ETLA
ELINKEINOELÄMÄN TUTKIMUSLAITOS
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
Lönnrotinkatu 4 B 00120 Helsinki Finland Tel. 358-9-609 900
Telefax 358-9-601 753 World Wide Web: http://www.etla.fi/

Keskusteluaiheita – Discussion papers

No. 1063

Aija Leiponen* – Constance E. Helfat**

WHEN DOES DISTRIBUTED INNOVATION ACTIVITY MAKE SENSE? LOCATION, DECENTRALIZATION, AND INNOVATION SUCCESS

* Department of Applied Economics and Management, 251 Warren Hall, Cornell University, Ithaca, NY 14853-7801

** Tuck School of Business at Dartmouth, 100 Tuck Hall, Hanover, NH 03755

This paper benefited from the comments of Janet Berkowitz, Bruno Cassiman, Maryann Feldman, Jasjit Singh, Harbir Singh, and presentations at Cornell University, an Academy of Management PDW, the Harvard Strategy Conference, the Roundtable for Engineering Entrepreneurship Conference, and the University of Maryland. Research support from ETLA (The Research Institute of the Finnish Economy) is gratefully acknowledged. This paper has been prepared as part of the BRIE-ETLA collaboration (www.brie-etla.org). The usual caveat applies.

ISSN 0781-6847 20.12.2006

LEIPONEN, Aija **– HELFAT,** Constance E., **WHEN DOES DISTRIBUTED INNOVATION ACTIVITY MAKE SENSE? LOCATION, DECENTRALIZATION, AND INNOVATION SUCCESS.** Helsinki: ETLA, Elinkeinoelämän Tutkimuslaitos, The Research Institute of the Finnish Economy, 2006, 32 p. (Keskusteluaiheita, Discussion Papers ISSN 0781-6847; no. 1063).

ABSTRACT: Companies face an expanding set of choices about where to locate their innovation activity, both within their home countries and abroad. This location choice also requires firms to make a simultaneous choice about the organizational structure of innovation activity: almost by definition, multiple locations per firm imply some degree of decentralization. Using firm-level data on innovation output and the location of research and development (R&D) activity, we shed new light on the question of whether firms that have multiple locations also have greater innovation success. Our results indicate that, on average, having distributed R&D activity is beneficial in terms of the extent and breadth of innovation success, and the effect is strongly related to the knowledge sourcing strategies that firms employ. These results are consistent with the interpretation that R&D location decisions are driven by the desire of firms to access a broad set of external sources of knowledge for innovation activities. We also find that the benefits of multiple R&D locations do not apply to novel (new-to-the-market) innovations. Our results suggest that when analyzing technological innovation, it is important to distinguish between novel and imitative innovations, since their determinants may differ.

KEY WORDS R&D, innovation, organizational decentralization, knowledge sourcing
JEL codes O32, L22

LEIPONEN, Aija **– HELFAT,** Constance E., **HAJAUTA VAI HALLITSE? INNOVAATIOTOIMINNAN SIJAINTI, ORGANISAATIO JA TUOTOS.** Helsinki: ETLA, Elinkeinoelämän Tutkimuslaitos, The Research Institute of the Finnish Economy, 2006, 32 s. (Keskusteluaiheita, Discussion Papers ISSN 0781-6847; no. 1063).

TIIVISTELMÄ: Suomalaisyritysten tutkimus- ja kehitystoiminta (T&K) on laajentunut maantieteellisesti viime vuosina, sekä ulkomaille että kotimaan sisällä. T&K -toiminnan hajauttamiseen liittyy hyötyjä ja kustannuksia – aiemmassa kirjallisuudessa on enimmäkseen korostettu hyötyjä. Tässä tutkimuksessa arvioidaan molempia organisaatioteorian ja suomalaisen yritysaineiston valossa. T&K:n hajauttamisen hyödyt liittyvät siihen, että yritykset pääsevät paikallisten hiljaisen tiedon lähteiden äärelle. Esimerkiksi yliopistojen tieteelliseen tutkimukseen ja markkinoiden ominaisuuksiin liittyy tietoa, jota on vaikea saada etäältä. T&K:n hajauttamisen kustannukset taas syntyvät tiedon ja osaamisen siirtämisestä eri T&K -toimipaikkojen välillä. Toimipaikkojen välisten synergioiden hyödyntäminen on vaikeaa, jos ne ovat hyvin etäällä toisistaan. Tutkimuksen empiiristen tulosten mukaan kotimaassa hajautetut T&K -organisaatiot lanseeraavat keskitettyjä organisaatioita todennäköisemmin ainakin yhden innovaation ja tuottavat enemmän innovaatioihin perustuvaa myyntiä. