, No. 4, pp. 464–473, 2004.

[30] PHL (Philadelphia International Airport), "2005 numbers show Philadelphia international airport is now the 9th busiest airport for aircraft operations," 2006, http://www.phl.org/news060228.html.

iB

Scientific
Research

Finding the Efficient Frontier for a Mixed Integer Portfolio Choice Problem Using a Multiobjective Algorithm

K. P. ANAGNOSTOPOULOS, G. MAMANIS

Democritus University of Thrace, Greece.
Email: kanagn@civil.duth.gr, gmamanis@pme.duth.gr

ABSTRACT

We propose a computational procedure to find the efficient frontier for the standard Markowitz mean-variance model with discrete variables. The integer constraints limit on the one hand the portfolio to contain a predetermined number of assets and, on the other hand, the proportion of the portfolio held in a given asset. We adapt the multiobjective algorithm NSGA for solving the problem. The algorithm ranks the solutions of each generation in layers based on Pareto non-domination. We have applied the procedure in sixty assets of ATHEX. We have also compared the algorithm with a single genetic algorithm. The computational results indicate that the procedure is promising for this class of problems.

1. Introduction

Every investor faces the problem of choice the appropriate assets in which he will invest his funds. To support such decisions, H. M. Markowitz set up some fifty years ago a quantitative framework, in which the selected portfolio is optimum with respect to both the expected return and the variance of return and maximizes the so-called utility function [1,2]. The optimal portfolio offers the highest level of expected return for a given level of risk and the minimum level of risk for a given level of return. All such portfolios are called efficient and constitute the efficient frontier. The assumption that asset returns follow the normal distribution allows the finding of efficient frontier via quadratic programming.

However, Markowitz mean-variance model has been criticised not only for the main assumptions it is based upon, but also because it neglects some important aspects of portfolio performance in real life situations. As a result, some other measures of risk have been used, e.g. Value-at-Risk [3,4]; and additional constraints were introduced in the standard model in order, for example, to avoid very small holdings, to restrict the total number of holdings and/or to take into consideration the roundlot of assets that can be bought or sold in a bunch [5].

Since these additional constraints lead to sets of discrete variables and constraints, the resulting optimi-

zation problem becomes quite complex as it exhibits multiple local extrema and discontinuities [4,6–8]. In such situations, especially in large-scale instances of the problem, classical optimization methods do not work efficiently and heuristic optimization techniques are the only alternatives for finding optimal or near-optimal solutions in a reasonable amount of time. Thus, researchers have experimented with the application of heuristic optimization techniques for finding the efficient frontier of the standard Markowitz model enriched with practical constraints. However, it must be noted that, although many metaheuristic algorithms have been developed in the past [9], "few authors seem to have investigated the application of local search metaheuristics for solving the portfolio selection problems" [7].

One of the first attempts for the use of heuristic optimization techniques to portfolio selection was made by Mansini and Speranza [10]. They have formulated the optimum portfolio choice with round lots as a mixed integer programming problem and they have proposed heuristics for its solution based upon the idea of constructing and solving mixed integer sub-problems, which consider subsets of the available investment choices. Chang *et al.* [6] have extended the standard Markowitz model to include cardinality constraints as

well as upper and lower bounds on the proportion of the portfolio invested in each asset. For finding the cardinality constrained efficient frontier the authors have applied three heuristic algorithms based upon genetic algorithms, tabu search and simulated annealing. For the same problem, Anagnostopoulos *et al.* [11] have also proposed a GRASP algorithm enhanced by a learning mechanism and a bias function for determining the next element to be introduced in the solution. Crama and Schyns [7] have also applied a simulated annealing algorithm but they have extended the model to contain not only cardinality constraints and upper and lower bounds, but also trading and turnover constraints. Jobst *et al.* [8] investigated the shape of the efficient frontier of the mean-variance model including buy-in thresholds, cardinality constraints and round lot restrictions using a branch-and-bound algorithm combined with heuristics.

In any case, the construction of the efficient frontier via quadratic programming requires the optimization problem to be solved several times for various values of return. In this paper we confront the standard Markowitz model with cardinality constraints as a bi-objective optimization problem in order to find the efficient frontier in a single execution of the algorithm. The problem is solved by a multiobjective genetic algorithm, which uses a non-dominated sorting procedure to select the best parents. To the best of our knowledge, a few of the related studies in the literature use a proper multiobjective algorithm to construct the Pareto front within the context of a portfolio selection problem such as the one considered in this work. The algorithm was applied in 60 assets of ATHEX and a comparison with a variant of the single (as opposed to multiobjective) genetic algorithm, which has been proposed by Chang *et al.* [6], was realized. The computational results indicate that the procedure is very promising for this class of problems.

The rest of this paper is organized as follows. In Section 2, after a short review of the Markowitz model, the portfolio selection is defined as a multiobjective combinatorial problem. An adaptation of the Nondominated Genetic Algorithm (NSGA) for solving the problem is presented in Section 3. Section 4 is devoted to our numerical results, and some concluding remarks are presented in Section 5.

2. The Formulation of the Problem

2.1 The Markowitz Mean-Variance Model

The problem of optimally selecting a portfolio among N assets was formulated by H.M. Markowitz in 1952. H. M. Markowitz based on the assumption that every investor has the desire to achieve a predetermined return and to minimize risk on investment. Mean or expected return is

employed as a measure of return and standard deviation or variance of return is employed as a measure of risk. Among all portfolios there are special ones for which it cannot be said that one is better than the other. All such portfolios that are Pareto-optimal (or non-dominated) offer the maximum level of return for a given level of risk, or equivalently, the minimum level of risk for a given level of return. The investor should select a portfolio among the efficient portfolios. The proper choice among efficient portfolios depends on the willingness and ability of the investor to assume risk.

However, the main problem is to find this efficient frontier. Under the assumption of the normality of returns, this can be done by solving a quadratic optimization problem for all possible values of ρ, i.e. the desired level of return. The set of all optimal solutions constitutes the mean-variance frontier. It is usually displayed as a curve in the plane where the vertical axis denotes portfolio's expected return, while the horizontal axis represents the variance of this return. Mathematically, the problem can be formulated as follows:

$$\min \ \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \sigma_{ij} \qquad (1)$$

subject to

$$\sum_{i=1}^{N} w_i r_i = \rho \qquad (2)$$

$$\sum_{i=1}^{N} w_i = 1 \qquad (3)$$

$$w_i \geq 0, \quad i = 1,...,N \qquad (4)$$

where:

w_i : the decision variable which denotes the proportion held of asset i

r_i : the expected return of asset i

σ_{ij} : the covariance between assets i and j

ρ : the desired level of return

N : the number of assets available

The objective function (1) minimizes the total variance (risk) associated with the portfolio, while Equation (2) ensures that the portfolio has an expected return of ρ. Equations (3) and (4) describe budget and non-negativity constraints respectively. Budget constraint ensures that 100% of the budget is invested in the portfolio, while non-negativity constraints ensure that no asset has a negative proportion.

An alternative form of the model is often used in practice (see, for example, [6,11] by removing the return constraint and replacing the objective Function (1) by

$$\min \ \lambda \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \sigma_{ij} - (1-\lambda) \sum_{i=1}^{N} w_i r_i \qquad (5)$$

Values of λ satisfying $0 \leq \lambda \leq 1$ represent an explicit tradeoff between risk and return, and generate solutions between the two extremes $\lambda = 0$ and $\lambda = 1$. To draw the efficient frontier, the problem is repeatedly solved using several values of λ.

2.2 The Multiobjective Optimization Model

For more realistic portfolio selection several extensions of Markowitz standard model have been proposed. In real financial decision-making, it is useful to avoid very small holdings, and to restrict the total number of assets. These requirements can be modeled as threshold and cardinality constraints. In general, both lead to sets of discrete variables and constraints.

Threshold and cardinality constraints can be added to the model using a binary variable z_i, which is equal to 1 if the asset i ($1 \leq i \leq N$) is held in the portfolio and 0 otherwise. Introducing finite upper and lower bounds ε_i, δ_i for the stock weight w_i, threshold constraints are represented by the following inequality:

$$\varepsilon_i z_i \leq w_i \leq \delta_i z_i, \quad i = 1,...,N$$

To facilitate portfolio management or to control transaction costs, some investors may wish to limit the number of assets held in their portfolio. The cardinality constraint, which limits the portfolio to contain predetermined number of assets K, can be added to the model by counting the binary variables z_i. This constraint is expressed by the following equation:

$$\sum_{i=1}^{N} z_i = K$$

When such constraints are added, the resulting mixed integer program becomes larger in size and computationally more complex than the standard mean-variance model.

In this paper we reformulate the quadratic optimization problem into a two-objective optimization problem. This allows us to find the efficient frontier in a single execution of the algorithm. The vector objective function has as elements the portfolio return and the variance of return. Moreover our model has been enriched with threshold and cardinality constraints.

The problem to be solved is formulated as follows:

$$\text{opt} \quad f(w) = \left[f_1(w), f_2(w) \right]$$

subject to

$$\sum_{i=1}^{N} w_i = 1$$

$$\sum_{i=1}^{N} z_i = K$$

$$\varepsilon_i z_i \leq w_i \leq \delta_i z_i, \quad i = 1,...,N$$

$$z_i = 0 - 1$$

The objective function $f_1(\mathbf{w})$ represents portfolio's return while the objective function $f_2(\mathbf{w})$ represents portfolio's variance of return. The N-vector $\mathbf{w}$ denotes the set of decision variables w_i.

3. The Multiobjective Algorithm

Multiobjective genetic algorithms have gained much attention last years in solving optimization problems with multiple objectives [12,13]. The primary reason of these studies is the unique feature of genetic algorithms to use a population of solutions. This allows multiple Pareto-optimal solutions to be found in a single simulation run. It appears that the first who tried to use genetic algorithms for finding the Pareto frontier in a multiobjective optimization problem was Schaffer [14]. Although his Vector Evaluated Genetic Algorithm (VEGA) gave encouraging results, it suffered from biasness towards some Pareto-optimal solutions. To overcome this problem, it is suggested the use of both techniques, a non-dominated sorting procedure to move a population toward the Pareto front and some kind of niching technique to keep the GA from converging to a single point on the front. Based on this suggestion a number of independent GA implementations have been proposed, for example the MultiObjective Genetic Algorithm (MOGA) [15] and the Niched-Pareto Genetic Algorithm (NPGA) [16].

Srinivas and Deb [17] proposed the Nondominated Genetic Algorithm (NSGA) which is based on several layers of classifications of individuals. Before selection, a procedure ranks the solutions of each generation in layers based on Pareto non-domination. Firstly, the nondominated individuals are identified so that to constitute the first nondominated front; and they are assigned a large dummy fitness value, which is proportional to population size, to provide an equal reproductive potential to all these nondominated individuals. To maintain diversity in the population classified individuals are shared with their dummy fitness values. Sharing is achieved by dividing each individual's dummy fitness value by a niche count which is proportional to the number of individuals one has in its neighborhood. The parameter niche count for every individual i in the front is calculated by the following equation:

$$c_i = \sum_{i=1}^{M} Sh(d_{ij})$$

where $Sh(d_{ij})$ is the sharing function, d_{ij} is the phenotypic distance between individuals i and j, and M is the number of individuals in the current front. Sharing function is expressed by the equation

$$\Phi(d_{ij}) = \begin{cases} 1 - (d_{ij}/\sigma_{Sh})^{\alpha}, & d_{ij} < \sigma_{Sh} \\ 0 \end{cases}$$

```
P ← randgeneratepopulation() /* Initial population P
generation ← 0
do while generation < maxgenerations
find the vector of decision variables for each individual i ∈ P
compute variance and return ∀i ∈ P
k ← 0
D_k ← ∅
F_k ← ∅ /* the k^th front of individuals
do until P = ∅ begin /*sorting procedure
k ← k+1
for all i ∈ P and for all j ≠ i ∈ P
if for any j, individual i is dominated by j then
D_k ← D_{k-1}∪{i}
else
F_k ← F_{k-1}∪{i}
end if
end for
P ← P – D_k
assign dummy fitness in each i ∈ F_k
apply sharing function in F_k
end do
P ← F_1∪…∪F_k
recombine P according to shared fitness value
mutate P
generation ← generation + 1
end do
```

Figure 1. Pseudocode of the algorithm

Table 1. Crossover example

Parent 1	10	20	30	40	50
	0.1	1	0	0.7	0.8
Parent 2	10	20	5	15	25
	0.9	0	0.6	0.4	1
Offspring 1	10	20	30	5	25
	0.1	0	0	0.6	1
Offspring 2	10	20	40	50	15
	0.9	1	0.7	0.8	0.4

where usually $\alpha = 1$, and σ_{sh} is the maximum distance allowed between two individuals.

Sharing function plays an important role in NSGA's performance, and it is strongly depended on the appropriate selection of the parameter σ_{sh}. The method proposed by Deb and Goldberg for estimating σ_{sh} seems do not to work efficiently in our problem. This is probably due to the additional integer constraints which limit the search space. Thus, the algorithm was executed several times for different values of the parameter σ_{sh}, which was kept smaller than the initial value computed by Deb and Goldberg's method, until the best efficient frontier was found. After sharing, these individuals are ignored temporarily and the second front of nondominated individuals is identified. These new set of points are assigned a new dummy fitness value which is kept smaller than the minimum shared fitness value of the first front (95% of the smallest shared fitness value of the previous front). The process continues until all individuals in the population

are classified.

The population is then reproduced according to the shared fitness value. A stochastic remainder proportionate selection is used in this approach. Since individuals in the first front have the maximum fitness value, they always get more copies than the rest of the population. This allows the search for nondominated regions and sharing helps to distribute the population over this region. The efficiency of NSGA lies in the way multiple objectives are reduced to a dummy fitness function using nondominated sorting procedure. Another aspect is that any number of objectives can be solved and both minimization and maximization problems can be handled [17]. The pseudocode of the algorithm is shown in Figure 1.

A crucial aspect in genetic algorithms is how to represent a solution. The chromosome is divided into two parts. The first part is a set A of K distinct assets and the second one is a set B that includes K real numbers associated with each asset i.

$$A = \{\alpha_1,...,\alpha_K\}, \quad \alpha_i \in \{1, N\}$$

$$B = \{n_{\alpha_1},...,n_{\alpha_K}\}, \quad 0 \le n_i \le 1, \quad i \in A$$

Then, in order to find the proportion of each asset, the free portfolio proportion is calculated as follows

$$fpp = 1 - \sum_{i=1}^{K} \varepsilon_{\alpha_i}$$

Thereafter, the proportion associated with each asset in the portfolio is calculated by the following equation

$$w_{\alpha_i} = \varepsilon_{\alpha_i} + \frac{n_{\alpha_i}}{\sum n_{\alpha_i}} fpp$$

In this way all the constraints are satisfied.

The offspring are generated by uniform crossover as described below. If an asset is present in both parents it is present in the children with the corresponding associated value n. The remaining non-common assets are then selected randomly to fulfill children's sets. An example can be seen in Table 1.

Children are also subject to mutation by multiplying by 0.9 or 1.1 (chosen with equal probability) the value n_i of a randomly selected asset i. The next generation of individuals completely replaces the current population.

