Market entry and priority of small and medium-sized enterprises in the software industry: An empirical analysis of cultural distance, geographical distance, and market size

Arto Ojala and Pasi Tyrväinen
University of Jyväskylä

Abstract

This paper investigates the influence of cultural distance, geographical distance, and three market size variables to the target country preference of SMEs in the software industry. In addition, we examine the shift of priorities in SMEs’ country selection by analyzing how these factors impact the selection of the first, second, and third target countries. Empirical findings of this study suggest that almost 70% of country choices can be explained by software market size and geographical distance alone. The findings adduced also that entry priority of SMEs shifts very fast from countries within a short geographical distance to countries with high purchasing power and within a greater geographical distance. Relations of these macro-level factors are discussed and compared to earlier qualitative and conceptual studies in this field. Implications for managers are also provided.

INTRODUCTION

Today, high-technology SMEs are playing increasingly important role in global markets. These firms, described as ‘born global’ or ‘international new ventures’, have a strong international focus from their inception (Bell et al. 2003; Coviello 2006; Oviatt and
McDougall 1994, 2005). Among the growing amount of research investigating to rapidly internationalizing firms, most of studies focus on network relationships (Bell 1995, Coviello and Munro 1997; Zain and Ng 2006), knowledge (Bell et al. 2003; Jones 1999; Oviatt and McDougall 2005), and resources (McDougall et al. 1994; Young et al. 2003) behind the internationalization behavior of entrepreneurial SMEs. While we acknowledge the importance of these studies using a firm specific data, there seems to be a lack of studies focusing on macro-level variables behind the country selection of high-technology SMEs.

As highlighted in several traditional internationalization theories (Davidson 1980; Dunning 1973, 1992; Johanson and Vahlne 1977; Johanson and Wiedersheim-Paul 1975; Luostarinen 1979), psychic or cultural distance, geographical distance, and market size have had a major impact to the target country selection. Although, these traditional theories are based on empirical findings of large multinational corporations, empirical findings (Bell 1995; Coviello and Munro 1997; Moen et al. 2004), and conceptual models (Bell et al. 2003; Jones and Coviello 2005) related to international entrepreneurship are somewhat inline with these theories. These studies imply that the above factors have an important role when firms choose their target countries, with the exception that the internationalization process of SMEs in high-technology sectors is very rapid and that the firms there use a number of different entry modes simultaneously (Coviello and Munro 1997, among others). However, these studies provide very little knowledge of interrelations between these factors (psychic or cultural distance, geographical distance, and market size) and the target country preference. In addition, these previous studies
(Bell 1995, 1997; Bell et al. 2003; Jones and Coviello 2005; Coviello and Munro 1995, 1997; Moen et al. 2004) indicate that the importance of psychic or cultural distance and geographical distance is the greatest in the first country selections while the importance of market size is more related to later country selections. Although this statement seems to be reasonable, these studies do not provide empirical evidence to this proposition.

For these above mentioned reasons, we conducted an empirical study to examine the influence of cultural distance, geographical distance, and three market size variables (GDP, GDP per capita, and the vertical market size) to the target country preference of SMEs in the software industry. In more detail, we examine the shift of priorities in SMEs’ country selection by analyzing the factors that best predict the selection of the first, second and third target countries. By analyzing these macro-level factors behind the target country selection gives us an opportunity to compare quantitative findings of this study to earlier qualitative findings related to internationalization of software firms (Bell 1995; Coviello and Munro 1997; Moen et al. 2004; Zain and Ng 2006). This enables theory development and complement of previous studies in this area. In addition, this study provides macro-level indicators which best specify the criteria for country selection. This is important for managers of SMEs because these firms are commonly characterized by having very limited resources, market knowledge, and network relationships at the time of the initial market entry (Bell 1995, 1997; Coviello 2006; Coviello and Munro 1997; Jones and Coviello 2005, among others). Thus, managers’ capabilities to acquire detailed micro-level market information are very limited at the time of their first market entries.
THEORY AND HYPOTHESIS DEVELOPMENT

Traditional internationalization theories (Davidson 1980; Dunning 1973, 1992; Johanson and Vahlne 1977; Johanson and Wiedersheim-Paul 1975; Luostarinen 1979) suggest that psychic or cultural distance, geographical distance, and market size have a major impact on target country selection. Although the impact of these three factors seems to be unquestionable, their interrelations are difficult to determine. Typically the impact of all these factors is not visible simultaneously due to the conditions of the case analyzed, and this is due to several reasons. Firstly, it is difficult to separate statistically the impact of low cultural and geographical distance, if these two factor correlate strongly with each other. In some cases, countries within a low cultural distance are not located geographically close to each other (for instance Australia and UK), which makes it possible to observe, separately, the impact of these two factors (Dow 2000). This is usually not the case, however. Secondly, the significance of cultural distance is also more difficult to observe if the home country is culturally close to most of the countries with large markets. Thirdly, it is also difficult to observe statistically significant impact of geographical distance, if most other countries are equally distant due to limited number of near neighbors, as in the case of Australia, or if the large markets favored by the firms are geographically close, such as in the case of the U.S. market to Canadian companies. These methodological reasons have contributed much to the lack of statistical evidence.
supporting the theories developed for the interrelationships of the three main factors impacting country selection.

Many studies (Davidson 1980; Erramilli 1991; Johanson and Vahlne 1977; Johanson and Wiedersheim-Paul 1975; Luostarinen 1979) also suggest that firms enter first to countries within a low geographical and/or psychic distance, and when they gain more international experience, they start to expand to countries with a larger market size and greater psychic and geographical distance. Although this seems to be a well accepted argument in the international business literature, very little empirical research have been done on how soon the focus changes from nearby countries within a low psychic and geographical distance to more distant countries with a large market size.