4. Computational Results

The algorithm has been implemented in Visual Basic and run on a personal computer Pentium 4 at 2.4 GHz. To construct the data set, 60 assets of big και medium capitalization from Athens Exchange were considered and weekly prices from 10-5-2005 to 12-5-2006 were used to calculate returns and covariances. The weekly
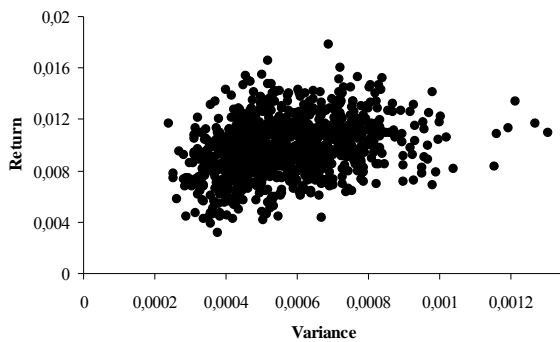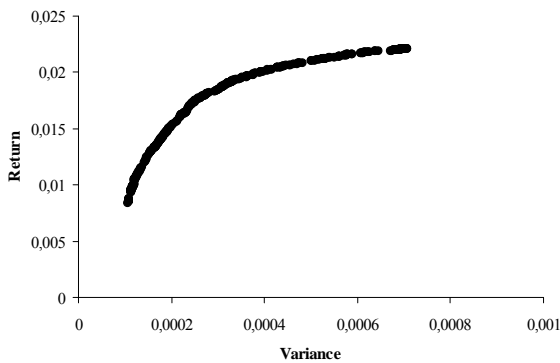
Figure 2. The initial population
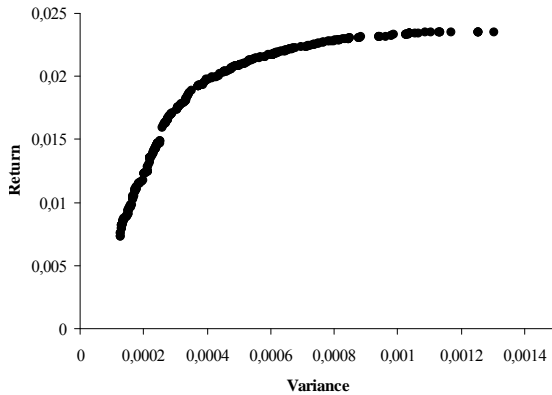

Figure 3. The efficient frontier, K = 10

Figure 4. The efficient frontier, K = 5
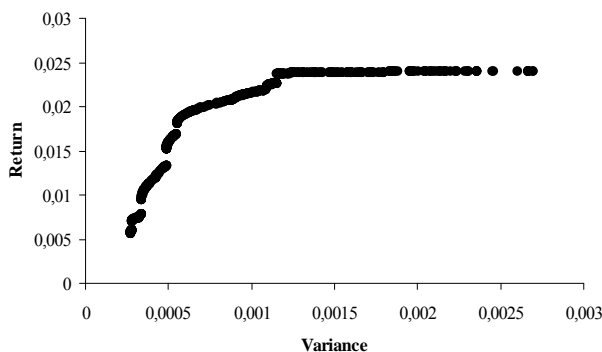

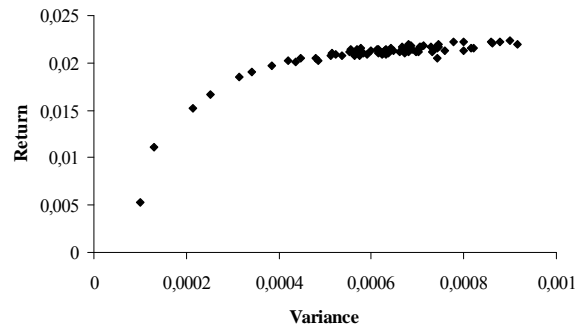**Figure 5. The efficient frontier, K = 2**

Figure 6. The efficient frontier generated by the single genetic algorithm, K = 10.

Table 2. Parameters of the problem

	$K = 10$	$K = 5$	$K = 2$
Population size	1000	1000	1000
Probability of crossover	0.7	0.7	0.7
Probability of mutation	0.1	0.1	0.1
σ_{sh}	0.25	0.15	0.05
Max number of generations	200	100	100
Efficient solutions	448 (44.8%)	385 (38.5%)	750 (75%)

return r_{it} of the asset i in the period t was calculated according the equation

$$r_{it} = \frac{\tau_{it}^e - \tau_{it}^b + d_{it}}{\tau_{it}^e}$$

where τ_{it}^e (τ_{it}^b) is the closing price of asset i at the end (beginning) of period t and d_{it} is the dividend paid to shareholders in period t.

We tried to find the efficient frontier for different values of K and especially for $K = 2, 5, 10$. For all these problems lower and upper bounds were 1% and 100% respectively, i.e., $\varepsilon_i = 0.1$, $\delta_i = 1$ $\forall i \in A$.

In order to see the algorithm performance, an initial population has been randomly generated (Figure 1). Figures 2, 3 and 4 represent the cardinality constrained efficient frontier for $K = 2, 5, 10$ respectively. As we can see from these outputs, the algorithm has found many Pareto-optimal points with good distribution along the efficient frontier. The number of generated points and their distribution are crucial aspects in multiobjective optimization.

If the multiobjective algorithm converges in a small region near or on the true Pareto-optimal front, the purpose of multiobjective optimization is not served. This is because, in such cases, many interesting solutions with large trade-offs among the objectives and parameter values have been probably undiscovered. Table 1 illustrates this distribution of points for each problem instance, together with important parameters of the algorithm.

We have also implemented a variant of the genetic algorithm proposed in [6]. The differences between their genetic algorithm and our algorithm are, on the one hand, the complete replacing of the solutions (as in our multiobjective algorithm) versus the partial replacing and, on the other hand, the rank selection versus the tournament selection. Because of limited space, only some of the obtained results are presented.

In order to compare the quality of solutions obtained by the multiobjective genetic algorithm and the single objective genetic algorithm, we use the technique proposed in [18]. The multiobjective genetic algorithm is considered not worse than the single objective if

$$\sum_{l=1}^{L} \left(s_l \left(\mathbf{w}^{ml} \right) - s_l \left(\mathbf{w}^{sl} \right) \right) \leq 0$$

where s_l is the scalarizing function, $\mathbf{w}^{sl}$ the best solution obtained by optimization of s_l with the single GA and $\mathbf{w}^{ml}$ the best solution on s_l selected from the set of Pareto-optimal solutions generated by the multiple objective GA. Figure 6 shows the solutions obtained by optimizing 81 objective functions ($l = 1,\ldots, 81$) with single objective GA, defined for values $\lambda = 0$ to 1 with step 0.0125 (see Equation (5)). Since

$$\sum_{l=1}^{L} \left(s_l \left(\mathbf{w}^{ml} \right) - s_l \left(\mathbf{w}^{sl} \right) \right) \leq -0.0265 \leq 0$$

we may compare the computational requirements of the two approaches. The effectiveness index is equal to

$$EI = \frac{CT_m}{CT_s} = 25$$

where CT_s is the average running time the single objective GA spent on optimization of s_l and CT_m the running time the multiobjective GA needs to generate the Pareto-optimal solutions $s_1 \ldots s_L$. These results are based on the Pareto front generated from the multiobjective algorithm with 200 generations. If the generations are equal to 50 (although the Pareto front is slightly inferior), the equation is still verified and the EI is equal to 6,025. Thus we can conclude that the generation of the Pareto-optimal solutions with NSGA is competitive both from the computational effectiveness point of view and the quality of the Pareto front.

5. Conclusions

Constraints in the size of the portfolio and in lower and upper bounds on the proportion of the portfolio held in a given asset transform the standard Markowitz model in a mixed integer optimization problem and create discontinuities in the efficient frontier. In this paper we adapt the multiobjective algorithm NSGA for finding the cardinality constrained efficient frontier.

We argue that the proposed procedure solves efficient-ly the cardinality constrained portfolio optimization problem as it generates in relatively short computational time a large number of Pareto-optimal solutions, which are uniformly distributed along the efficient frontier. Even if the efficient frontier is not continuous and, then, competition among solutions may lead to extinction of some sub-regions, the algorithm finds a large number of Pareto-optimal solutions in every segment. On the other hand, the procedure is in general time consuming, since the quality of solutions depends on the population size, but this shortcoming is balanced by the fact that the efficient solutions are obtained after a small number of generations. Finally, a further difficulty is the appropriate selection of σ_{sh} as the algorithm performance is highly dependent on this value.

Constraints in the size of portfolio and in lower and upper bounds on the proportion of the portfolio in a given asset help the decision maker to facilitate its portfolio management; and to avoid excessive transaction costs on one hand, and to avoid holding very small/large amounts of any particular asset on the other. It is empirically known that much of the portfolio risk can be diversified by holding a rather small number of assets [4,19]. We have solved for the efficient frontier following the tradition of standard Markowitz approach, however, focusing on the case where the investor wants to invest in exactly K out of N number of assets. Furthermore, portfolios with positions in assets with very small amounts have been excluded through the use of threshold constraints. The resulting efficient frontier gives the best possible trade-off of risk against return for a particular number of assets (K). The investor then examines the trade-off points in the possibilities curve and selects the one particular point of interest. This may be the point with the lowest variance but having the lowest return, located in the lower left part of the frontier; or it may be the point with the maximum expected return but with the maximum risk, located in the right upper part of the frontier; or it may be any intermediate point. The proper selection of the particular point depends on the investor's willingness to assume risk. In the next step, the investor implements the one particular portfolio whose image is the point in the nondominated frontier. Furthermore, solving for different values of K, the trade-off between risk, return and the number of assets of the portfolio could be examined.

Currently our research focus on a generalization of the cardinality constrained mean-variance problem, by including class constraints that limit the proportion of the portfolio that can be invested in assets in each class, such as bank stocks, telecommunication stocks etc. For its solution, procedures of the so called second generation multiobjective genetic algorithms are tested.

REFERENCES

[1] Markowitz, H. M., "Portfolio selection," Journal of Finance, No. 7, pp. 77–91, 1952.

[2] Markowitz, H. M., "Portfolio selection, efficient diversification of investments," Blackwell, Cambridge MA and Oxford UK, 1990.

[3] Benati, S. and Rizzi, R., "A mixed integer linear programming formulation of the optimal mean/value-at-risk portfolio problem," European Journal of Operational Research, Vol. 176, No. 1, pp. 423–434, 2007.

[4] Gilli, M. and Këllezi, E., "A global optimization heuristic for portfolio choice with VaR and expected shortfall," in Kontoghiorghes, E. J., Rustem, B., and Siokos, S., Eds., Computational Methods in Decision-Making, Economics and Finance, Applied Optimization Series, Kluwer Academic Publishers, pp. 167–183, 2002.

[5] Mitra, G., "A review of portfolio planning: Models and systems," in Satchell, S., Scowcroft, A., Eds., Advances in Portfolio Construction and Implementation, Butterworth-Heinemann, Amsterdam, pp. 1–39, 2003.

[6] Chang, T. J., Meade, N., Beasley, J. E., and Sharaiha, Y. M., "Heuristics for cardinality constrained portfolio optimization," Computers & Operations Research, 27, pp. 1271–1302, 2000.

[7] Crama, Y. and Schyns, M., "Simulated annealing for complex portfolio selection problems," European Journal of Operational Research, No. 150, pp. 546–571, 2003.

[8] Jobst, N. J., Horniman, M. D., Lucas, C. A., and Mitra, G., "Computational aspects of alternative portfolio selection models in the presence of discrete asset choice constraints," Quantitative Finance, Vol. 1, pp. 1–13, 2001.

[9] Blum, C. and Roli, A., "Metaheuristics in combinatorial optimization: Overview and conceptual comparison," ACM Computing Surveys, Vol. 35, No 3, pp. 268–308, 2003.

[10] Mansini, R. and Speranza, M. G., "Heuristic algorithms for the portfolio selection problem with minimum transaction lots," European Journal of Operational Research,

[11] Anagnostopoulos, K. P., Chatzoglou, P. D., and Katsavounis, S., "A reactive greedy randomized adaptive search procedure for a mixed integer portfolio optimization problem," in P.D. Chatzoglou, Ed., Proceedings of the 2nd International Conference on Accounting and Finance in Transition (CD), 2004.

[12] Coello, C. A., "An updated survey of GA-based multi-objective optimization techniques," ACM Computing Surveys, Vol. 32, No 2, pp. 109–143, 2000.

[13] Deb, K., "Multiobjective genetic algorithms: Problem difficulties and construction of test problems," Evolutionary Computation, Vol. 7, No. 3, pp. 205–230, 1999.

[14] Schaffer, J. D., "Multiple objective optimization with vector evaluated genetic algorithms," in Grefenstette, J. J., Ed., Proceedings of the First International Conference on Genetic Algorithms and Their Applications, Lawrence Erlbaum, Hillsdale, New Jersey, pp. 93–100, 1985.

[15] Fonseca, C. M., and Fleming, P. J., "Genetic algorithms for multiobjective optimization: Formulation, discussion, and generalization," in Forrest, S., Ed., Proceedings of the Fifth International Conference on Genetic Algorithms, Morgan Kaufmann, San Mateo, California, pp. 416–423, 1993.

[16] Horn, J., Nafpliotis, N., and Goldberg, G. E., "A niched Pareto genetic algorithm for multiobjective optimization, Piscataway," IEEE Service Center, NJ, pp. 82–87, 1994.

[17] Srinivas, N. and Deb, K., "Multiobjective optimization using nondominated sorting in genetic algorithms," Evolutionary Computation, Vol. 2, No. 3, pp. 221–248, 1994.

[18] Jaszkiewicz, A., "On the computational effectiveness of multiple objective metaheuristics," in Proceedings of the Fourth International Conference on Multi-Objective Programming and Goal Programming MOPGP'00, Theory & Applications, Berlin-Heidelberg, Springer–Verlag, May 29–June 1, 2000.

[19] Maringer, D., "Portfolio management with heuristic optimization," Springer, New York, 2005.

Vol. 114, No. 2, pp. 219–233, 1999.

iB

Scientific
Research

What Types of Small and Medium-Sized Businesses are Utilizing New Financial Products?[*]

Nobuyoshi YAMORI

Graduate School of Economics, Nagoya University, Chikusa-ku, Nagoya, Aichi, Japan.
Email: yamori@soec.nagoya-u.ac.jp

ABSTRACT

The increased diversification of fund rising methods among small and medium-sized businesses has been a major policy challenge in recent years, and private financial institutions are proactively striving to disseminate new financial technologies. However, this does not necessarily mean that every small and medium-sized business benefits from such technologies. It is difficult to analyze this aspect based on ready-made data. Fortunately, this paper can analyze the current status and challenges of the utilization of new financial products among small and medium-sized businesses by using unique survey questionnaires (Kansai RIETI Questionnaires). The results of responses from more than 2,000 companies showed that most companies began utilizing new financial products due to introductions by main banks, and it was seen that the diversification of fund rising methods among small and medium-sized businesses has developed as a result of efforts made by policy-making authorities and financial institutions. However, the rate of utilization of each financial method is as low as a few percentage points, and extremely small businesses or cash-strapped companies has not improved in terms of fund risings. Improved diversification was mainly found in excellent small and medium-sized businesses.

1. Introduction

The increased diversification of fund rising methods among small and medium-sized businesses has been a major policy challenge in recent years. According to the "Status of Progress of the Action Program for the Promotion of Functional Enhancement of Community-Based Finance" published by the Financial Service Agency, private financial institutions have been making effort in the areas of 1) financing methods such as intellectual property-collateralized loans focusing on business values, loan-collateralizing movables or transfer of receivables, nonrecourse loans (loans for which the nonexempt property of performance obligation is limited to those subject to financing), or project finance loans; and 2) securitization such as the issuance of asset-backed securities utilizing local CLO (Collateralized Loan Obligation) or accounts receivable held by small and medium-sized businesses. Generally speaking, substantial achievements have generally been obtained.

However, such efforts are not necessarily equally effective for every small and medium-sized business, and currently, many are still suffering amidst the severe financial environment. Therefore, in order to enable the utilization/dissemination of new financial products to lead toward a real facilitation of financing for small and medium-sized businesses, consideration of what may further be required must be sought. Because it is difficult to analyze such aspects with ordinary data, more detailed corporate data such as questionnaires is necessary. Now, in this paper, the current status and challenges of utilizing new financial products among small and medium-sized businesses are analyzed using survey questionnaires (Kansai RIETI Questionnaires) that the RIETI has implemented.

2. Kansai RIETI Questionnaires

2.1 Summary of Questionnaires

I implemented a survey called "Corporate awareness survey regarding corporate financing in the Kansai Region" in June 2005 as collaborative research with the

Table 1. Utilization of diversified fund procurement methods

Total no. of responses	Leasing	Installment	Lump sum settlement method (transfer of receivables as collateral)	Factoring	Borrowing against accounts receivable (without accounts receivable registration)	Borrowing against accounts receivable (with accounts receivable registration)	Securitization of accounts receivable (liquidation)	Secured privately-subscribed bonds	Credit Guarantee Corporation-backed privately subscribed bonds	Securitization of bonds for financing using Credit Guarantee Corporation	Issuance of publicly subscribed bonds and CP, etc.	Venture capital (contribution from VC)	Contribution from financial institutions other than VC	Procurement from equity market targeted at venture companies	Other
1129	710	127	55	211	28	9	29	62	121	21	24	7	7	2	157
100%	62.9	11.2	4.9	18.7	2.5	0.8	2.6	5.5	10.7	1.9	2.1	0.6	0.6	0.2	13.9

Local Financing Study Group (General manager of the project: Professor Yoshiro Tsutsui, Osaka University) of the Research Institute of Economy, Trade and Industry (RIETI). Survey sheets were sent to 9,000 companies with headquarters located in Osaka, Hyogo, and Kyoto prefectures[1]. Survey sheets were mailed out on June 6, 2005, and the 2,041 companies from which we received responses (return rate: 22.68%) became subjects in the analysis. However, the actual number of companies that responded varied per questionnaire.

The survey sheets that were mailed to the companies consisted of 54 questions. The questions were divided into eight major sections. Part 1 is the attribution of questionnaire respondents. Part 2 is the attribution of respondent companies. Part 3 consists of questions regarding the general management of respondent companies. General management policies are queried in Part 4, while the relationship with main banks is asked about in Part 5. Part 6 deals with questions regarding general bank transactions. Part 7 consists of questions on how small and medium-sized businesses evaluate financial institutions in terms of loan screening and borrowers monitoring, and Part 8 deals with questions regarding the public credit-guarantee system.

2.2 Current Status of Utilization of New Financial Products and Corporate Size

In the RIETI Questionnaire, current utilization of various fund procurement methods were asked as shown in Table 1 (Question 19). According to those questions, leasing

showed an outstanding utilization rate, and 62.9% of the respondent companies utilized leases. Factoring, installment, and Credit Guarantee Corporation-backed privately subscribed bonds follow in this order, and it is clear that fund procurement from non-bank institutions plays a certain significant role. However, such financial methods are not discussed in this paper, in order to facilitate a greater focus on the fund-providing functions of banking institutions.

The five items from the lump sum settlement method (transfer of accounts receivable as collateral) to the securitization of account receivables (liquidation) in Table 1 are, in a broad sense, fund procurement methods utilizing accounts receivable (hereinafter referred to as "Procurement utilizing Accounts Receivable"). Furthermore, secured privately subscribed bonds and Credit Guarantee Corporation-backed privately subscribed bonds are fund procurement by privately subscribed bonds. Then, we investigate how this "Procurement utilizing Accounts Receivable" and "Fund Procurement by Privately Subscribed Bonds" have been utilized and by what types of companies. Since some companies utilize multiple methods[2], excluding duplications, 287 companies have experience with Procurement utilizing Accounts Receivable, while 179 companies have experience in utilizing privately subscribed bonds.

First, respondent companies were divided into ten groups, depending on the corporate size as determined by the number of employees, and the utilization rate of Procurement utilizing Accounts Receivable and "Fund Procurement by Privately Subscribed Bonds" were calculated for graphic description (Figure 1). With regard to privately subscribed bonds, the effect of size is significantly clear. In other words, there is a tendency in which the larger a company is, the higher will be the utilization rate, excluding the largest companies (210 or more employees)[3].

On the other hand, with regard to the utilization rate of

---

[1]For a detailed explanation, refer to Yamori [1]. Yamori [2], Uchida, Udell, and Yamori [3,4] wrote a series of papers based on this Kansai RIETI Questionnaire.

[2]For example, regarding "Procurement utilizing Accounts Receivable," some companies use both "lump sum settlement method" and "factoring."

[3]It is understood that the utilization rate of privately subscribed bonds is low for companies in the largest group because they tend to issue publicly subscribed bonds.
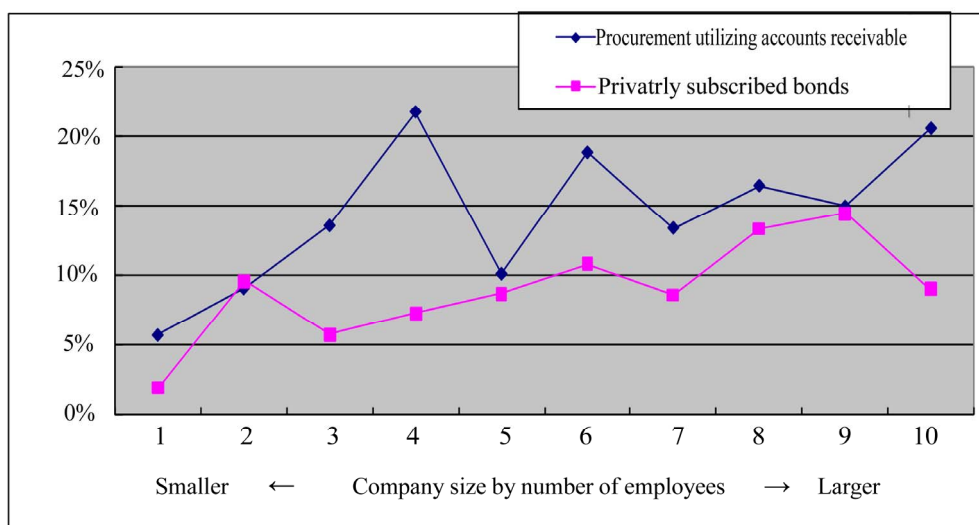
Note: Employee size 1 = 8 or less, Size 2 = 22 or less, Size 3 = 30 or less, Size 4 = 40 or less, Size 5 = 56 or less, Size 6 = 70 or less, Size 7 = 90 or less, Size 8 = 127 or less, Size 9 = 210 or less, and Size 10 = 211 or more. The number of companies belong to each group is about 200

Figure 1. Corporate size and utilization rates of fund procurement methods

Table 2. Status of utilization of new financial products

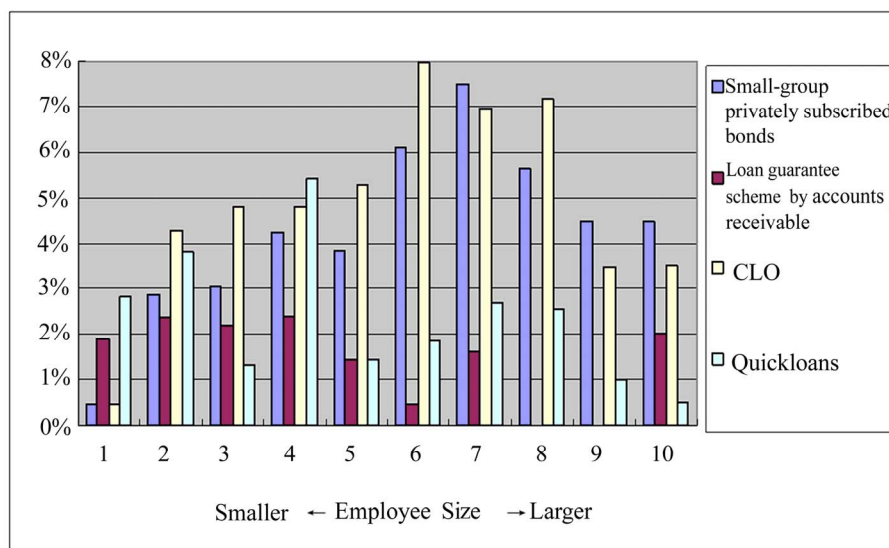|	Total no. of responses	Small-group privately subscribed bonds	Loan guarantee scheme by accounts receivable	CLO	Quick loans
Aware of them	961	590	692	275	427
Utilized within the last 3 years	230	85	29	99	46
Introduction by a main bank was the trigger for utilization	198	76	34	83	36

Figure 2. Employee size and status of utilization of new financial products

Table 3. Utilization rates of new financial products by small and medium-sized businesses and type of operation of main banks

	Employee size	City bank	Regional/second-tier regional bank	credit association (Shinkin Bank)
Small-group privately subscribed bonds	22 or less	2.8%	1.0%	1.0%
	40 or less	5.6%	1.5%	0.0%
	Overall	5.7%	2.7%	2.3%
Loan guarantee scheme by accounts receivable	22 or less	1.4%	2.1%	3.9%
	40 or less	1.0%	3.0%	6.3%
	Overall	0.8%	2.1%	3.5%
CLO	22 or less	4.8%	2.1%	1.0%
	40 or less	7.1%	3.0%	3.1%
	Overall	7.2%	2.4%	2.3%
Quick loan	22 or less	2.8%	3.1%	6.9%
	40 or less	2.5%	6.0%	1.6%
	Overall	1.5%	3.0%	4.2%

Table 4. Evaluation of main banks and status of utilization of new financial products

		Evaluation of main bank (1: Excellent to 5: Very poor)				
		1	2	3	4	5
Know your company	Overall	28.1%	48.7%	18.4%	3.7%	1.2%
	Small-group privately subscribed bonds	38.8%	43.8%	15.0%	2.5%	0.0%
	Loan guarantee scheme by accounts receivable	19.2%	61.5%	15.4%	3.8%	0.0%
	CLO	17.6%	65.9%	14.3%	1.1%	1.1%
	Quick loans	19.5%	58.5%	17.1%	2.4%	2.4%
Stable provision of funds	Overall	26.2%	43.6%	23.2%	4.4%	2.5%
	Small-group privately subscribed bonds	31.3%	45.0%	21.3%	1.3%	1.3%
	Loan guarantee scheme by accounts receivable	30.8%	46.2%	15.4%	0.0%	7.7%
	CLO	31.9%	53.8%	12.1%	1.1%	1.1%
	Quick loans	14.6%	58.5%	14.6%	7.3%	4.9%
Provision of funds at a low interest rate	Overall	10.7%	38.0%	36.7%	10.8%	3.9%
	Small-group privately subscribed bonds	10.0%	45.0%	30.0%	12.5%	2.5%
	Loan guarantee scheme by accounts receivable	7.7%	50.0%	26.9%	11.5%	3.8%
	CLO	13.2%	39.6%	39.6%	6.6%	1.1%
	Quick loans	4.9%	31.7%	43.9%	17.1%	2.4%

Procurement utilizing Accounts Receivable, the larger the size, the higher is the rate in the four groups having up to 40 employees, but no clear trend is seen in other size groups. What is clearly found is that those with employee sizes 1 and 2 (that is, companies with 22 or fewer employees) have a lower utilization rate of Procurement utilizing Accounts Receivable compared to other size companies. Conversely, according to responses to question asking about difficulties in fund procurement during the past year (Question 10), small businesses responded more frequently that "procurement is difficult"[4]. Therefore, the diversification of fund procurement has not significantly disseminated among small businesses, despite the fact that they suffer more from fund procurement difficulties. This trend is also seen in the pretax profit status. Namely, while 9.2% of the companies with two consecutive business years in the red utilized Procurement utilizing Accounts Receivable, 14.7% of the companies with two consecutive business years in the black utilized this method. The same applies to privately subscribed bonds; specifically, while the utilization rate for companies with two consecutive business years in the red was 2.8%, companies with two consecutive business years in the black showed a rate of 9.8%[5]. This indicates that the diversification of fund procurement is popular in well-run companies with comfortable fund management.

2.3 Roles of Main Banks in the Utilization of New Financial Products

We ask the roles of main banks as the trigger for deciding to use new financial products of "(1) small-group privately subscribed bonds," "(2) loan guarantee scheme by accounts receivable," "(3) CLO," and "(4) Quick loan," for the past 3 years. Increased dissemination of these products has been supported by policies for facilitating finance for small and medium-sized businesses. In the RIETI questionnaires, the "small-group privately subscribed bonds" is defined as corporate bonds that solicit undertaking subject to less than 50 private individuals (those who have a relationship with the company, such as executives, employees, and clients), the "loan guarantee scheme by accounts receivable" as the scheme

that Credit Guarantee Corporation guarantees when borrowing from financial institutions is made, by having accounts receivable as collateral, the "CLO" as the fund procurement from the financial market by selling securities backed by loans to investors, and the "quick loan" as the product in which financial institutions quickly determine whether to finance or not by using the credit scoring model, etc.

While the result is summarized in Table 2, in the case of "(1) small-group privately subscribed bonds," about 90% of those (85 companies) that had used within the previous 3 years responded that "the trigger of use was an introduction by a main bank." In sum, many were triggered basically by the introduction from main banks for any new financial product, although the introduction rate was slightly lower in the case of a quick loan[6]. This indicates that it is difficult for small and medium-sized businesses that have poor financial information or knowledge to consider adoption of new financial methods by themselves, and the roles of financial institutions are significantly larger.

Figure 2 shows the utilization rates calculated by the same employee size classification as Figure 1 for these four financial products. By looking at this figure, in the case of small-group privately subscribed bonds, the utilization rate is higher in the group of companies with employee sizes 6 to 8, and almost no utilization is implemented in the smallest size groups. CLO has the same tendency as the small-group privately subscribed bonds. On the other hand, the loan guarantee scheme based on accounts receivable shows an exceptionally high rate in the largest size group, but other than that, the utilization rate is clearly quite high in the smaller employee size groups. However, the rate remains at about 2% at the highest. The same can be said for quick loans, showing a higher utilization rate in smaller employee size groups. Furthermore, although it is not shown in a figure or table, when the status of utilization of these four financial methods in the current earnings situation was investigated, companies with poor performance utilized the loan guarantee scheme by accounts receivable or quick loans, while companies with good performance often utilized small-group privately subscribed bonds or CLO.

2.4 Relationship between Diversification of Fund Procurement and Main Banks

Next, the utilization rate by main banks categories regarding the four financing methods analyzed in Table 2 was investigated. As a result (Table 3), with regard to privately subscribed bonds and CLO, the utilization rate was high among companies for which the main banks were city banks, and conversely, concerning the loan guarantee scheme based on accounts receivable or quick loans, the utilization rate was higher among the firms for which the main banks were regional banks or credit as-

---

[4]For example, while the rate for "difficult" in employee size 1 is 39.7% and 29.6% in employee size 2, the largest companies showed a rate of only 9.4%, which indicates an overall average value of 22.5%.

[5]Clearly, this does not mean that firms with two consecutive business years in the red have leeway for funding. According to the results of questions regarding funding difficulties, while 61.6% of the companies with two consecutive business years in the red responded, "difficult," 14.9% of the companies with two consecutive business years in the black responded similarly.

[6]The number of instances of introduction is larger than the number of instances of utilization in the "(2) loan guarantee scheme by accounts receivable." It is possible that there may have been false responses. However, as the question was asking "in the last 3 years" for utilization it is believed that companies that utilized earlier than that period responded as such.

sociations (or Shinkin Bank).

It is possible that this may simply reflect differences by type of operation of average sized companies that are used as main banks. To control this size factor, Table 3 also shows the figures by companies with 22 or fewer employees and those with 23 to 40 employees. This shows that companies for which the main banks were city banks utilized privately subscribed bonds or CLO more than those for which the main banks were regional banks or credit associations, even within the same employee size group. This indicates that not only the corporate situations but also the enthusiasm or capabilities of financial institutions affect the diversification of fund procurement.

The rate of satisfaction with current main banks was also asked in our survey questionnaires (Question 35). The relationship between the evaluation of current main banks and the use/non-use of four financial methods was investigated according to the three items of "know your company," "provision of funds at a low interest rate," and "stable provision of funds" among them.

The results are summarized in Table 4. First, from the perspective of "know your company," companies that utilized privately subscribed bonds gave a high rating to the main banks. Conversely, the evaluation of the main banks was relatively poor among companies that utilized quick loans. It is understandable in the case of quick loans, because the element of "knowing your clients" is not as necessary as in relationship banking due to the loan scoring method. However, as approximately 80% was introduced by a main bank according to the previous response, the companies were dissatisfied, assuming that they were asked to utilize quick loans because the main bank did not know them well. Alternatively, it appears as though some companies have no other choice but to use the quick loans of other banks, as the main bank does not respond well enough.