Next, we go through literature that is related to the impact of these three factors on target country selection and priority. In each sub-section, we first review more general literature related to the phenomena under investigation and based on current research findings in the field of international entrepreneurship. After that, we draw a hypothesis based on the literature presented.

**Cultural distance and target country selection**

An increasing amount of international business and management studies have focused on cultural or psychic distance between the home and target country. As noted in studies of Dow and Karunaratna (2006) and Sousa and Bradley (2006) many authors have used
these terms interchangeably to explain differences in cultural values among countries either at a national or individual level. Beckerman (1956) introduced the concept of psychic distance between countries, and it became popular after Nordic studies (Johanson and Vahlne 1977; Johanson and Wiedersheim-Paul 1975). These studies conceptualized the internationalization process as step-wise progress where a firm starts its internationalization by expanding to countries nearby, and when it gains more international experience, it then starts to develop activities in countries within a greater psychic distance. Factors that determine the psychic distance between countries were defined as differences in language, culture, political system, level of education, level of industrial development etc. (Johanson and Wiedersheim-Paul 1975). These pioneering studies, have been followed by a growing amount of research focusing on the concept of psychic distance.

However, measurement of psychic distance has been complicated: various authors (Conway and Swift 2000; Dow and Karunaratna 2006; Evans and Mavondo 2002; Johanson and Wiedersheim-Paul 1975; Swift 1999) have suggested different factors that may influence to psychic distance, and there seems to be a lack of agreement in this respect. In addition, most of these factors are difficult to measure by using publicly available statistics. For these reasons, many authors (e.g., Benito and Gripsrud 1992; Gomez-Mejia and Palich 1997; Morosini et al. 1998; Steensma et al. 2000) have focused only on measurement of cultural distance even though it is only one of the dimensions of psychic distance as highlighted in Dow and Karunaratna (2006) and Evans and Mavondo (2002). For instance, Evans and Mavondo (2002) define a summary construct of Psychic
Distance, which contains Cultural Distance and Business Distance. In their study, measuring organizational performance, use of these two dimensions produces approximately as good explanatory power as the single dimension of Psychic Distance. However, if both of these two dimensions are disaggregated into five individual factors for regression analysis, the resulting model of ten dimensions has much higher explanatory power.

Findings in studies operationalizing cultural distance based on the Hofstede’s (1980; 2001) cultural dimensions and the composite index of Kogut and Singh (1988) diverge. Many authors have criticized the usage of cultural distance (Dow and Karunaratna 2006; Shenkar 2001; Tihanyi et al. 2005) to estimate cultural values between countries and have failed to find support for it (Dow 2000; Dow and Karunaratna 2006; Tihanyi et al. 2005). However, others (Evans and Mavondo 2002; Jones and Teegen 2001; Morosini et al. 1998; Rothaermel et al. 2006) have found further evidence for the Hofstede’s cultural dimensions. This same discrepancy occurs in studies investigating managerial decision making in the country selection process. Study of Robertson and Wood (2001) adduced that cultural environment in a target country was the least important factor affecting the managers’ decision making process. In contrast, Edwards and Buckley (1998) found that cultural familiarity was the most important factor for Australian managers’ entry decision to enter the UK market.

Studies in the field of international entrepreneurship (Arenius, 2005; Bell 1995; Chetty and Campbell-Hunt 2004; Coviello and Munro 1997; Moen et al. 2004; Spence 2003)
have focused mainly on qualitative studies and evaluated psychic distances between countries by using the concept of Johanson and Vahlne (1977) without any systematic measurement. The common finding in these studies was that entrepreneurial SMEs do not follow incremental step-wise progressing internationalization as suggested in the Uppsala model (Johanson and Vahlne 1977; Johanson and Wiedersheim-Paul 1975). The study of Chetty and Campbell-Hunt (2004) focusing on SMEs in New Zealand proposed that entrepreneurial firms select their first target markets by entering to countries within a low psychic distance, such as Australia and the UK. However, having selected the first market, these firms quickly start entering countries within a greater psychic distant. This conclusion is also consistent with the findings of Coviello and Munro (1997), Moen et al. (2004), and Spence (2003). Bell (1995) conducted a survey involving 88 small software firms from Finland, Ireland, and Norway. He found that 50-70 percent of firms enter countries within a low psychic distance, indicating psychic distance as a key factor for country selection. In terms of cultural distance, Rothaermel et al. (2006) discovered support to the hypothesis that a greater cultural distance increases the probability of market entry of U.S. Internet firms.

Altogether, both traditional internationalization theories (Johanson and Vahlne 1977; Johanson and Wiedersheim-Paul 1975; Luostarinen 1979) and current findings in international entrepreneurship literature (Bell 1995; Chetty and Campbell-Hunt 2004; Coviello and Munro 1997; Moen et al. 2004) suggest that firms enter first countries within a low psychic or cultural distance and thereafter start to develop their operations in countries within a greater cultural distance. However, very little empirical evidence
exists on how fast these firms’ focus changes from culturally familiar countries to countries within a greater cultural distance. Therefore, we hypothesize that firms favor counties within a low cultural distance, but only the first country selected is culturally closer than later countries. Accordingly:

**Hypothesis 1a:** Cultural distance is inversely related to the target country selection of SMEs.

**Hypothesis 1b:** The first country entered by SMEs is culturally closer than countries entered later.

**Geographical distance and target country selection**

Besides cultural distance, geographical distance has been found to be a significant factor when a firm selects its target countries. Many researchers (e.g., Chetty 1999; Clark and Pugh 2001; Dow 2000; Luostarinen 1979; Srivastava and Green 1986) have found support to the concept that geographical distance has significant impact when firms select their target countries. In their study, Clark and Pugh (2001) investigated British firms’ target country selection. They found that the first three countries entered were significantly closer than countries entered later in terms of geographical distance. This was argued to be due to the lower economic and managerial costs in geographically closer countries. Similarly, Chetty (1999) found that managers of New Zealand SMEs favored Australian market in their market entry decision due to a familiar business environment and low operation costs. The study of Dow (2000) also highlighted that low
geographical distance was viewed favorably in the country selection of Australian SMEs because of lower transportation costs. In their study, Srivastava and Green (1986) adduced that geographical distance has the most significant impact on trade intensity between countries. Although there seems to be a strong support to the concept that firms internationalize their operations first to countries within a relatively low geographical distance, the study of Terpstra and Yu (1988) failed to find support to this concept. Their results suggest that U.S. advertising agencies tend to internationalize their operations with larger firm size and favor countries with a large market size instead of geographical distance.

The concept of geographical distance is also somewhat complex in software industry. In many cases, software products can be sold and distributed electronically around the world by using the Internet, which makes the geographical distance less important. However, studies investigating internationalization of software firms (Bell 1995, Coviello and Munro 1997; Moen et al. 2004) have found evidence that software firms enter first to countries within a low psychic and geographical distance before expanding to countries within a greater distance. This internationalization behavior is related to the fact that most of the software products require intensive liaison with customers in order to specify requirements, install the software, and give after-sales support (see e.g., Ojala and Tyrväinen 2006).

In summary, the above literature suggests that, in their target country selection, firms favor countries within a low geographical distance for two reasons. Firstly, the
environment in these countries is usually familiar in terms of language, culture, business practices etc. Secondly, it is less expensive to operate in nearby countries than countries within a greater geographical distant. However, similarly with cultural distance described above (hypothesis 1a and 1b), very little empirical evidence exists on the rate firms change from countries within a low geographical distant to countries within a greater geographical distance. Thus, we hypothesize that firms favor counties within a low geographical distance, but these firms’ focus moves from the first country entered to countries within a greater geographical distance.

**Hypothesis 2a:** Geographical distance is inversely related to the country selection of SMEs.

**Hypothesis 2b:** The first country entered by SMEs is geographically closer to them than later entered countries.

**Market size and target country selection**

A firm’s major objective in expanding internationally is to get access to markets that offer a large customer base for its products and services. Thus, firms tend to favor countries with attractive markets because these countries provide opportunities for greater returns (Agarwal and Ramaswami 1992; Dunning 1992). Terpstra and Yu (1988) found that a large market size of the host country (GDP) was the main determinant of U.S. advertising agencies’ foreign investments. The study of Clark and Pugh (2001) proposed that the order of market entry of British firms is strongly related to affluence (GDP per
capita) in the target country rather than to market size (GDP). This is in line with findings of Rothaermel et al. (2006) proposing that, for U.S. Internet firms selecting their target countries, a high level of GDP per capita in the target country increases the attractiveness of the country and decreases the effects of country risk, cultural distance, and national cultural values. From the managerial point of view, many studies support the view that managers tend to favor countries with a large market size in their country selection process (Brewer 2001; O’Farrell et al. 1997; Robertson and Wood 2001).

Although there seems to be clear agreement that firms prefer large markets, the studies do not give an answer to how soon these firms enter to these countries. Moreover, many of the studies suggesting that a large market size has a major impact for country selection have been conducted using sample of U.S. multinational firms (Agarwal and Ramaswami 1992; Culem 1988; Davidson 1980; Rothaermel et al. 2006; Terpstra and Yu 1988). We can assume that firms from countries with large domestic markets are able to develop their resources and then move straight to countries with a large market size, whereas limited domestic markets drives firms internationalize their operations early on (Bell et al. 2003; Spence 2003). Most of the qualitative studies (Bell 1995, Coviello and Munro 1997; Moen et al. 2004) in the field of international entrepreneurship have found support to that entrepreneurial firms focus first to countries within a low psychic and/or geographical distance and larger markets are selected later on. Shift in the focus from nearby countries to countries with a large market size seems to be somewhat unclear. Current literature also provides evidence that market size in terms of GDP and GDP per capita has a strong impact on the selection of a target country whereas the impact of a vertical market size
has been largely ignored in the previous studies. If we add the impact of the vertical market size (the software market size), relations of these constructs are not so clear. For instance, China and India have very high GDP, but the size of the software market in these countries is relatively low. By using these three market size variables, we test the best indicator for target country selection and hypothesize that, instead of GDP and GDP per capita, a large software market size in the target country has a positive impact on the selection of the target country. In addition, based on the qualitative studies related to internationalization of software firms (Bell 1995; Coviello and Munro 1997; Moen et al. 2004) we also hypothesize that SMEs in the software industry do not enter first to countries with a large market size, and that the second and third countries selected are positively related to a large market size (GDP). Accordingly:

**Hypothesis 3a:** A large market size (GDP) of the target country is positively related to the country selection of SMEs.

**Hypothesis 3b:** Affluence of the target country (GDP per capita) is positively related to the country selection of SMEs.

**Hypothesis 3c:** A large size of the software market in the target country is positively related to the country selection of SMEs.

**Hypothesis 3d:** The second country entered by SMEs is positively related to large market size (GDP).

**Hypothesis 3e:** The third country entered by SMEs is positively related to large market size (GDP).
METHODOLOGY

Scope of the Study

The empirical study targeted on finding the shift of priorities in the SMEs’ country selection strategy from a low cultural and geographical distance to a large market size, analyzing the three factors mentioned above and based on objective empirical data. This puts special requirements for the set-up: all the three above-mentioned factors need to be visible in a single empirical setting to a statistically significant degree. As discussed earlier, reaching this goal is easier if the country of origin is neither geographically isolated nor next to a very large market. The country of origin should also have culturally heterogeneous neighboring countries and should preferably not be in the same cultural cluster as most of the largest software markets. Choosing a relatively small country for our purpose makes it easier to obtain an extensive sample, but is likely to reduce the size of the companies. This is due to the early need to internationalize because of the small domestic market.