On the other hand, it is interesting that companies utilizing privately subscribed bonds gave high ratings for main banks, stating that the bank knew them well. "Knowing your company" is the basis of relationship banking, and it is well known that this is essential in lending for small and medium-sized businesses. However, companies that have close relationships with banks implement non-loan fund procurement, utilizing the security market such as with privately subscribed bonds. The possibility that the relationship with clients may have

strengthened during the process of issuing privately subscribed bonds cannot be denied, judging from other questions in these questionnaires[7]. Therefore, this result reflects that there is almost no second market for privately subscribed bonds at present but they are rather simply a replacement for bank lending. This indicates that the issuance of privately subscribed bonds has a severe limitation as a financing method for small and medium-sized businesses which have financial difficulties.

With regard to the "stable provision of funds," companies that utilize privately subscribed bonds, the credit-guarantee system based on accounts receivable and CLO, gave high ratings for main banks. This seems an unexpected result, because companies that receive unstable fund provision from main banks must have a stronger need for such new financial products. However, it is possible that the evaluation of "stable provision of funds" has been improved due to such methods, by considering that new financial products are a part of the lending capabilities of main banks. According to the results of the questionnaire in this study, the latter situation may be the case.

Currently, excellent clients use them in the process of further improvement of fund procurement, and we are not in a situation in which they are helpful in fund procurement for companies that have difficulties borrowing funds from the main banks. On the other hand, companies that utilize quick loans give a poor rating for the main banks. It is believed that they feel as though they have no other choice but to use quick loans because of the unstable provision of funds from the main banks.

The last aspect of "provision of funds at a low interest rate" showed a similar tendency as has been mentioned above, and companies utilizing quick loans particularly gave poor ratings.

3. Conclusions

In this paper, the current status of diversification of fund procurement for small and medium-sized businesses in Japan was analyzed based on a survey questionnaire that the RIETI implemented in 2005. According to the questionnaire, many companies responded that the trigger for utilizing new financial products was an introduction by a main bank, and I was able to verify that fund procurement methods for small and medium-sized businesses have diversified as a result of efforts made by the policy-making authorities and financial institutions. However, the utilization rate of each financing method remained at a few percentage points, and diversification has not significantly improved small businesses as a whole and, particularly, those that have financing difficulties. Furthermore, although in relative terms, the quick loan method is being utilized by small businesses with poor performance, this does not mean that diversification has improved financial environments of small firms, but rather that it appears to be a result of the fact that strong

---

[7]Question 45 asked whether the main bank has strengthened "the depth of knowledge of your company itself" in the past 2 years. The rate at which "strengthened" was selected among companies utilizing privately subscribed bonds was 15.7%, while that among companies utilizing quick loans was 23.8%, which indicates that "strengthened" was not necessarily selected more frequently among companies utilizing privately subscribed bonds. In other words, it can be interpreted that the results of Table 5 indicate that privately subscribed bonds were issued to companies that were basically "well known."

relationship with banks has become unavailable to firms with poor performance.

On the other hand, companies that utilize privately subscribed bonds or CLO likely rate the main banks very highly from the viewpoint that they "know our company very well" and they think that they keep good relationships with the main banks. Although some improvement of the situation of small and medium-sized businesses with financing difficulties had been realized, current increased diversification has found mainly among excellent small and medium-sized businesses with strong financial conditions. Therefore, we are not at a stage in which the utilization of new financial products contributes considerably to improvement of the financial environment of small and medium-sized businesses as a whole. Furthermore, the results of the questionnaire indicate that privately subscribed bonds are simply a modified form of bank lending at the moment, and it can be said that we are not at a stage in which the issuance of privately subscribed bonds significantly contributes to the construction of a multi-stream financial system.

Finally, although the number of dealings with new financial products by private financial institutions has been steadily increasing, further efforts must be made to facilitate the diversification of fund procurement for small and medium-sized businesses amidst severe conditions.

REFERENCES

[1] Yamori, N., 2006, "Financial services that corporations want and challenges of financing for small and medium-sized businesses: Focusing on corporate awareness survey concerning corporate finance in the Kansai Region (in Japanese)," RIETI Discussion Paper 06–J–003.

[2] Yamori, N., 2007, "Japanese corporate finance: What factors affect the financial decisions of Japanese firms?" Alan N. Kendall, ed., International Business and Finance Issues, Nova Science Publishers, Inc., pp. 81–102.

[3] Uchida, H., Udell, G. F., and Yamori, N., "Loan officers and relationship lending," RIETI Discussion Paper Series 06–E–031, June 2006.

[4] Uchida, H., Udell, G. F., and Yamori, N., "How do Japanese banks discipline small and medium-sized borrowers?: An investigation of the deployment of lending technologies," International Finance Review (Institutional approach to global corporate governance), Vol. 9, pp. 57–80, 2008.

Scientific
Research

A Note on Secondary Buyouts-Creating Value or Recycling Capital

Jana KITZMANN[1], Dirk SCHIERECK[2]

[1]European Business School, Department of Finance, Accounting and Real Estate, Oestrich-Winkel, Germany; [2]Tech University Darmstadt, Department of Business Administrations, Economics and Law, Darmstadt, Germany.
Email: schiereck@bwl.tu-darmstadt.de

ABSTRACT

This paper analyzes whether secondary buyouts of private equity (PE) investors in general create value and therefore are a suitable alternative to exit strategies like trade sales and IPOs. Theoretically, two conflicting approaches might explain the use of secondary buyouts as an exit channel of private equity investors: the capital recycling effects and different potential sources of value creation. We present empirical tests of these approaches. The profitability of secondary buyouts is assessed by a comparison of exit multiples realized with secondary buyouts and trade sales. The results are not unequivocal, but overall we interpret our findings in a way that awards secondary buyouts a profitability that is not significantly different from trade sales. Therefore, we argue that secondary buyouts have the potential for adding value that arise from different sources like the reduction of agency costs or the functions of the financial investor. Secondary buyouts should thus not be seen as a second best alternative for recycling the PE investors' capital in situations where alternative—and supposedly more attractive—exit channels are unavailable.

Keywords: Management Buyouts, Performance, Private Equity, Secondary Buyout

1. Introduction

The profitability of a private equity (PE) investments critically depends on the realization of a successful exit strategy to terminate an investment [1,2] Throughout the 1990s and until the year 2000, PE investments were primarily exited through initial public offerings (IPOs) [3]. Later, this once highly attractive exit channel for PE investors was not longer attractive or available anymore. Because of the sluggish economy and the dwindling acquisition funds of corporations, the sale of equity shares to trade buyers also decreased. Thus, due to highly difficult market conditions, the traditional investment exit routes, IPOs and trade sales, turned out to be highly challenging [4].

Our study analyzes the return realized by *secondary buyouts*—the sale of a PE investor's equity stake in a portfolio company to another PE fund—as one possible exit alternative. The lack of alternative exit opportunities may be one reason why secondary buyouts as a means of rolling over PE investments were more and more taken into consideration over the last years [5–7]. But the bright future prospects which promise the secondary market for PE shareholdings to further grow significantly [8], underlines that it is not only the high liquidity of PE

investors, who are looking to invest their free funds, that drives the market [4]. One specific characteristic of a secondary sale is that only the equity of the financial investor is sold; the entrepreneur remains in possession of his equity share of the company [9]. Obviously, a secondary buyout will only be realized if the purchaser of the minority equity share can also expect the realization of an appropriate risk-adjusted return on this investment. Thus, as a prerequisite, the retiring financial investor must not have been able to use all available growth potentials of the company into a financial gain or the new investor has to possess specific supplementary know-how. From that perspective, a secondary buyout either can still generate value or is only used when other sale alternatives are not suitable for the portfolio company. It remains ex ante not clear why other financial investors should "see potential in a company their competitor is already trying to exit, unless their views are widely divergent?" [10].

Thus, a secondary buyout may only serve the PE investor to recycle his funds invested. However, an argument for wealth creation can be seen in that a secondary buyout is typically used during the expansion phase of a

portfolio company, because different financial investors engage in different lifecycle phases of a company. Thus, the acquiring investor may still be able to use his different expertise in order to maximize firm value. We study if secondary buyouts in general create wealth, where these potentials for value can be found, or whether secondary buyouts are merely used as a second best alternative in order to recycle capital in a situation where alternative exit channels are unavailable. Overall, the potential profitability of secondary buyouts will be derived by a comparison of secondary buyouts to alternative exit strategies, using data from worldwide exit transactions from 1999 until 2004.

PE investments can be defined as minority participations in the equity of companies that are not publicly listed at a stock market [11]. These kinds of investments have "a strict finite life with an expected duration of a few years [11][1] and the nature of the investment usually requires a high degree of direct involvement by the investor." [12,13] The overriding importance of an efficient exit strategy thereby results from the fact that most of the PE investor's return arises in the form of capital gains at the time of the exit, and not through cash flows during the investment period [8,9].

The paper is organized as follows: Chapter 2 gives an introduction to exit channels fro PE investments. After that, Chapter 3 focuses on secondary buyouts and how they impact exit values. Chapter 3 describes different sources of capital recycling effects caused by secondary buyouts. Opposing to that, Chapter 4 continues by introducing several wealth creating hypothesis from secondary buyouts. Chapter 5 represents the main part of this paper, with an empirical examination of secondary buyouts' profitability compared to alternative exit methods. Finally, Chapter 6 will give a brief summary as well as a short outlook.

2. Secondary Buyout Exits and Their Impact on Exit Values

The most important characteristic of a secondary buyout exit is that the equity share of the financial investor is passed on to a second PE investor. Thus, the wealth creating potential of a secondary buyout results from the value-adding services and the involvement of the new financial investor in the portfolio company. Since the company remains independent, both the entrepreneur's and the PE investor's efforts will contribute to the success of the venture [14,15].

The financial investor serves several functions that are for the benefit of the portfolio company's value. His value-adding services include "a combination of financial capital, monitoring and advisory services, and reputational capital", [6] and he makes significant contribu-

tions beyond the money provided to the portfolio companies [16,17]. The financial investor often brings in his expertise in financial engineering and helps in negotiating with potential corporate customers; he contributes his experience and contacts. Overall, an intensive involvement is crucial for portfolio companies since an increase of the firm value can be achieved through the investor's activities [18–20].

3. Capital Recycling By Secondary Buyout Exits

The term capital recycling should comprise all cases where a secondary buyout is either chosen by the retiring PE investor because an IPO or a trade sale is not available, or where a secondary buyout results in a situation where no real value is added by the acquiring financial investor, i.e. only wealth redistribution effects are used to increase exit value.

Secondary buyouts as second best alternatives for exit -The PE investor's value added will decline over the time of the investment, as the portfolio firm matures [21]. A hypothesis by Cumming and MacIntosh states that a PE investor will exit from an investment when the projected marginal value added—resulting from his investment efforts—is less than the projected marginal cost of these efforts. Thus, the exit will occur at a point in time when the financial investor's skill set is exhausted and the firm value cannot be further increased [22]. Deriving from this hypothesis, a central argument for secondary buyouts to be only a means of capital recycling is that a second financial investor can hardly be able to realize a positive return on investment since the former investor—now seeking the exit–has already exploited the means of increasing firm value [23]. Thus, secondary buyouts can be seen as a second best alternative in order to recycle the selling financial investor's capital and reinvest it in new ventures [6]. An indication can be seen in the fact that PE participations often trade at a discount in the market for secondary buyout companies.

Re-capitalizations (recaps) can be seen as a special form of a secondary buyout. It is used especially for companies with high free cash flows [24]. In a recap, the company is newly levered, based e.g. on the current EBITDA. The residual portion of the equity which exceeds the total firm value at that time flows back to the PE investor, so that he can pull out his invested equity [24]. Thus, a recap does not involve the sale to a new PE investor. Re-leveraging investments can be realized e.g. via securitisations, high yield debt offerings or sales and lease-backs [5]. A re-capitalization or re-leveraging of the portfolio company can be seen as a way of recycling capital, since no new "real" value is added by the transaction. The refinancing by the PE investor may be seen as a means of effecting exit for those portfolio companies for which other, more profitable, transactions were not

---

[1]PE investments usually have a duration of up to ten years.

available [8].

4. Value Creation Effects

The central question of this paper is, whether secondary buyouts are a qualified exit channel through which wealth can be created, so that the financial investor can realize a positive capital gain through the termination of his investment [20]. In this paper, it is assumed that wealth will be created if the buyer of the equity shares is able to add value to the portfolio company. If he is, he will be in the position of paying a price that enables the seller to realize a successful exit, that is comparable to an IPO or a trade sale exit. Value creation by the new owner can occur through several means. As already pointed out, it is the new owner's ability to resolve information asymmetries in valuing the firm, and his ability of monitoring the management and of resolving agency problems after the exit, that has an influence on the choice of exit. Since the acquiring PE investor may be superior in these abilities, a secondary buyout may turn out to be a favourable exit compared to other alternatives.

Reduction of agency costs—Corporate managers are the agent for their equity investors as the principals, a relationship that is usually impaired by *agency costs*. It is assumed that the managers, if left alone, will attempt to act in their own self-interest [25,26]. Such an agency conflict is also possible for portfolio companies financed with PE [27]. In a secondary buyout, the portfolio company's equity is acquired by a new financial investor, and it is this characteristic that is a potential source of value creation. The new PE investor is said to be better at overseeing and guiding the enterprise than a pool of small investors, thereby mitigating agency costs [28]. When compared to an IPO as exit channel, a secondary buyout has a clear advantage. In a public company, management's tendency to pursue its own interests is combined with the dispersed stockholders' inability to observe and precisely assess management's effort and its outcomes [29,30]. This conflict "creates a moral hazard and can lead to departures from the generally assumed principle of value maximization." [31] A secondary buyout exit, in contrast, results in ownership structures that are more efficient in resolving agency conflicts. They have the potential to increase firm value because the governance structure of the portfolio company after the exit provides stronger incentives for managers to operate firms efficiently [32][2].

---

[2]The argumentation for LBOs can be applied to secondary buyout companies, since both types are characterized by a limited number of outside investors.

[3]According to Kieschnick (1989), monitoring in the private secondary buyout company becomes even more effective due to the avoided free-rider problem in public corporations, where investments made in monitoring managerial actions benefit all shareholders so that any individual shareholders has too little incentive to invest in these monitoring activities.

Monitoring by the new owners-After a secondary buyout, the portfolio company is largely controlled by the new PE investor, which results in a potential superiority of secondary buyouts over e.g. IPOs from the monitoring aspect of reducing agency costs. Since the financial investor has a high equity interest in the corporation, he has a strong incentive to monitor the management so that he can ensure that it follows only the most profitable strategies that increase firm value. Because of his substantial equity position, he is more interested in monitoring the managers' actions [33,34] than dispersed stockholders in a public firm are [35][3]. Additionally, due to his expertise in the field of PE investments, he has a comparative advantage in monitoring the management. As a consequence of this diligent monitoring, the private secondary buyout company is likely to show better performance due to superior management decisions and commitment [36–40].

Reduction of information asymmetries - The problem of information asymmetries acknowledges that information is distributed unevenly in the business environment; management and investors have different information about a corporation's situation and prospects [41]. Information asymmetries are particularly developed in public companies where management has a better idea of the actual and potential performance of the firm than outside—and often uninformed—investors. In contrast to this, a secondary buyout company is controlled by the financial investor as an inside investor who not only provides capital but who also is actively involved in the portfolio company and is very well informed about the firm's future and investment opportunities. PE investors have a strongly developed experience in the concerned industries and are better able to reduce information asymmetries than most stockowners in a public company, thereby reducing one component of the agency cost [38,42–44].

Reduction of the overinvestment problem and finance staging—The overinvestment problem results from the limited liability of the entrepreneur. Since he uses the money of investors in his venture, a situation may arise where the entrepreneur wants to continue to invest in a project even when it has a negative net present value. As long as there is some probability that the investment will turn out to be successful, the entrepreneur will invest as long as he will be provided with capital [41]. This overinvestment problem can be solved if the financing comes from the PE investor after the secondary buyout. He plays the role of an inside investor and observes private information about the project's profitability [38]. For solving the overinvestment problem, the financial investor uses a *staged capital commitment*, i.e. he commits only a fraction of the capital needed for achieving certain milestones. Subsequent financing is dependent on the successful completion of these intermediate objectives

[38,39,45].