We selected small and medium-sized Finnish software firms with foreign operations as the target group of this study. In this study, SMEs refer to firms fewer than 250 employees. This limiting value is based on European Union’s and Finnish government’s definitions for SMEs (OECD, 2003), and it gives a good base for comparison with earlier studies conducted in the software industry (Bell 1995, 1997; Coviello and Munro 1997; Moen et al. 2004; Zain and Ng 2006). Due to Finland’s limited domestic market,
internationalization is generally a common growth strategy for Finnish high technology SMEs (Autio et al. 2000). Selecting a small country with a limited home market gave us a good opportunity test the influences of cultural distance, geographical distance, and market size to the country selection and priority. Finland has many neighboring countries within a low geographical and cultural distance belonging to same Nordic cluster (Ronen and Shenkar 1985) and also a neighbor within a greater cultural distance (Russia). Finland also locates fairly far from countries with large market sizes. This allows us to test hypothesis that firms tend first enter countries within a low cultural and geographical distance. Cultural differences within Finland and its Nordic neighbors are relatively low, thus cultural homogeneity in these countries helps us overcome the criticism of Shenkar (2001) related to ‘the assumption of spatial homogeneity’. In addition, selecting a small country for this study enables us to get an extensive sample, covering over 80% of the total foreign income of the software industry in Finland.

The software sector was selected as the target industry due to the fact that software firms are commonly used in international entrepreneurship literature as the individual target sector (Bell 1995, 1997; Coviello 2006; Coviello and Munro 1995, 1997; Moen et al. 2004; Zain and Ng 2006). Moreover, focusing on one single sector in this study helps compare our quantitative findings to earlier qualitative studies in the software industry (Bell 1995, 1997; Coviello and Munro 1997; Moen et al. 2004; Zain and Ng 2006) and use the industry specific variables, such as the vertical market size. Despite of a single industry focus, the software industry shares common characteristics with other high-
technology industries (Spence 2003) and service sector (O’Farrell et al. 1997). Thus, the sample is comparable, with some caution, to larger settings.

Although the sample consist of Finnish software firms, findings can be generalized to other countries due, to a large extent, to the similar internationalization behavior of software firms highlighted in many studies (Bell 1995; Coviello and Munro 1997; Loane et al. 2004; McNaughton 1996; Spence 2003; Zain and Ng 2006). These studies propose that software firms originating from Canada, Finland, Ireland, New Zealand, Norway, Malaysia, and the U.S. are generally using similar routes, networks, and strategies in their internationalization processes.

The source data in this study is based on secondary data provided by the Finnish National Software Industry Survey 2003 investigating the status of software product firms, i.e., firms that design, sell, and maintain software products. The source data was collected, using a mail questionnaire, during 2003. In addition, the largest firms were systematically approached afterwards by telephone requests to attend the survey. Altogether, 165 firms responded to the survey, a figure which presents approximately 17% of the whole industry in Finland. Although the total response rate of the survey was quite low (17%), the sample covers the target group of this study fairly well, as the firms that responded generated over 80% of the total foreign income of the software industry in Finland. From the source database of these 165 firm responses, it can be seen that 73 firms had foreign income and 53 of them had responded to the question related to the first three foreign countries the firms had entered. We excluded all the firms having 250 or more employees
and ended up with the final sample of this study consisting of 51 SMEs (with the average of 27 employees), which had international operations and had identified at least the first foreign market entered.

**The Dependent Variables**

The dependent variables are calculated from the source data as follows: First Country Count (FirstCC<sub>j</sub>), Second Country Count (SecondCC<sub>j</sub>) and Third Country Count (ThirdCC<sub>j</sub>) correspond to the sum of entries to country <i>j</i> as the first, second, and third foreign market of a firm.

The summary variable for the attractiveness of a country was derived from the mean establishment rank (MR<sub>j</sub>) of country <i>j</i> (Nordström and Vahlne 1994) based on the three first entry choices given by the firms to the country and using entry order four for the other countries. This produces the smallest rank for Sweden (1.62) but values between 3 and 4 for most of the countries. This highly skewed dependent variable is corrected with a linear logarithmic transformation producing the main dependent variable Country Preference of country <i>j</i> (CP<sub>j</sub>) with higher values for attractive countries and lower values for others:

\[
(1) \quad CP_j = \ln[4 - MR_j]
\]

**The Independent Variables**
Cultural Distance

The first independent variable is based on the composite index of Kogut and Singh (1988) and Hofstede’s (1980, 2001) four cultural dimensions (Power Distance, Uncertainty Avoidance, Individualism, and Masculinity). We first calculated the cultural distance score of Finland (f) for each of Hofstede’s (2001) four cultural dimensions (i), namely uncertainty avoidance, individualism, masculinity, and power distance. Next we calculated a composite score Cultural Distance (CD) for each country j as an average of the four factors, each factor normalized with the variance of the dimension i (Vi) with respect to distance to Finland. Algebraically:

\[
CD_j = \sum_{i=1}^{4} \frac{(I_{ij} - I_{if})^2}{4 * V_i}
\]

The shortcomings of Hofstede’s (2001) data concerning the 26 target countries was mitigated by using generic Arabic values for Arab Emirates and using the average of Poland, Russia, and Estonia for Latvia and Lithuania, based on the distribution of nationalities within the population (The World Factbook 2006) of the latter two countries. The three countries within the lowest cultural distance were Estonia, Norway and the Netherlands while the three most distant ones were China, Singapore and Japan. To take into account Shenkar’s (2001) recommendations, we tested also Hofstede’s fifth dimension ‘Long Term Orientation’ by adding it to the composite index of Kogut and
Singh (1988). In addition, we also tested the all five cultural dimensions individually to justify the use of the aggregate index. In both cases the results were very close to the ones presented here and provide little added value\(^2\). Despite of criticism toward operationalization of Hofstede’s (1980; 2001) cultural dimensions and composite index of Kogut and Singh (1988), the usage of this method gives us a possibility to compare and test our findings to earlier studies using the same method.

**Geographical Distance**

Geographical distances between the target country and Finland were calculated based on the geographical coordinates of the countries found in The World Factbook (2006). The geographical coordinates include latitude and longitude figures and determine the geographical center of the country. The final independent variable, Geographical Distance (LnGDist\(_j\)), was produced using a logarithmic transformation as follows:

\[
(3) \quad \text{LnGDist}_j = \ln \text{[distance in thousands of kilometers from country } j \text{ to Finland]}
\]

The geographically closest country to Finland, Estonia, got the value -.59. The highest value was 2.75 for Australia. Because delivery of software products does not follow physical routes, the geographical coordinates gave a neutral way to measure the distance from Finland to the target country.
Software Market Size

The software market size of country j was used as the third independent variable. The source data was mainly based on the statistics of European Information Technology Observatory (EITO) from the year 2005 (EITO 2006). A logarithmic transformation was applied for the highly skewed values to produce the independent variable Software Market Size (LnSWMSIZE_j) as follows:

\[
(4) \quad \text{LnSWMSIZE}_j = \ln [1000 \times \text{Country j software product market share}]
\]

The U.S. market being 45%, gained the highest value of 6.11, while the smallest software markets, in Baltic countries, got values around -1.5. For countries not listed by EITO, the software market size estimates were collected from various data sources (IDC 2006, Yates 2003).

Gross Domestic Product and GDP per Capita

Gross Domestic Product figures and the values for GDP per capita were obtained from The World Factbook (2006). A logarithmic transformation was used for correcting the unbalanced distribution of national GDP to produce the independent variable Gross Domestic Product (LnGDP_j), while the more balanced distribution of source data was used as such for the independent variable GDP per Capita (GDPperCapita_j):
(5) \[ \text{LnGDP}_j = \ln [\text{Country } j \text{ cross domestic product in billions of USD}] \]

(6) \[ \text{GDPperCapita}_j = \text{Country } j \text{ GDP per capita in thousands of USD} \]

RESULTS

Analysis Strategy

Impact of cultural distance, geographical distance, and market size to the first three market entries was analyzed using the bivariate correlation analysis and linear stepwise multivariable regression analysis. Bivariate correlations were first observed to verify the ability of single independent variables to explain variation of the dependent variables. Bivariate correlations between the dependent variables and t-tests were used to observe any possible change of countries which were targeted as the first, second and third market entry.

Stepwise creation of linear multivariable models for each independent variable was used to capture both the attraction effect of market size and the negative effects of cultural and geographical distances. As we targeted on analyzing the shift in factor importance from first to third country selection, we treated the independent variables separately instead of analyzing the moderation effect of market size on the negative factors (see e.g.
Rothaermel et al. 2006). The models created were evaluated based on their ability to explain the variation of target market selection in the target group (R² and adjusted R² values), statistical significance of the model and F value (ratio of means square of regression and residual). Collinearity effects of alternative market size indicators were controlled by stepwise creation of regression models, along which the drop of F value was monitored. Only VIF values below 4 were accepted (max 4.016 in Models 4 and 5 in Table 2A). The Durbin-Watson test was used for ensuring non-autocorrelation of residuals. The statistically significant and close to significant standard coefficients (beta) were identified from each model passing the collinearity check. Bivariate correlations between independent variables were later on used as possible explanations for statistically close to significant standard coefficients.

**Bivariate Correlations**

Table 1A represents some descriptive statistics and bivariate Pearson correlations of the variables. All four dependent variables (1-4 in Table 1A) correlate positively at a statistically significant level (p < .01) as expected, with the exception of the First Country Count and Third Country Count, which do not show a statistically significant correlation.

Among the independent variables (5-9 in Table 1A), there are two statistically significant positive correlations with Gross Domestic Product; these are Geographical Distance and Software Market Size. At the level of p < .05 there are also positive bivariate correlations
between Geographical Distance and Cultural Distance, Geographical Distance and Software Market Size, as well as between Gross Domestic Product and GDP per Capita.

Two independent variables have statistically significant negative correlations with Country Preference, namely Geographical Distance and Cultural Distance. Other independent variables have statistically insignificant correlations with Country Preference, although Software Market Size is close (p = .051). Neither of the distances have significant correlations with the First, Second and Third Country Count except Cultural Distance and Second Country Count with a level of p = .017. Instead, Software Market Size has a positive correlation with the Second Country Count (p = .008) and GDP per Capita with Third Country Count (p = .009). In addition, Second and Third Country Count correlate on a lower level (p = .031 and p = .03) with GDP per Capita and Software Market Size, respectively.

[Add Table 1A round here.]

**Regression Analysis**

The analysis is elaborated for each dependent variable using stepwise multivariable linear regression analysis. Table 2A displays alternative models for Country Preference and Table 3A presents alternative models for First, Second and Third Country Selection. Figure 1 summarizes the correlations found in the regression analysis (directed arrows) and presents the bivariate correlations between independent variables and between
dependent variables listed in Table 1A. The correlations of Country Selections and Country Preference are omitted to avoid unnecessary complexity. The five independent variables are presented in the middle. Arrows to Country Preference (on the left) indicate the statistically very significant correlations (***) of Geographic Distance (with negative impact indicated with “-“) and Software Market Size as well as the significant negative impact of Cultural Distance (cf. Table 2A). Arrows to the First, Second and Third Country count indicate the impact of each independent variable to each of the first three country selections separately (cf. Table 3A). Asterisk labels on the arrows indicate significance of the evidence. Altogether, the regression models were able to explain over 73% of the variation in Country Preference of the sample (explaining 70% of the target population) and about 27% (20%), 56% (50%) and 44% (34%) of the first, second and third country selections of the sample (and target population), respectively.