Transaction synergies, economies of scope and learning curve effects—In general, PE investors have lower than average information-gathering costs. They can benefit from *transaction synergies*, i.e. they gather the information about potential portfolio companies on behalf of a number of investors in PE funds [43,44,46]. They can use *economies of scope* because they invest in different ventures, thereby being able to creating a network of referral sources, service providers like attorneys and accountants and industry contacts. Also, a PE investor profits from *learning curve effects* in that he can use information produced for one proposal for subsequent proposals, thereby dramatically reducing information-gathering costs. These three sources of savings are especially developed for investors who specialize in certain industries [28,47]. Because of the transaction synergies, the economies of scope and the learning curve effects, a PE investor acquiring a portfolio company in a secondary buyout might be able to bid higher prices, thereby increasing the rate of return that the retiring financial investor can realize. Additionally, the new PE investor may be able to use his expertise and information base in order to guide this knowledge into value-adding actions for the portfolio company.

5. Data and Methodology

The following empirical examination aims at testing the introduced theories—capital recycling and value creating theories for secondary buyouts—for practical relevance.

The purpose of the following empirical examination is the analysis of the profitability of different exit transactions. Profitability in this context is defined as the value that the retiring PE investor can realize with the exit transaction. Thus a higher exit profitability will lead to a higher realized return on investment for the financial investor and is therefore in accordance with the object of exit value maximization. The comparison of different exit channels' profitability will be used in order to determine whether secondary buyouts are value creating or not. It is assumed that secondary buyouts can create wealth if they provide comparable profitability to those exit strategies which are widely described to be superior exit forms in the literature. If the profitability of secondary buyouts turns out to be inferior to alternative exit strategies, they are assumed to be merely a means of recycling capital. This would also indicate that secondary buyouts are chosen primarily for lower quality portfolio companies.

In order to be able to compare alternative exit strategies with each other and in order to draw a conclusion whether secondary buyouts are competitive exits compared to IPOs, trade sales, or MBOs, a ranking order hypothesis for the four different exit strategies is developed in a first step. IPO exits appear to be the most desirable exit for high quality, rapid growth firms for which PE investors derive most of their returns. The listing at a public market is unique in its ability to supply capital on a large scale and to spread risk. IPOs have the potential to maximally enhance the PE investor's reputation and that of the portfolio company. The IPO may result in a public profile for the firm that facilitates the sale of its products or services and that assists in future capital raising efforts. As a result, the IPO exit is hypothesized to be the most desirable form of exit. Of course, only portfolio firms that surmount specific hurdle sizes and growth rates are suitable candidates for IPOs, so that other exit alternatives must be taken into consideration for companies that are not saleable in the public market [22].

Trade sales are considered to be the second best form of exit, following IPOs. One of the great advantages is the ability of the acquirer to generate transaction synergies [23]. A trade sale will be especially attractive when the firm's technology is highly complementary to technologies possessed by the strategic acquirers. Additionally, the strategic acquirer in a trade sale will be uniquely well positioned to monitor and discipline the management of the portfolio company [34,48].

MBOs are hypothesized to be an inferior form of exit. MBOs result in no transaction synergies and are typically used for investments with no potential for alternative exits. Even though post-exit managerial incentives will be high, managers may tend to prefer non-financial objectives, in particular indulging a leisure preference. Additionally, since the entrepreneurs usually do not have the funds to affect a MBO, they will need large amounts of debt financing, which in turn will only be available at great cost [6].

Secondary buyouts, finally, are hypothesized to be on average inferior to an IPO and a trade sale, but superior to a MBO. Since the acquirer will purchase less than 100 percent of the portfolio company's equity in a secondary buyout, the incentive and ability to monitor post-exit will be less than in connection with a trade sale, thereby lowering the value of the purchase to the acquirer. Secondary buyouts may be slightly inferior with respect to trade sales in reducing agency costs. Since the portfolio company is usually integrated into another company after a trade sale, and since the entire equity stake is acquired by the buyer, the new owners have the total monitoring and decision control over the operations, which often results in a high control premium [24].

However, secondary buyouts have more capacity than MBOs to generate high returns, since a high valuing acquirer may be content to purchase the equity shares of the retiring PE investor, and may be willing to pay a high multiple to effect this purchase. Additionally, a secondary buyout will bring on board a new actively involved monitor with specialized knowledge [49]. The different functions and tasks of the new financial investor will

often enhance the value of the firm and will hence increase the buyer's willingness to pay [22]. Overall, the alternative exit channels can be ranked in the order IPO, trade sale, secondary buyout, and MBO. If this ranking order is confirmed in the following examination of exit multiples, this could be assessed as an indication for secondary buyouts to be merely a means of capital recycling. If this ranking order cannot be confirmed, the empirical examination would be supporting the value adding potential of secondary buyouts.

Examination of exit multiples realized with secondary buyouts—The data comprises four different exit multiples—turnover multiples, EBITDA multiples, EBIT multiples and earnings multiples – from secondary buyout, trade sale and MBO transactions worldwide, recorded by www.mergermarket.com. The examined exit transactions took place during a time period from 1999 to 2004. In total, the data set comprises 63 secondary buyout transactions, 36 MBO transactions, and 96 trade sales [50] In addition to the exit multiples for each registered transaction, the available data includes the industry group for the respective portfolio companies as well as the turnover, EBITDA, EBIT and earnings of the portfolio companies at the time of the exit transactions. This allows a categorization of the total data set into industry groups and into groups of companies with comparable levels of turnover and EBITDA levels respectively.

In order to make a statement about the profitability of secondary buyouts, four different exit multiples for secondary buyout transactions will be compared to exit multiples for MBOs and trade sales. Since IPOs are assumed to be a superior form of exit [6,49,51], and since exit multiples for private exit transactions are hardly comparable to exit multiples realized with IPOs [39][4], the paper confines itself to a comparison between the three mentioned methods of private exits.

In a first step, the total data set will be analyzed; thus, no grouping or clustering will be used. The mean multiples for the three exit channels will be compared for every multiple category so that there will be four different examinations.

In a second step, this procedure will be applied to the data set categorized into the different *industry groups*, in order to gain more specific insights and results. In a third step, the clusters of companies with *comparable levels of turnover and EBITDA* will be analyzed. In order to examine whether there is sufficient evidence to conclude that the multiples for different exit strategies differ significantly from each other, an *analysis of variance* (ANOVA) will be conducted following the described comparisons. Finally, the results of the empirical examination will be tested for their accordance with the hypothesized ranking order so that a statement about secondary buyouts' value creation potential can be made as

[4]Multiples of PE backed IPOs could e.g. be influenced by reputation and signalling effects rather than performance per se.

a central conclusion.

In the empirical examination, exit multiples are used in order to determine the profitability of secondary buyouts in comparison to trade sales and management buyouts. It is assumed that in a case where secondary buyouts show comparable or significantly higher exit multiples, this kind of exit strategy has the potential for wealth creation and is not only used as a second best alternative to recycle capital. The examined data comprises four different exit multiples: turnover multiples, EBITDA multiples, EBIT multiples and earnings multiples. The exit multiples can be calculated by dividing the value of the exit transaction by the turnover, EBITDA, EBIT or earnings of the portfolio company at the time of the exit. Turnover, EBITDA, EBIT or earnings serve as *reference values* for the respective multiple. In a case where several portfolio companies have comparable reference values, the higher the exit multiple, the higher will be the value of the exit transaction. Therefore, a high exit multiple would indicate that the buyer of the equity shares attaches value to the portfolio company and was ready to pay a relatively high price, making the respective exit a value adding one. As a result, it is possible to compare exit multiples from different exit methods in order to determine which exit strategy is the most profitable one, resulting in the highest values of exit transaction.

A prerequisite for this procedure of examination comes from the use of multiples. Multiples should only be used among comparable companies, e.g. within *peer groups*, in order to ensure that these firms being used for comparison have similar characteristics [52]. For the examination in this paper, the companies whose exit multiples are being compared should have similar turnover, EBITDA, EBIT and earnings levels respectively. Otherwise, similar exit multiples for different exit strategies could arise even if the transaction values are highly different from each other. Thus, higher multiples would not unequivocally indicate higher transaction values and thus higher profitability. To sum up, a multiple's value can be driven by two influence factors, the reference value and the transaction value. In order to be able to compare the different exit channels' profitability unrestrictedly, the first influence factor has to be eliminated.

When determining the members of a peer group, i.e. the comparable firms, one should control for all the variables that can influence the multiple, in theory [52]. Because of the limited data set available for the empirical examination, the total data set is analyzed in a first step, even though it is assumed that the results are potentially biased because of possible differences in the reference values.

In a second step, peer groups are formed by dividing the transactions data into industry groups according to the industry sector of the portfolio companies. By this, at least a minimum of comparability between the exit multiples is achieved. However, when analyzing the results

iB

of this examination, one should always keep in mind that they could be biased. The fact that no two companies are exactly similar in terms of reference values, the definition of "comparable" firms is somewhat "vague" [52].

In a third step, several clusters are formed from the total data set based on the reference values. In other words, within one cluster, the analyzed exit multiples come from portfolio companies with similar turnover and EBITDA levels, respectively. By this *clustering*, the problem of missing clearness can be eliminated by some extent, i.e. a high multiple would necessarily indicate higher transaction values. This clustered analysis was only conducted for the turnover and EBITDA multiples since the clustered data sets for the remaining two multiples were too small in order to conduct a meaningful ANOVA.

In order to draw a conclusion about the competitiveness of secondary buyouts, one has to analyze the transaction multiples from each of the three exit categories. The objective thereby is to compare the means. With an analysis of variance (ANOVA) it is possible to determine whether there is sufficient evidence to conclude that the transaction multiples from at least one exit strategy on average differ significantly from at least one other exit strategy's multiples. Since we have three data sets—multiples for secondary buyouts, MBOs and trade sales-, ANOVA is used, because this method uses the F distribution and can test the significance among multiple data sets simultaneously. Since we have one factor as the independent variable—namely the multiple—and three levels—secondary buyouts, MBOs and trade sales—we can use a single factor ANOVA. When a single factor is employed, the levels of the factor are called the *treatments* of the examination. Since ANOVA assumes that all of the data sets are normally distributed, the normality of the different data sets was assessed in a first step. This was done by computing the intervals $\bar{x} \pm s$, $\bar{x} \pm 2s$, and $\bar{x} \pm 3s$ -where $\bar{x}$ stands for the mean and s stands for the standard deviation-and by determining the percentages of measurements falling in each. This method is based on the properties of a normal distribution. Since the percentages were tested to be approximately equal to 68%, 95%, and 100% respectively, a normal distribution of the analyzed data sets could be assumed. ANOVA can

therefore be applied. For all ANOVA examinations in this paper, a significance level of 0,05 is used. In the ANOVA examinations, two hypotheses are tested:

H_0: There is no significant difference among the exit multiples for the three exit categories.

H_1: At least one of the exit channels shows multiples significantly different from at least on other exit channel's multiples.

6. Results

As ANOVA only shows that a difference exists between at least two treatments, this leads to the question which of the means differ. A graphical representation is used to answer this question. First, the 95% confidence intervals on the mean are calculated for each data series. With the confidence intervals, upper and lower bounds according to the chosen confidence level are determined. The results are shown in an effects plot. The graphical representation emphasizes which means show a significant difference. If the intervals of two exit channels do not overlap, this means that their multiple means differ statistically significantly.

In Table 1, an overview is given of the results of the empirical analyses of exit multiples. First, the examination of the *total data* set will be analyzed. Important to notice is, that secondary buyouts show the highest mean for turnover, EBIT and earnings multiples. Only for EBITDA multiples trade sales have a slightly higher mean. However, the differences found for the EBITDA multiples are statistically not significant for a significance level of $\alpha = 0,05$. In contrast, the F values for turnover, EBIT and earnings multiples were higher than the F crit. values, so that the null hypotheses had to be rejected in these cases. Thus, significant differences were found where multiple means were highest for secondary buyouts. The effects plots drawn for the four multiple categories show that there is a significant difference between secondary buyouts and MBOs for the turnover multiples and for the EBIT multiples; and there is a significant difference between secondary buyouts and MBOs and between secondary buyouts and trade sales for the earnings multiples. In the mentioned cases, the bars-representing the confidence intervals added on and

Table 1. Results summary of the empirical analysis of exit multiples—Total data set

Data Set	Mean			Significance	
	Secondaries	MBOs	Trade Sales	Significance level of 0,05.	
Total					
Turnover Multiples	**1,378**	0,853	1,110	**Yes***	SB - MBO
EBITDA Multiples	10,712	7,600	**10,996**	No	
EBIT Multiples	**15,008**	9,753	13,741	**Yes**	SB - MBO
Earnings Multiples	**30,828**	11,763	18,142	**Yes***	SB - MBO; SB - TS

* Significant differences also for a significance level of 0,01.

Table 2. Summary of results of the empirical analysis of exit multiples—Industry Groups

Data Set	Mean			Significance
	Secondaries	MBOs	Trade Sales	Significance level of 0,05
Industry Groups				
Consumer Retail				
Turnover Multiples	**1,586**	0,920	0,800	No
EBIT Multiples	**20,340**	12,933	11,100	No
Manufacturing, Industrial Products				
Turnover Multiples	1,027	0,740	**1,233**	No
EBITDA Multiples	**9,320**	7,200	7,300	No
Construction				
Turnover Multiples	1,100	0,740	**1,550**	No
Medical				
Turnover Multiples	**2,933**	1,633	1,375	No
EBITDA Multiples	**16,075**	7,850	10,750	No
EBIT Multiples	**26,125**	10,700	25,733	No
Earnings Multiples	**58,225**	13,233	39,625	No
Leisure				
Turnover Multiples	**3,050**	0,733	1,375	**Yes** SB - MBO
EBITDA Multiples	11,000	4,850	**16,333**	No
EBIT Multiples	14,200	6,433	**23,740**	No
Services				
Turnover Multiples	**2,050**	1,525	1,443	No
EBITDA Multiples	**15,000**	8,467	14,208	No
EBIT Multiples	**24,600**	13,650	18,107	No

Table 3. Summary of results of the empirical analysis of exit multiples–Clusters

Data Set	Mean			Significance
	Secondaries	MBOs	Trade Sales	Significance level of 0,05
Turnover Multiples				
0 to 20 mio. £	**3,2667**	1,6125	1,7640	No
to 40 mio. £	**1,6167**	0,9200	1,4944	No
to 60 mio. £	**1,8000**	0,5333	0,8357	No
to 80 mio. £	**2,1857**	0,9500	0,7111	No
to 100 mio. £	**2,3000**	1,2667	0,4500	No
EBITDA Multiples				
0 to 5 mio. £	12,6500	7,5000	**13,0643**	No
to 10 mio £	**12,5000**	10,7500	8,4833	No
20 to 40 mio. £	8,2636	6,6000	**9,5400**	No

subtracted from the means—do not overlap.

Concerning the examination of the data divided up into *industry groups*, results are similar to the total data set results in that secondary buyouts in most cases showed the highest multiple means. Trade sales had the highest means in only four of the 15 data sets analyzed; means for MBOs were always lower than those of at least one alternative exit channel. However, when analyzing the exit multiples for the industry groups, onlyfor the turnover multiples in the leisure category can be

iB

found a statistical significance between secondary buyouts and management buyouts. For all other data sets, the ANOVA showed no significant differences in the multiple means. The results for this second examination method are summarized in Table 2.

Concerning the clustered ANOVA analyses summarized in Table 3, again secondary buyouts showed the highest means in most of the analyzed cases, but no significant differences between secondary buyouts, trade sales and MBOs could be found at all.

To sum up the results, it can be said that in all three different methods of examination, secondary buyouts showed the highest multiple means in most cases. In all other cases, trade sales had the highest average means. Management buyouts were always inferior to at least on other category of exit strategy. However, significant differences between the multiple means for secondary buyouts, trade sales and MBOs were only found in the examination of the total data set and for the industry group *leisure* for the turnover multiple. Concerning the total data set analysis, significant differences were proven between secondary buyouts and MBOs in three cases (turnover multiples, EBITDA multiples, and earnings multiples) and between secondary buyouts and trade sales in one case (earnings multiples). Relating to the established ranking order hypothesis, these results are more or less in harmony with the hypothesized superiority of secondary buyouts over MBOs. In contrast, the results for the relationship between secondary buyouts and trade sales seem to deviate from the ranking order. While the theory predicts trade sales to be a more profitable exit channel than secondary buyouts, the results of the empirical examination indicate that secondary buyouts are at least of equal quality, showing comparable or even higher exit multiples than trade sales. However, these first results are probably of little explanatory power because the total data set may not exclusively comprise comparable companies with similar reference values. Therefore, the results should be seen with caution and one should always keep in mind the assumptions that have been made.

In contrast, the industry group examination and the analyses of the clustered groups should yield more reliable results. For both methods no significance—with exception of the industry group *leisure* for the turnover multiple–between the exit multiples of the three exit channels was discovered. This result is also in contradiction with the hypothesized ranking order. Secondary buyouts not only seem to be a competitive exit strategy— as opposed to trade sales—but also provide multiples that are as high or even higher as the supposedly superior trade sale exits. This would indicate that secondary buyouts result in transaction values that are comparable to those of trade sales. While MBOs had consistently lower multiple means than secondary buyouts and/or trade sales,

no statistical significance could be found here either. Thus, according to these results, management buyouts could not be ranked as an inferior exit strategy per se.

When considering the results of the two last examination methods, some restrictions have to be made here as well. It is true that the grouping into industry sectors and the clustering resulted in a comparison of a more peer group like data set, but simultaneously reduced the data sets dramatically. In some cases, there were only three multiples left for one exit category. Thus, the small data sets could turn out as a problem by introducing biases because they are not representative anymore [53].

The discussed empirical examination is suitable to make concrete statements about the profitability of secondary buyouts. All three different methods applied validate the discovery that secondary buyouts are a competitive exit strategy compared to trade sales, and are therefore not only a second best alternative that is chosen when alternative exits are not available in order to recycle capital. However, potential biases in the examinations cause the existence of two different potential explanations for these results. One possibility is that secondary buyouts indeed have value adding potentials that could arise from the various sources already described in detail. Due to the new financial investor's functions and the value creation effects, secondary buyout candidates are highly valued by the potential buyer, which in turn drives up the transaction values realized in secondary buyouts. As a result, secondary buyout transaction values are comparable to those realized with trade sales, the exit method that is widely regarded as a superior form of exit in the literature [49]. Since therefore the retiring PE investor would be able to realize the same return on investment with these two exit alternatives, he should be indifferent when choosing between these two—in a case where portfolio characteristics would allow the availability of both. Also, the results indicate that secondary buyouts are a superior form of exit compared to management buyouts. This result is in accordance with the literature, describing MBOs as a suitable exit form mainly for moderately successful companies with stable cash flows [24].

Another possibility is that the empirical results were driven by potential biases in the data sets. Distortions can especially be suspected for the results of the examinations of the total data set because of the limited comparability of the exit multiples due to unequal reference values. Additionally, problems may arise due to the varying amount of data for the three different exit channels compared in the examinations. The amount of transaction data for MBOs was much smaller than for secondary buyouts and trade sales, which may cause biases due to the restricted representativeness. Additionally, the classification of the exit multiples according to the portfolio companies' industry groups and reference values resulted in relatively small data sets, causing similar potential

biases and therefore limited explanatory power. Thus, it remains possible that secondary buyouts indeed do not have value creation potential and are therefore only a means of recycling capital. However, the fact that in no case was found a significant superiority of trade sales over secondary buyouts could be seen as a hint that reality shows a tendency towards the first explanation. Overall, all the findings should be considered as merely hypothetical.

7. Conclusions and Outlook

The central subject of examination in this paper has been the question whether secondary buyouts in general create value and are therefore a suitable alternative to exit strategies like trade sales and MBOs. In a theoretical part, two main hypotheses were discussed: the capital recycling effects of secondary buyouts and different potential sources of value creation. These theories were tested in an empirical examination.

It was assumed that secondary buyouts would be potentially value adding exit channels, if their profitability would be comparable to that of trade sales. The profitability was therefore assessed by a comparison of exit multiples realized with secondary buyouts and trade sales. The results were not unequivocal, but could be interpreted in a way that awards secondary buyouts a profitability that is not significantly different from trade sales as an exit strategy that is often assumed to be superior with respect to the realized returns on investment. Therefore, it could be argued that secondary buyouts have the potential for adding value that could arise from different sources like the reduction of agency costs, the functions of the financial investor etc. as introduced in the theoretical part of this paper. Secondary buyouts should thus not be seen as a second best alternative for recycling the PE investors' capital in situations where alternative—and supposedly more attractive—exit channels are unavailable. This argumentation can be underpinned with the expected development of the market for secondary buyouts. The worldwide market for secondary buyout portfolio companies currently comprises a volume of about USD 10 bill. Experts estimate that this volume will increase to over 25 percent of all effected transactions worldwide in the coming years, and will thus play a highly important role in the financial markets. Additionally, completed transactions in secondary buyouts which have been highly successful, can be seen as prove that secondary buyouts are not merely a temporary solution.

REFERENCES

[1] Thies, A., "Steuerliche optimierung bei buy outs durch pe-gesellschaften," in M&A Review, pp. 479–484, November 2003.

[2] Bance, A., "Special paper–why and how to invest in PE, an EVCA investor relations committee paper," European PE & Venture Capital Association, 2002, http://www.evca.com.

[3] Bascha, A. and Walz, U., "Financing practices in the German venture capital industry: an empirical assessment," Working Paper, University of Tübingen, February 2001, available on the Social Science Research Network (SSRN) Electronic Library at: http://ssrn.com/abstract=267572.

[4] Köhler, P., and SCHÖNAUER, F., "Bankfusionen lassen Angebot am Beteiligungsmarkt anschwellen, Überschneidungen in den Firmenportfolios—Kreditinstitute geben Private-Equity-Geschäft ab," in Handelsblatt, No. 17, pp. 26, January 26, 2004.

[5] Sonter, D. and Fountain, K., "Alternative exits," in Venture Economics, w/o y., available at: http://www.ventureeconomics.com.

[6] Black, B. S. and Gilson, R. J., "Venture capital and the structure of capital markets: Bank versus stock markets," in Journal of Financial Economics, No. 47, pp. 243–277, 1997.

[7] Sormani, A., "Exits: The IPO alternative," in EVJC, pp. 20–23, November 2002, http://www.evjc.com.

[8] Wright, M., Chiplin, B., Thompson, S., and Robbie, K., "Realisation of venture capital investments: The case of management buyouts in the UK," in The Service Industries Journal, No. 10, 499–520, 1990.

[9] Cumming, D. J. and Fleming, G., "A law and finance analysis of venture capital exits in emerging markets," Working Paper, University of Alberta and Australian National University, February 2003.

[10] Laine, M. and Torstila, S. "The exit rates of liquidated venture capital funds," May 2003, available on the Social Science Research Network (SSRN) Electronic Library at: http://ssrn.com/abstract=417641.

[11] Timmreck, C., "Beteiligungsgeschäft in Deutschland," in M&A Review, No. 5, pp. 225–229, 2003.

[12] Wang, C. K. and Sim, V. Y. L., "Exit strategies of venture capital-backed companies in Singapore," in Venture Capital, No. 3, pp. 337–358, 2001.

[13] Sahlmann, W. A., "The structure and governance of venture-capital organizations," in Journal of Financial Economics, No. 27, pp. 473–521, 1990.

[14] Hellmann, T., "IPOs, acquisitions and the use of convertible securities in venture capital," Working Paper, Graduate School of Business, Stanford University, December 2000, available on the Social Science Research Network (SSRN) Electronic Library at: http://ssrn.com/abstract=257608.

[15] Houben, E., "Venture capital, double-sided adverse selection, and double-sided moral hazard," December 2002, available on the Social Science Research Network (SSRN) Electronic Library at: http://ssrn.com/abstract=365841.

[16] Sapienza, H. J., 'When do venture capitalists add value?" in Journal of Business Venturing, No. 7, pp. 9–27, 1992.

[17] Cumming, D., Franzke, S., and Walz, U., "Kann es eine

Venture Capital-Industrie ohne Börsengänge geben? Notwendigkeit der Finanzierung von Innovationen—IPO ist attraktivste Exit-Lösung—Pool erfahrener VCs steht für den nächsten Aufschwung zur Verfügung," in Börsen-Zeitung, No. 10, pp. 19, January 16, 2004,.

[18] Sapienza, H., Manigart, S., and Vermeir, W., "Venture capitalist governance and value-added in four countries," in Journal of Business Venturing, No. 11, pp. 439–469, 1996.

[19] Schwienbacher, A., "An empirical analysis of venture capital exits in Europe and in the United States," Berlin Meeting Discussion Paper, July 2002, available on the Social Science Research Network (SSRN) Electronic Library at: http://ssrn.com/abstract=302001.

[20] Prester, M., "Exit-strategien deutscher venture capital Gesellschaften," Betriebswirtschaftliche Schriftenreihe Band 87, Münster, 2002.

[21] Smith, G. D., "Control over exit in venture capital relationships," 2001, available on the Social Science Research Network (SSRN) Electronic Library at: http://ssrn.com/abstract=272231.

[22] Cumming, D. J. and Macintosh, J. G., "Venture capital exits in Canada and the United States," in University of Toronto Law Journal, Vol. 53, pp. 101–200, 2003, available on the Social Science Research Network (SSRN) Electronic Library at: http://ssrn.com/abstract=321641.

[23] Wall, J. and Smith, J., "Better exits, results of a survey of the venture capital exit market and guidance on how venture capitalists can improve exit performance," Price Waterhouse Corporate Finance, 1998.

[24] Achleitner, A., "Handbuch investment banking," 2 Editon, Wiesbaden, 2000.

[25] Jensen, M. C., "Agency costs of free cash flow, corporate finance, and takeovers," in American Economic Review, No. 76, pp. 323 – 329, 1986.

[26] Ross, S. A., Westerfield, R. W., and Jaffe, J. F., "Corporate finance," 6 Editon, New York, 2002.

[27] Barney, J., Busenitz, L., Fiet, J., and Moesel, D., "The structure of venture capital governance: An organisational economic analysis of relations between venture capital firms and new ventures," in Academy of Management Proceedings, pp. 64–68, 1989.

[28] Fried, V. H. and Hisrich, R. D., "Towards a model of venture capital investment decision making," in Financial Management, pp. 28–37, Autumn 1994.

[29] Jensen, M. C., "Eclipse of the public corporation," in Harvard Business Review, No. 67, pp. 5, 1989.

[30] Lehn, K. and Poulsen, A., "Free cash flow and stockholder gains in going private transactions," in: The Journal of Finance, No. 44, pp. 771–787, 1989.

[31] Frankfurter, G. M. and Gunay, E., "Management buyouts: The sources and sharing of wealth between insiders and outside shareholders," in Quarterly Review of Economics and Finance, No. 32, pp. 82–95, 1992.

[32] Liebeskind, J., Wiersema, M., and Hansen, G., "LBOs, corporate restructuring, and the incentive-intensity hy-

pothesis," in Financial Management, Vol. 21, No. 1, pp. 73–88, Spring 1992.

[33] Brous, P. A. and Kini, O., "The valuation effects of equity issues and the level of institutional ownership: Evidence from analysts' earnings forecasts," in Financial Management, pp. 33–46, Spring 1994.

[34] Chen, S. S., Ho, K. W., Lee, C. F., and Yeo, G. H. H., "Wealth effects of pe placements: Evidence from singapore," in The Financial Review, No. 37, 2002, pp. 165–184.

[35] Kieschnick, R. L., "Management buyouts of public corporations: An analysis of prior characteristics," in AMIHUD, Yakov: Leveraged management buyouts—Causes and consequences, New York, pp. 35–67, 1989.

[36] Deangelo, H. and Deangelo, L., "Management buyouts of publicly traded corporations," in Financial Analysts Journal, pp. 38–49, May–June 1987.

[37] Deangelo, H., Deangelo, L., and Rice, E. Y., "Going private: Minority freezeouts and stockholder wealth," in Journal of Law & Economics, No. 27, pp. 367–402, 1984a.

[38] Admati, A. R. and Pfleiderer, P., "Robust financial contracting and the role of venture capitalists," in: The Journal of Finance, No. 49, pp. 371–402, 1994.

[39] Barry, C. B., "New directions in research on venture capital finance," in Financial Management, No. 23, pp. 3–15, 1994.

[40] Opler, T. and Titman, S., "The determinants of leveraged buyout activity: Free cash flow vs. financial distress costs," in: The Journal of Finance, No. 48, 1985–1999.

[41] Martola, J. M., "Developing a framework for venture capitalists to increase the value of investments," Working Paper, Helsinki University of Technology, February 2002.

[42] Gompers, P. A., "Optimal investment, monitoring, and the staging of venture capital," in Journal of Finance, No. 50, pp. 1461–1489, 1995.

[43] Triantis, George G., "Financial contract design in the world of venture capital," 2001, available on the Social Science Research Network (SSRN) Electronic Library at: http://papers.ssrn.com/paper.taf?abstract_id=258392.

[44] Dushnitsky, G. and Lenox, M. J., "When does corporate venture capital investment create firm value?" Working Paper, September 2003.

[45] Siemes, M., "Going private unter Beteiligung von Finanzinvestoren in Deutschland," Dissertation European Business School Oestrich-Winkel, Auflage, Wiesbaden, 2003.

[46] Chan, Y., "On the positive role of financial intermediation in allocations of venture capital in a market with imperfect information," in Journal of Finance, No. 38, pp. 1543–1561, 1983.

[47] Norton, E. and Tenenbaum, B. H., "Specialization versus diversification as a venture capital investment strategy," in Journal of Business Venturing, No. 8, pp. 431–442, 1993.

iB

[48] Cumming, D. J. and Macintosh, J. G., "The extent of venture capital exits: Evidence from Canada and the United States," University of Toronto Law and Economics Research Paper, No. 01–03, October 2000, available on the Social Science Research Network (SSRN) Electronic Library at: http://ssrn.com/abstract=250519.

[49] Cumming, D. J. and Macintosh, J. G., "A cross-country comparison of full and partial capital exits," in University of Toronto Law and Economics Research Paper No. 01–04, May 2002, available on the Social Science Research Network (SSRN) Electronic Library at: http://ssrn.com/abstract=268557.

[50] w/o a, Deals: Result Summaries, available at: http://www.mergermarket.com/deals/dealSummary.asp?npid=500&contextid =4788927, date of retrieve: 26.01.2004.

[51] Cole, S., "Exits and entrances," in CA Magazine, Vol. 130, No. 7, pp. 41–42, 1997.

[52] Damodaran, A., "Investment valuation—tools and techniques for determining the value of any asset," New York, 1996.

[53] Mcclave, J. T., Benson, P. G., and Sincich, T., "Statistics for business and economics," 8 Edition, New York, 2001.

◆❖ Scientific
❖◆ Research

Study of Green Behavior with a Focus on Mexican Individuals

Eva CONRAUD-KOELLNER[1], Luis Arturo RIVAS-TOVAR[2]

[1]Universidad de Guanajuato, Guanajuato, Mexico; [2]Instituto Politecnico Nacional, Mexico City, Mexico.
Email: evac@quijote.ugto.mx, larivas33@hotmail.com

ABSTRACT

Studies of green behavior have awakened a growing interest, since the accelerated environmental degradation is partly attributed to a lack of such behavior. However it is not an easy task to modify consumer habits and influence the behavior of individuals. This paper summarizes variables extracted from four models of green behavior in five core groups: orientation man-nature, perceived control, ecological knowledge, personal consequences and environmental consequences, with a special focus on the Mexican individuals as an example of application of environmental marketing strategies. Analysis shows that in spite of efforts done in environmental management in Mexico, one of the 12 environmentally megadiverse country of the world, there is a big gap between official programs to protect environment and social participation. In the particular case of Mexico, it is recommended to relate environmental issues to survival process, in order to apply efficient environmental marketing strategies.