Figure 1.
Correlations of the variables impacting market selection of SMEs in the software industry.
Hypothesis Evaluation

The hypothesis stated earlier can now be tested based on the results available. In Model 1 of Table 2A the negative correlation of Cultural Distance alone explains nearly 30% of the variation in Country Preference of the sample ($R^2 = 29.4\%$, equaling to 26.4\% for the target population, Adj. $R^2$). In Model 2, the combination of Cultural Distance and Software Market Size explain 46\% (41.3\%) of the Country Preference of the sample (and population, respectively). Based on this, we accept Hypothesis 1a; cultural distance is inversely related to target country selection of SMEs.

However, adding Geographical Distance to Model 2 in Table 2A produces Model 3, where the negative impact of Cultural Distance is no more at a statistically significant level, while the negative correlation of Geographical Distance and positive impact of Software Market Size are on a high level of confidence. Model 3 explains 70\% of Country Preference of the target population. Further, if we exclude Cultural Distance, Geographical Distance and Software Market Size in Model 6 still explain almost 70\% of Country Preference with a higher F value. This implies that Cultural Distance adds little explaining power and the four dimensions (Power Distance, Uncertainty Avoidance, Individualism, and Masculinity) embedded in Cultural Distance have some similarities with Geographical Distance, as can be seen also in the correlation of these variables in Table 1A. We can also confirm Hypothesis 2a, stating that geographical distance is
inversely related to the country selection of SMEs in the software industry and Hypothesis 3c, proposing that the large size of the software market in the target country is positively related to market selection of SMEs (discussed in detail later on). Geographical Distance and Software Market Size together also explain 20.6% of the First Country Choices (Model 2 in Table 3A).

In Table 1A, we can see that the correlation of Cultural Distance with First Country Count is weaker than its correlation with Second Country Count and Third Country Count. Also, the mean of Cultural Distance of the first countries entered (M = 1.011, SD = .518) is even higher than the cultural distance of the second countries entered (M = .979, SD = .457). Notably all these means are below the mean of Cultural Distance of the 26 countries (Table 1A). Based on this, Hypothesis 1b is strongly rejected, i.e. in this population the first country selected is not culturally closer than the later countries.

From Table 1A, we can see that the negative correlation of Geographical Distance is highest with First Country Count, although the correlations with Second and Third Country Count are also negative. Geographical Distance and Software Market Size together also explain best the first country choices of the target population (Model 2 in Table 3A), while model containing Cultural Distance and Software Market Size outperforms this combination in explaining the Second Country Count and we found no statistically significant correlations of Geographical Distance with the Third Country Count in any of the models. Also the independent samples’ t-test for average Geographical Distances between first countries entered (M = .85, N = 51) vs. second or
third countries entered (M = 1.18, N = 72) shows the difference to have some significance (at level of .057, just below *). Thus, there exists some evidence that the first country selected is geographically closer than later markets (Hypothesis 2b).

Out of the five independent variables, we find no statistically significant support for GDP’s correlation with Country Preference (Hypothesis 3a) or any of the three country selections. No statistically significant correlation between GDP per Capita and Country Preference was found either. However, as a single variable, GDP per Capita explains 22.2% of the Third Country Selection (Model 9, Table 3A), which gives partial support for Hypothesis 3b with respect to Third Country Selection, i.e. affluence of the target country (GDP per capita) is positively related to third country selection of SMEs.

The strong negative impact of Geographic Distance and the positive impact of Software Market Size explains Country Preference well, as explained earlier (Models 3 and 7 in Table 2A). In addition, Software Market Size correlates positively with all Country Counts (Models 2, 4-8, and 11 in Table 3A). Thus, we confirm Hypothesis 3c: The large size of the software market in the target country is positively related to country selection of SMEs.

As noticed earlier, we did not find any significant correlation between large market size (GDP) and any of the dependent variables. Therefore, we must reject both Hypothesis 3b and 3e. Instead, we found a strong correlation between large Software Market Size and Second Country Count (Table 3A, Model 6 explaining over 50%). We also found a
correlation between GDP per Capita and the Third Country Selection (Model 9, Table 3A).

**CONCLUSION**

By analyzing objective empirical data, this study targeted on finding the shift in priorities in SMEs’ country selection from cultural and geographical distance to market size. The empirical results show that it is possible to choose an industry segment and a country of origin from which all the three major factors impacting target country selection are statistically observable. In this context it is also possible to identify the shift of emphasis from the distance factors to the market size factors in the first three target country selections.

Based on the statistical results, geographical distance and software market size in the target country best explain country preference, up to 70% in the target population. Thus, SMEs in the software industry seem to favor countries within a low geographical distance and which have with large software markets. A more detailed investigation of the source data also highlights the fact that countries with small software markets tend to be selected only if they are located nearby. Interestingly, the large size of the software market in the target country was the most important determinant for country preference, instead of GDP (Terpstra and Yu 1988) or GDP per capita (Clark and Pugh 2001; Rothenberg et al. 2006). This finding adds the important factor ‘vertical market size’ to the current literature suggesting that it, instead of commonly used variables GDP and GDP per capita,
gives the best explanation for country preference. The strong correlation of GDP with software market size and geographical distance suggests that it may have served as a proxy approximating these two primary variables in some of the previous studies. Our findings also suggest that cultural distance between the home and host country can only partly explain country preference. Cultural distance is a reasonable single variable approximation for this population. Nevertheless, it is outperformed by geographical distance and the software market size in a target country. Thus, we do not find support to studies (Jones and Teegen 2001; Rothaermel et al. 2006) proposing that cultural distance is an important determinant in target country selection.