1. Introduction

Studies of green behavior have awakened a growing interest, since the accelerated environmental degradation is partly attributed to a lack of such behavior. However it is not an easy task to modify consumer habits and influence the behavior of individuals. The review of literature regarding green behavior reveals preference for variables related to values or attitudes (or changes in attitudes) or external variables of the individuals, mainly the situational, that occupy a prominent position in these studies [1–8]. Studies focus especially on social standards and the physical environment where specific behaviors occur, observing how these are fomented or obstructed by the conditions of the environment [9,10]. This paper summarizes variables extracted from four models of green behavior in five core groups: orientation man-nature, perceived control, ecological knowledge, personal consequences and environmental consequences, which are described below, with a special focus on the Mexican individuals as an example of application of environmental marketing strategies.

2. Models of Green Behavior

The four models found in recent specialized literature are briefly described here below:

The conceptual model of the behavior of ecological purchase, developed in 2004 by Chan and Lau [1] suggests that the values that traditionally govern the relation of the individuals with nature (orientation man-nature) influence directly the amount of knowledge that a person has about environmental topics (ecological knowledge), and the level of emotional attachment and commitment that an individual shows towards the environment (ecological feeling). And they conclude that purchase behavior depends on the individual's purchase intention of green products. This model can be described in the following Figure 1.

The model of the environmentally responsible purchase developed by Follows and Jobber [2], predicts the purchase of a specific type of environmentally responsible product based on three variables: self importance (or pro-social) values, conservation values, and individualistic values. These refer to the factors that motivate people to put their own interests over any other thing. Follows and Jobber propose that the goals of an individual person's self-indulgence and self reward are consistent with the interest in the individual consequences, and act in a negative way on the environment. The model is presented in the following Figure 2.

The model of the Theory of Behavior developed by Ajzen adapted to the study of the consumer purchase
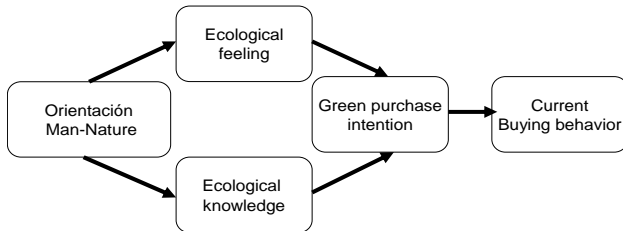
iB

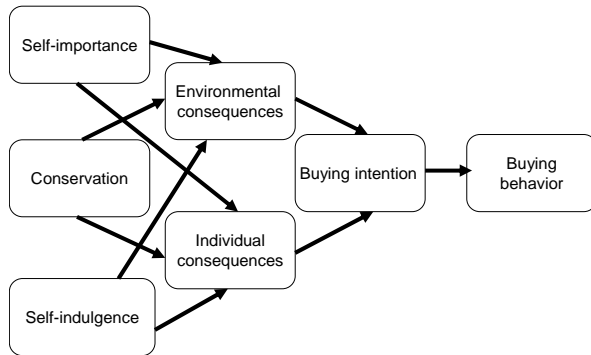**Figure 1. Conceptual model of the behavior of ecological purchase**

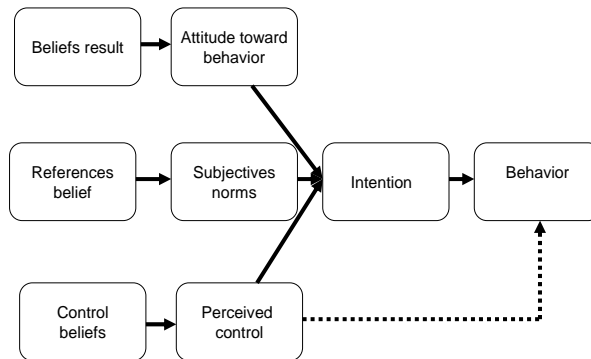**Figure 2. Model of the environmentally responsible purchase**

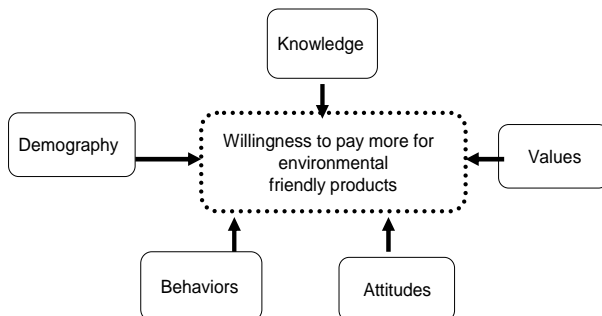**Figure 3. Model of the theory of behavior**

Figure 4. Model of the ecological behavior of purchase and the willingness to pay a higher price

behavior in green markets in 1999 by Stavros P. Kala-

fatis [3] and their team of collaborators. The variables that this model proposes are: the beliefs, the result and the attitude towards the behavior, the reference beliefs and subjective norms; and these last ones are the guidelines that individuals are taking in consideration when they decide what they must or must not do, as presented in Figure 3.

The model of the ecological behavior of purchase and the willingness pays a higher price developed by Laroche, Bergeron and Barbaro–Forleo [4]. It focuses on the study of the factors that determine the willingness of individuals to pay a higher price for environment friendly products. These factors can be grouped in three categories: demographic characteristics, level of knowledge, and values and attitudes that consumers have (or adopt) towards environmental topics. The model is presented in Figure 4.

Table 1 shows five main common variables extracted from these models to provide further insights about green behavior: orientation man-nature, perceived control, ecological knowledge, personal consequences, and environmental consequences; with the purpose of eventually applying them in environmental marketing strategies.

For this specific study these five variables are being related to some interesting results found in surveys applied to Mexican individuals regarding their green behavior, in order to suggest some recommendations on how to launch an efficient environmental marketing campaign in Mexico.

3. Orientation Man-Nature

This variable represents the values that govern the relationship between men and nature. The degree of awareness towards environmental problems is revealed as one of the key factors of green behavior. It is therefore important for an organization to determine the level of awareness in its potential market, especially regarding environmental issues directly related to its activity or product. This level of awareness is driven by ecological consciousness, which is present in most of the models of green behavior [11,12]. Barreiro *et al*. [13] divide consciousness in cognitive and emotional aspects; that is, those related to the ecological knowledge of individuals and to their environmental awareness, focused on general ecological attitudes. Green consciousness is understood as a feeling, a representation, an image, a purpose, an attitude, a course of action, and a trend of behavior whose main value is to protect the environment and improve quality of life [14]. As mentioned before, this green consciousness is the starting point for the generation of green consumer behavior. Ecological behavior is defined as "the action that one person does, either individually or collectively, in favor of the conservation of natural resources and aimed to obtain a better quality of

iB

Table 1. Variables of this study

Models	Theory of the behavior of the ecological purchase Kalafatis S.P. et. Al (1999)	Model of the environmentally responsible purchase Follows-Jobber (2000)	Conceptual model of the behavior of ecological purchase Chan –Lau (2000)	Model of the behavior of ecological purchase and willingness to pay a higher price Laroche et. Al(2001)
Independent variables	Beliefs result	Self-importance	Orientation man-nature	Demography
	References beliefs	Conservation	Ecological feeling	Knowledge
	Control beliefs	Self-indulgence	Ecological knowledge	Values
	Behavior	Environmental consequences		Behaviors
	Subjective norms	Individual consequences		Attitudes
	Perceived control			
Dependant variables	Intention	Buying intention	Green buying intention	Willingness to pay a higher price
	Behavior	Purchase behavior	Current purchase behavior	

the environment"[15]. Barreiro *et al.* [13] emphasize that the environmental consciousness of individuals is divided into two main attitudes: a first related to the concern that suggests the protection of the environment, with special references to issues regarding the pollution of that environment; and a second one related to the impression that individuals have of who is truly responsible for protecting that environment and should, by their actions, be concerned about making decisions that will help to solve these problems; in this case, they detected that individuals believe those who should indeed be responsible are countries, governments and, the society in general, assuming that they could not themselves do anything to protect the environment [16].

On the contrary, values can significantly reduce green behavior, although they would hardly eliminate it entirely. Stern, Dietz and Kalof [17] classify the values in three groups: values of selfish orientation, that represent the concern for oneself; values of biosphere orientation, that represent the concern for non-human species and the biosphere as a whole; and values of social or altruistic orientation, that show concern for others. Under this scheme, if people are more oriented towards biosphere values, one could expect that their beliefs about ecological behavior would consider the consequences of some behavior on the environment, whereas if people privilege selfish values, their beliefs would be based on the evaluation of the possible consequences for themselves. In contrast, if people are concerned about others, their beliefs should also be oriented towards the consequences of their behavior on others.

Some studies point to women as more concerned about environmental issues and ecological behaviors [17–19]. However, Zelezni, Chua and Aldrich [20] indicate that

environmentalism does not start in adulthood, and refute the argument that women are more concerned with environmental issues because of their maternity and their interest in protecting their family from environmental threats. A revision of the literature [21–24] also highlights evidence of green conscience in western countries, reflected in the consumer's decisions to buy environmental friendly products and actions taken by companies and governments to sustain this interest and sensibility with the generation of green products on one hand, and the adoption of laws to protect the environment on the other hand.. However, Elliott [25] emphasizes the difficulty companies can have in implementing green measures and even some resistance by consumers to accepting these green technologies, and suggests that these be accompanied with social, economics and political strategies. Rivera-Camino [26], Straughan R. D and Roberts J. A., [27], and Tadajewski y Wagner-Tsukamoto [28] agree to the above, showing in addition the importance of considering demographic, anthropological and psychological aspects to more effectively identify the profile of the targeted market segment in order to apply efficient environmental marketing strategies.

4. Perceived Control

This variable determines the degree of responsibility individual feel towards their possibilities to reduce pollution problems. Attitudes and general environmental concerns are not necessarily predictors of attitudes and behavior regarding specific or isolated acts. In fact, a person may show a favorable attitude towards the environment in general and a negative attitude towards recycling, e.g. by considering this process uncomfortable. It is also quite possible that each concrete ecological behavior

(purchase, recycling, etc) could involve specific variables and influence in different directions and means. In her study of values and attitudes on environmental pollution in Mexico, Durand [16] reveals that in the specific case of Mexicans, fewer are willing to make an effort to collaborate in the war against pollution than those who are not. On the other hand, when they believe that pollution is not a serious issue, the number of Mexicans who are willing to collaborate will increase. In other words, the willingness of Mexicans to act apparently occurs in inverse proportion to the perceived gravity of the problem, which probably means that under very negative environmental circumstances the effort of collaboration by individuals loses its meaning or is considered worthless with regard to the size of the challenges. This is an important result as to the way the environmental problem is being publicized, since presenting problems such as pollution or deforestation to the people in an alarming way without relating them to potential solutions at their reach may paralyze or decrease their capacity to act. Truly, the concerns of people and their interests are associated with their daily lives and their particular world. According to Hernandez *et al*. [29] the perception of ecological crime is very different, whether the non-environmental behavior is affecting one's own convenience and/or banking accounts. As to the willingness to recycle, Diaz and Palacio [30] show that the recycling behavior of consumers can not only be better represented through a model of habit or low-involvement routine "know-do-feel" than through the classical hierarchy "know-feel-do", but it can also be represented through other models such as the hedonic "feel-do-know" and the reverse learning "do-feel-know". In addition, the age and zone of residence and to a lesser extent gender, level of studies, and income affect the way in which the consumers recycle. Also domestic conditions and the existence of legal and economic incentives foment a slightly different recycling response. Pato [31] describes the possibility of multiples causes that influence a proactive ecological behavior, characterized by actions related with the preservation and conservation of the environment, and with buying decisions and use of products considered environmental friendly, or a decision to reject products perceived as harmful to the environment. Thereon Pato y Tamayo [32] indicate that the proactive behavior suggests that individual and collective actions benefit the collective indistinctly, and would imply equality of recognition of the persons subject to these actions. Similarly these actions generally seek to improve conditions and quality of life, and suggest respect of other people, and even other forms of life. They add that it is different to claim that a specific value does not produce any effect on an ecological behavior more than saying that such value influences the

beliefs people have on the environment, making them receptive to act in a pro or anti ecological manner. D'Souza and Taghian [33] reiterate this assertion, by proving the relationship between consumers's appreciative of and reactions to green advertisement, and their low or high feeling of commitment to the environment. Regarding the later it is interesting to mention the conclusion of a survey applied in Europe by Munuera *at al*. [34], who emphasize the importance of cultural background in the decision to purchase green products.

As mentioned earlier, the decision to buy green products may be affected by the surroundings and socio-economic circumstances. There are also some controversies in terms of the perception consumers have of green products. According to a survey launched at the beginning of 2008 by the marketing agency Ecoalign[1], even though there is a great concern for the environment, consumers think that many forms of green technology (renewable and recycled material) are expensive, difficult to understand and take care of, and not very esthetically attractive. This is being confirmed by Elliott [25]. In contrast, according to the Mango Sprouts[2] agency's survey, consumers are not only willing to pay up to 5% more for environmentally friendly products [35–38] , but are looking to support companies and distribution centers that have adopted green practices with their purchases.

5. Ecological Knowledge

This variable defines the sets of ecological knowledge that an individuals have of environmental topics. According to Barreiro *et al*. [13], the knowledge of environmental problems can be in a continuum ranging from mere knowledge of their existence to the detailed knowledge of their causes, effects, affected areas, actors responsible for the problem, solutions, and agents responsible for these latest. Such knowledge can derive mainly from three sources:
  • Cultural tradition:
  It is called the ecological ethnocentrism;
  • Diffusion of scientific knowledge on these issues:
  Increasingly, environmental issues are more present in the media, school curriculum and numerous recreational activities. This results in "socialization" with new and greater environmental content.
  • Personal experience of problems, their risks and effects:
  Personal experience only counts if the person has developed some sort of cognitive scheme based on their experience. This scheme must include some insight into its definition, its causes, effects and extent, and those responsible for provoking and resolving the problems.

The degree of information someone has on a given environmental problem will largely determine their opinion regarding this problem. What is true is that not all environmental problems require the same amount of previous

information, since "familiarity" with some of them provides this knowledge through one's own experience. For that the scale of the problem, i.e. its proximity and/or reach, its intensity or frequency, and its compatibility with other problems and social needs are the main features that condition the knowledge and the appreciation of the problems—and eventually, the corresponding behavior. Ecological knowledge is also supported by past behavior and perceptions about recycling and second-hand products purchase. According to research from Aguilar *et al.* [39] past behavior directly influences the intention to recycle, as part of green behavior. He emphasizes the significant and independent contribution of past behavior to the variance of the intention, once variables such as attitudes, social norms or perception of behavioral control are taken into account. Gregson & Crewe [40] also describe second-hand purchasing as an alternative form of green consumption, implying therefore a change in the perception of used goods, otherwise interpreted as waste.

6. Individual Consequences

This variable determines the degree of knowledge that a person has of the individual consequences of a pollution problem, or on the contrary, of a more friendly action on the environment. In relation to the individual consequences, Durand [16] contradicts previous positions expressed by Inglehart [41], suggesting that environmentalists' values emerge when caring for the environment is transformed into a new component of security or individual well-being, i.e. when it becomes part of a survival process. Thus, it does not only occur in wealthy sectors, but also in less favored ones. According to Inglehart [41], from the perspective of the theory of post-material value changes, materialistic values linked with economic security are substituted by postmaterial values or associated to emotional, esthetic or intellectual needs. Supported by the hypothesis of scarcity, the author explains that in those societies that have met their basic needs, individuals and social groups grant less importance to materialistic values (economical success, public and individual security) and come to worry about things that are now more complicated to get, usually related to the quality of life and social relationships affected by industrial development. It's only once men and women have satisfied those needs that they consider prioritary or live according to those values they consider primary, that they will eventually seek new goals, and new values will emerge to guide them towards these goals. In postmaterial societies the environmental problem is being reacted to by concern for global warming, pollution, and transgenesis. While in less prosperous societies, problems such as bad harvests, flood, and diseases are standing out. In this case, the concern for environmental issues may occur in nonindustrial countries as a component of personal security,

that is, as a new material value. Durand [16] argues that it is likely that in non postindustrial societies, such as Mexico, environmental values emerge from the coexistence of interest and priorities that respond to the confluence of both industrializing and post industrial structures. From the survey the author did on environmental pollution in Mexico [16], interesting results emerged concerning individual consequences that the surveyed Mexicans perceived. Among other interesting findings was that the majority of the respondents questioned the existence of environmental problems, that is, they are not sure if those who talk about pollution are overreacting, or are fighting a real problem. They consider that some social purposes, such as the need to create jobs may justify the existence of environmental pollution. Another interesting point that Durand shows [16] is the correlation between level of education and income with the awareness of environmental problems. Indeed the results of the survey show that the environmentalism index only relates to those groups that have a better standard of living. The author also points out that having environmentals values does not depend on having either postmaterial values or a high standard of living. It surely increases with socioeconomical status, but analysis reveals that environmentalism in less advantaged sectors is accompanied with material values, i.e. tied to survival.