In the detailed investigation, we tested impacts of cultural distance, geographical distance, and the three market size variables in the market priority among the first three countries entered. Findings in this study indicate that geographical distance has the most impact on the first country choice, as suggested in previous studies (Chetty 1999; Clark and Pugh 2001; Dow 2000; Srivastava and Green 1986). This result also confirms earlier qualitative findings related to market selection of software firms (Bell 1995; Coviello and Munro 1997; Moen et al. 2004) suggesting that software firms enter first to countries within a low geographical distance. The reason for this might be limited market knowledge, low financial and human resources, and narrow set of network relationships of SMEs in the software industry when they make the first market entry (Coviello and Munro 1997). In our sample, the large size of the software market was found to be the most important factor for the second country selection. In addition, low cultural distance correlated with the second country rather than with the first country. Based on this, the
criticism for using Hofstede’s cultural dimensions (Dow 2000; Dow and Karunaratna 2006; Shenkar 2001; Tihanyi et al. 2005) seems to be justified. In third country selection, GDP per capita in the target country was the most important single determinant while also the software market size correlated with third country selection.

In summary, the findings indicate that entry priority of small and medium-sized software firms shifts from geographically close countries to countries with high purchasing power and greater geographical distance. The reason for this might lie on increasing network relationships which increase competitive capabilities and knowledge of foreign markets (Coviello 2006; Coviello and Munro 1995, 1997). Findings of this study complement earlier qualitative findings related to small and medium-sized software firms (Bell 1995; Coviello and Munro 1997; Moen et al. 2004) adding the fact that, after the first country has been selected, firms’ focus move to countries with large market size for their products. In addition, these empirical findings validate the conceptual framework of Bell et al. (2003) suggesting that knowledge-intensive firms start to seek market opportunities in leading markets in a very early phase of their internationalization process. Prior to this study, these two findings have not been combined as a coherent model quantifying the change in market priority from the first country (geographically close) to second and third countries (leading markets). Although this study uses only macro-level variables, it seems to indicate that increasing market knowledge, resources, and network connections highlighted in previous qualitative (Coviello 2006; Coviello and Munro 1997) and conceptual studies (Bell et al. 2003; Jones and Coviello 2005; Oviatt and McDougall 1994, 2005) are the factors behind these macro-level changes. These factors enable firms’
to move in a very early phase in their internationalization process from geographically close countries to countries with high market potential.

**Managerial implications**

Software industry can be characterized as a highly global market, where high-technology SMEs try to rapidly acquire market leadership or major market shares in opening new market segments. These firms, cited as ‘born global’ or ‘international new ventures’, have a strong international focus, while they often lack sufficient network relationships, knowledge, resources or time (Bell 1995, 1997; Coviello 2006; Coviello and Munro 1997; Jones and Coviello 2005; Oviatt and McDougall 2005, among others) needed for well organized internationalization based on detailed data of the new, emerging global market opportunities. In this situation, the two macro-level variables to look for are the software market size and the geographical distance of the potential target country. Based on our results, the managers strongly prefer countries with large software product market over countries with large GDP and high GDP per capita when choosing the first countries to enter.

Based on some earlier studies (Bell 1995; Coviello and Munro 1995, 1997; Moen et al. 2004), entering first to nearby countries helps also develop network relationship and foreign market knowledge, as well as develop sufficient resources to enter countries within a greater geographical distance and larger market size. Operating first in a nearby country facilitates organizing critical customer support activities, such as requirement
specification, installation, and after-sales support, in the target country. Conducting these important activities in a sales process of software products first in a nearby country is easier due to lower operational cost and environmental familiarity. When this has been done once, the same operation model can be cost-efficiently copied and complemented in other countries. Majority of the firms seem to follow this approach. If no better information is available, managers can use the ratio of a country’s software product market share percentage per the distance to the country (in thousands kilometers) as the first approximation for initial market selection. Internationalization is often started from a nearby country, where the ratio is over 0.1, while the preferred second and third countries have ratios close to or over 1.

Limitation and Further Research

In this study, we focused only to factual macro-level variables without investing firm specific variables such as a firm size, existence of client, or experience or nationality of managers. Further quantitative research is needed to examine the influence of these firm specific variables to the target country selection. The results are also limited by fact that this study was based on only SMEs in a single industry. Although this gave us a possibility to complement earlier case findings related to software firms, the results can be generalized for other industries only to a certain extent. Further research is needed to test the generalizability of these findings in other vertical markets. For instance, the size of the software market in the target country was found to be the most important determinant for country preference. Thus, further research should investigate how well
vertical market size determines the country preference in other industries. Although we can assume that the influence of variables used in this study can be generalized to firms originating from other countries, their occurrence might vary depending on the country of origin. Thus, more comparative studies are needed to verify the results with firms originating from other countries. In this study we focused on the three main factors of country selection with variants of the market size metrics, using step-wise regression analysis. The resulting models explained 70% of the variation in the target population and about 74% of the sample ($R^2$ values). Higher explanation rates can be achieved by adding other independent variables (e.g. country risk factors) in further studies, while the risk with multicolinearity problems rises significantly along rising $R^2$ values. Studies in similar context with multi-dimensional indicators of psychic distance would also add to the literature as only four Hofstede’s (1980, 2001) cultural dimensions were operationalized in this study. The temporal dimension was also excluded from this study. We did not consider the pace at which firms entered foreign markets or whether this had some relation with market choice. Further studies for investigating the timeframes of entrepreneurial SMEs’ internationalization processes could be useful also for validating earlier case findings (e.g., Coviello and Munro 1997) and the conceptual model of Jones and Coviello (2005). The entry modes that firms used in their target countries were also excluded from this study. Thus, more investigations are needed to examine influences of cultural distance, geographical distance, and market size on software firms’ entry mode selection. Finally, further research should identify different internationalization approaches, such as ‘nearby countries’ first or ‘large markets’ first approaches. Special emphasis should be paid on the impact of the approaches on the performance of firms in
their internationalization processes to find out how successful firms enter international markets.