7. Environmental Consequences

This variable determines the degree of knowledge that individuals have of the environmental consequences of a lifestyle or a behavior. Barreiro *et al.* [13] manifest that the global dimension of the problem of environmental degradation has converted it into a "global" problem, clearly exposing the insufficient measures adopted individually by each nation. Conservation of the natural environment is a concern that transcends borders; therefore a planetwide joint action becomes necessary to achieve the combination of economic growth and the environment conservation once and for all. Since the end of the 20th century, the green house effect, air and see pollution, and the destruction of the ozone layer have been some of the problems perceived by a growing number of population sectors, who are making worthy efforts to produce, through economical, social and political strategies, radical changes in the production process of industrialized societies characterized by high consumption. Durand [16] indicates that the appreciation of environmental care can be manifested in actions or incorporated in lifestyles when factors that determine wellbeing, in the short or long term, can be combined with the desire to preserve nature, which in western countries usually occurs when the material development goes beyond the subsistence level and the environmental values do not interfere with acquiring food or maintaining health.. The history of environmentalism as a social movement is consistent

with the ideas that changes in social structure are causing changes in values and priorities. Durand [16] claims that the strengthening and the dispersion of environmental values can not be understood only as a logical consequence of the intensification of environmental issues in local and global dimensions that are being pointed out more vigorously since the mid-20th century, or of our scientific and technological capacity to perceive problems inaccessible to our senses. On the contrary, it has to be seen as an action and reaction of the configuration of a new social order.

An interesting fact that Durand [16] also reveals is the lack of correlation between environmental values and city size in Mexico. The inhabitants of cities such as Mexico City, Guadalajara and Monterrey, which have high levels of pollution, do not show a greater concern; neither did they show a more cooperative attitude than inhabitants of less polluted and smaller cities. This phenomenon can be explained as an "adaptation" involving a denial process and suppression of consciousness of environmental risk, in exchange for obtaining other basic or superficial benefits, subordinating the conservation value to others that put more emphasis on short-term than long-term welfare. Durand [16] concludes that living in environmentally degraded areas seem not to involve the emergence of values linked to environmental care, the commitment of action or even the perception of the deteriorated environment.

8. A Specific Approach of the Mexican Context

In the environmental field Mexico is a country of contrasts. First of all it is one of the 12 environmentally megadiverse countries of the world. In fact it ranks third among the countries with the greatest biological diversity. It is in the first place with respect to its fauna of reptiles (717 species), second in mammals (451 species), and fourth in amphibians (282 species) and flowering plants (around 25,000 species). 32% of its national vertebrate fauna are endemic to Mexico, of which 52% are shared only with other Mesoamerican countries [42,43,45]. The high biological diversity of Mexico is a combined result of both variations in topography and climate found in its territory. These are mixed with each other, creating a mosaic of environmental conditions and micro environments, in addition to the complex geographical background of the area, in particular the southeast part of the country, better known as the Nuclear Central America [44]. There has been major progress in the field of environmental management. Mexico occupies one of the most outstanding places in the world in caring for protected areas, that are being run by the "Comisión Nacional de Areas Naturales Protegidas" (CONANP, National commission of protected natural areas). In 2001 the CONANP was managing 61 protected areas, today it takes care of 171 (23,877.977 ha), more than 97% of the total protected area of the country. As a result of the effort of the CONANP to preserve the ecosystems and the biodiversity of Mexico, 98.2% of the protected areas of the country currently enjoy an annual operating program, with skilled human resources and equipment to apply strategies of conservation and development [46].

Worthwhile mentioning aside the above mentioned progress are however some very serious gaps in other fields of environmental management. In terms of legal issues, even though there is a federal law, the environmental rules and norms show a lack of legal consistency as well as some considerable lack of interpretation. Environmental education is also far for being uniform: there are in reality 32 different programs. In terms of economy, while there has been some interesting mechanisms installed to foment the protection, such as payment for environmental services and servitude, some perverse subsidies still remain, the most serious one is for the gasoline that is rather an award for polluting, and among others consequences this subsidies has impeded the technological transitions towards the modernization of the fleet. In terms of organization, the structures of environmental organizations are supported by heavy and inefficient pyramidal function. Likewise some lack of action in metropolitan areas can be observed that provokes the deterioration of the environment. The social participation also shows great backwardness. Indeed it is limited to some hired technical advisory committees that rather allow validating some previous decision taken. On the other hand the non profit organizations are playing an increasingly role but are still at an initial process and do not constitute a social impact that would limit the environmental deterioration.

9. Evaluation of the Green Behavior of the Mexican Individuals with Respect to the Studied Variables

As a summary of the deficits in environmental management in Mexico, we could describe it as following. There is a poor orientation man-nature, a very low perceived control by the population that believes that it is the government's responsibility to take care of the environment. The ecological knowledge is very poor too. Indeed there is not one Phd program in environmental education and there are barely three masters programs related to the subject in the 1,200 universities of the country. The personal consequences are appreciated in the upper-level class only and finally the environmental consequences are poorly appreciated in the country. Despite being the country with most national communication launched, there is a big gap between the official programs and the citizen participation. In addition, the highly unequal social structure allows distinguishing at least three cultural

and social groups:

1) People related to the post-industrial countries, for either working with them of for economic or cultural links.

2) People belonging to the medium class that barely ever leave the country.

3) And at least 40% of Mexicans that are classified as poor, not only because of their lack of patrimony, but also for nutritional deficiency, and for whom the green behavior, with the exception of a few, represent a difficult concept to understand.

For the first group, one could say that their practices of environmental care are very similar to other post industrialized countries. For the second group of the population, one could claim that there is a growing concern for environmental aspects and finally for the least favored sector of the population, with the exception of a few, the majority has little knowledge about the subject.

10. Conclusions

The revision of the variables of the green behavior from the perspective of Mexican individuals shows the relevance of adding aspects of development and social justice in the environmental speech that promotes in Mexico the preservation of the natural environment. The existence of environmental values in sectors that are still struggling for achieving decent standard of living, show that those values are not superfluous components in the vision of the future, but it seems that the connection between a better future and the environmental responsibility is not yet clear to most of Mexicans. It is therefore necessary to expand the interest on environmental issues among Mexicans and generate values in the particular context of different population sectors, so that they can acquire consistency and gain a real sense. It is indeed perceived that in the case of Mexicans, the importance of the environmental component in the quality of life in the short and long term has not been achieved yet. This situation is probably due to the scarcity of economical resources to implement campaigns and projects focused towards the environmental sustainability beyond the local context, but it is maybe also due to the kind of environmentalist speeches that are being publicized that have a strong ecological bias, i.e. they ponderate environmental conservation and the links between the development and the quality of life. Whereas the search for solving more basic and immediate problems, such as acquiring food, maintaining health and keeping a job, or survival issues, should find its sustain in predominant environmental speeches, worried so far essentially for the protection of species and the simple no alteration of the ecosystems.

REFERENCES

[1] Chan, R. and Lau, L., "Antecedents of green purchases: A survey in China," Journal of Consumer Marketing, Vol. 20, No. 49, pp. 338–357, 2000.

[2] Follows, S. and Jobber, D., "Environmentally responsible purchase behavior: A test of a consumer model," European Journal of Marketing, No. 5–6, pp. 723–746, 2000.

[3] Kalafatis S., Pollard M., East R., and Tsogas M., "Green marketing and Ajzen's theory of planned behavior: A cross-market examination," Journal of consumer marketing, Vol. 16, No. 5, pp. 44–46, 1999.

[4] Laroche, M., Bergeron J., and Barbaro-Forleo, G., "Targeting consumers who are willing to pay more for environmentally friendly products," Journal of Consumer Marketing, Vol. 18, No. 6, pp. 503–520, 2001.

[5] Lopez, A. G., "La preocupación por la calidad del miedo ambiente, Un modelo cognitivo sobre la conducta ecológica," PH.D. dissertation, Universidad Complutense de Madrid, 2002.

[6] García-Mira, R. and y Real-Deus, E., "Valores, actitudes y creencias: Hacia un modelo predictivo del ambientalismo," Medio ambiente y Comportamiento humano, Vol. 2, No. 1, pp. 21–43, 2001.

[7] Corraliza, J. A. and Martín, R., "Estilos de vida, actitudes y comportamientos ambientales," Medio ambiente y comportamiento humano, Vol. 1, No. 1, pp. 31–56, 2000.

[8] Werner, C. M. and Makela, E., "Motivations and behaviors that support recycling," Journal of environmental psychology, Vol. 18, No. 4, pp. 373–386, 1998.

[9] Corraliza, J. A. and Berenguer, J., "Environmental values, beliefs and actions: A situational approach," Environment and Behavior, Vol. 32, No. 6, pp. 832–848, 2000.

[10] Gurau, C. and Ranchhod, A., "'International green marketing: A comparative study of british and romanian firms'," International Marketing Review, Vol. 22, No. 5, pp. 547–562, 2005.

[11] Bigné, E.,"El consumidor verde: bases de un modelo de comportamiento", Esic-Market, Vol. 96, pp. 29–43, april 1997.

[12] Chan, R., "Consumer responses to environmental advertising in China," Intelligence & Planning, Vol. 22, No. 4, pp. 427–437, 2004.

[13] Barreiro, J. M., Lopez, M. A., Losada, F., and Ruzo, E., "Análisis de las dimensiones cognoscitiva y afectiva del comportamiento ecológico del consumidor," Revista galega de economía, Vol. 11, No. 2, pp. 1–21, 2002.

[14] Faix, W. G., Kurz, R., and Wichert, F., "Innovation zwischen Ökonomie und Ökologie," Moderne Industrie Verlag, Landsberg, Lech, 1995.

[15] Castro, R., "Naturaleza y función de las actitudes ambientales," Estudios de psicología, Vol. 22, No. 1, pp. 11–22, 2001.

[16] Durand, V. M. and Durand, L., "Valores y actitudes sobre la contaminación ambiental en México. Reflexiones en torno al posmaterialismo," Revista mexicana de sociología, Vol. 66, No. 3. pp. 511–535, 2004.

[17] Stern, P. C., Dietz T., and Kalof, L., "Value orientations and environmental concern," Environment and Behavior, No. 25, pp. 332–348, 1993.

[18] Schultz, P. W., "The structure of environmental concern:

concern for self, other people, and the biosphere," Journal of Environmental psychology, Vol. 21, No. 4, pp. 327–339, 2001.

[19] Dietz, T., Kalof, L. and Stern, P., "Gender, values, and environmentalism," Social Science Quartely, Vol. 83, No. 1, pp. 351–364, 2002.

[20] Zelezny, L. C., Chua, P. P., and Aldrich, C., "Elaborating on gender differences in environmentalism," Journal of environmental Psychology, Vol. 18, No. 4, pp. 373–386, 2000.

[21] Connelly, J. and Smith, G., "Politics and the environment: From theory to practice," London, Routledge, 2002.

[22] Wagner, S. A., "Understanding green consumer behaviour: A qualitative cognitive approach," London, Rouledge, 2002.

[23] Szarka, J., "The shaping of environmental policy in France," Contemporary France, New York, Vol. 6, Berghahn, 2002.

[24] Lafferty, W. M. and Meadowcroft, J. R., "Implementing sustainable development: Strategies and initiatives in high consumption societies," Oxford, UK, Oxford University Press, 2000.

[25] Elliott, D., "Energy, society, and environment: Technology for a sustainable future," Routledge, Introductions to Environment Series, New York, Routledge, NY, 2003.

[26] Rivera-Camino, J., "Re-evaluating green marketing strategy: A stakeholder perspective," European Journal of marketing. Vol. 41, No. 11–12, pp. 1328–1358, 2007.

[27] Straughan, R. D. and Roberts, J. A., "Environmental segmentation alternatives: A look at green consumer behavior in the new millennium," Journal of Consumer marketing, Vol. 16, No. 6, pp. 558–575, 1999.

[28] Tadajewski, M. and Wagner-Tsukamoto, S., "Anthropology and consumer research: qualitative insights into green consumer behavior," Qualitative market research: An international journal, Vol. 9, No. 1, pp. 8–25, 2006.

[29] Hernández, B., Martin, A., Hess, S., Martínez, J., Suárez, E., Salazar, M., *et al*., "Análisis multidimensional de la percepción del delito ecológico," Medio Ambiente y Comportamiento Humano, No. 6, pp. 51–70, 2005.

[30] Díaz, M. and Palacio, B., "El proceso de adopción de la conducta de reciclado: Modelos explicativos y variables moderadoras," Cuadernos de Economía y Dirección de la Empresa, No. 28, pp. 55–86, 2006.

[31] Pato, C., "Comportamento ecológico: Relações com valores pessoais e crenças ambientais," PH.d. dissertation, Universidade de Brasília, 2004.

[32] Pato, C. and Tamayo A., "Valores, creencias ambientales y comportamiento ecológico de activismo," Medio Ambiente y Comportamiento Humano, Vol. 7, No. 1, pp. 51–66, 2006.

[33] D'Souza, C. and Taghian, M., "Green advertising effects on attitude and choice of advertising themes," Asia Pacific Journal of Marketing and Logistics, Vol. 17, No. 3, pp. 51–66, 2005.

[34] Munuera, J. L. and Pemartin, M., "El consumidor europeo de productos ecológicos," Distribución y consumo, pp. 50–64, January–February 2006.

[35] Kassahye, W. W., „Green dilemma," Marketing Intelligence & Planning, Vol. 19, No. 6, pp. 444–455, 2001.

[36] Coddington, W., "Environmental marketing: Positive strategies for reaching green consumers," Mc Graw Hill, New York, NY, 1993.

[37] Davis, J. J., "Strategies for environmental advertising," Journal of Consumer marketing, Vol. 10, No. 2, pp. 19–36, 1993.

[38] Ottman, J., "Green marketing, challenges and opportunities for the new marketing age," NTC Business Books, Loncolnwood, IL, 1992.

[39] Aguilar, M., Monteoliva, A., and Garcia, J. M., "Influencias de las normas, los valores, las creencias proambientales y la conducta pasada sobre la intención de reciclar," Medio Ambiente y Comportamiento Humano, Vol. 6, No. 1, pp. 23–36, 2005.

[40] Gregson, N. and Crewe, L., "Second-hand cultures, Materializing culture," Berg, Oxford, 2003.

[41] Ingelhart, R., "Modernization and postmodernization. cultural, economic and political change in 43 societies," New Jersey, Princeton University Press, 1996.

[42] Halffter, G., "La diversidad de Iberoamérica," Acta Zoológica Mexicana, Inst. de Ecología, SEDESOL y CYTED, México, Vol. Especial, 1992.

[43] Retana, O. G. and Lorenzo, C., "Lista de los mamíferos terrestres de Chiapas: Endemismo y estado de conservación," Acta Zool, Mex., (n.s), No. 85, pp. 25–49. 2002.

[44] Flores, V. O. and Gerez, P., "Biodiversidad y conservación en México: Vertebrados, vegetación y uso del suelo, 2ª edición," CONABIO-UNAM, México, pp. 439, 1994.

[45] Toledo, V. M., "La diversidad biológica de México," Ciencia y Desarrollo, Mexico, 1988.

[46] Comision Nacional de Areas Protegidas, CONANP, 2009, Extracted on the 24 of october, 2009, http://www.conanp.gob.mx/