Notes

1. In this study, the aim is to use objective and commonly applied indicators to which data is available. In addition, there seem to be a lack of commonly accepted factors influencing the psychic distance. Thus, due to data constraints, instead of using a single or multiple dimensions of psychic distance in the internationalization framework, a single dimension of cultural distance consisting of four factors is used there.

2. In this study, we used Hofstede’s (1980, 2001) four cultural dimensions (Power Distance, Uncertainty Avoidance, Individualism, and Masculinity) and the composite index of Kogut and Singh (1988). In line with Shenkar’s (2001) suggestion, we ran the regression tests also with an indicator, which adds Hofstede’s fifth dimension ‘Long Term Orientation’ to the indicator. In general, the results were rather similar to the ones reported here. The five-dimensional composite cultural distance was still the best single variable for explaining Market Preference of the population (23.1%) although not as good as the four-dimensional variant reported here (26.4%). A similar slight drop of explanation power was visible in all the models of Table 2A when the five-dimensional variant was used. However, in Model 4 it increased the model’s capability to explain 74.5% of Market Preference in the population. With the five-dimensional variable, average Cultural Distance to first countries entered was still higher than the average to
the second countries entered (.920 and .868, respectively). Due to the slightly lower explaining power of the five-dimensional variable, the combination of Geographic Distance and Software Market Size now outperforms the combination of five-dimensional Cultural Distance and Software Market Size in explaining Second Country Count in the target set (48.6% vs. 47.0%). Altogether, the slightly poorer performance of the five-dimensional variant did not justify giving up the ability to compare the results of this research with previous literature.

REFERENCES


### Table 1A
Descriptive Statistics and Bivariate Correlation Matrix

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. First Country Count</td>
<td>1.96</td>
<td>3.42</td>
<td>0</td>
<td>17</td>
<td>1.54</td>
<td>1.92</td>
<td>0</td>
<td>7</td>
<td>.514</td>
<td>6.415</td>
<td>.493</td>
<td>3.520</td>
</tr>
<tr>
<td>2. Second Country Count</td>
<td>1.54</td>
<td>1.92</td>
<td>0</td>
<td>7</td>
<td>1.23</td>
<td>1.45</td>
<td>0</td>
<td>5</td>
<td>.179</td>
<td>.599**</td>
<td>.493</td>
<td>.380</td>
</tr>
<tr>
<td>3. Third Country Count</td>
<td>1.23</td>
<td>1.45</td>
<td>0</td>
<td>5</td>
<td>-1.53</td>
<td>1.09</td>
<td>-3.26</td>
<td>.87</td>
<td>.712**</td>
<td>.803**</td>
<td>.520**</td>
<td>.465*</td>
</tr>
<tr>
<td>4. Country Preference</td>
<td>1.51</td>
<td>0.95</td>
<td>.04</td>
<td>3.44</td>
<td>-.313</td>
<td>-.675</td>
<td>-.512**</td>
<td>.493*</td>
<td>.512**</td>
<td>.465*</td>
<td>.380</td>
<td>-.542**</td>
</tr>
<tr>
<td>5. Cultural Distance</td>
<td>1.42</td>
<td>0.99</td>
<td>-.59</td>
<td>2.75</td>
<td>-.345</td>
<td>-.191</td>
<td>.509**</td>
<td>.426*</td>
<td>.387</td>
<td>.037</td>
<td>.415*</td>
<td>.620**</td>
</tr>
<tr>
<td>6. Geographical Distance</td>
<td>1.92</td>
<td>1.90</td>
<td>-1.69</td>
<td>6.11</td>
<td>.209</td>
<td>.627</td>
<td>3.10</td>
<td>9.42</td>
<td>.018</td>
<td>.218</td>
<td>.121</td>
<td>.117</td>
</tr>
<tr>
<td>7. Software Market Size</td>
<td>24.76</td>
<td>12.36</td>
<td>3.30</td>
<td>43.40</td>
<td>.136</td>
<td>.424*</td>
<td>.503**</td>
<td>.272</td>
<td>-.288</td>
<td>-.003</td>
<td>.428*</td>
<td>-.091</td>
</tr>
</tbody>
</table>

Notes: N = 51  
** statistically significant p < .01, * statistically significant at the level p < .05 (2-tailed)
Table 2A
Regression Models for Country Preference

<table>
<thead>
<tr>
<th>Model</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2</td>
<td>.294</td>
<td>.46</td>
<td>.736</td>
<td>.751</td>
<td>.751</td>
<td>.697</td>
</tr>
<tr>
<td>Adj. R2</td>
<td>.264</td>
<td>.413</td>
<td>.700</td>
<td>.704</td>
<td>.689</td>
<td>.671</td>
</tr>
<tr>
<td>Significance</td>
<td>.004</td>
<td>.001</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
</tbody>
</table>

**Standardized coefficients**

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cultural Distance</td>
<td>-.542**</td>
<td>-.557**</td>
<td>-.232</td>
<td>-.264</td>
<td>-.265</td>
<td></td>
</tr>
<tr>
<td>Geographical Distance</td>
<td>-.679***</td>
<td>-.747***</td>
<td>-.745***</td>
<td>-.813***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software Market Size</td>
<td>.408*</td>
<td>.678***</td>
<td>.507*</td>
<td>.514</td>
<td>.725***</td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>.250</td>
<td>.243</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP per Capita</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.005</td>
<td></td>
</tr>
</tbody>
</table>

The models 1-3 are produced with stepwise addition of independent variables to explain Country Preference (Formula 1). Model 5 uses all the independent variables as defined by Formulas 2-6. Model 4 uses other variables except GDP per Capita. Model 6 uses only two independent logarithmic variables, Geographical Distance and Software Market Size.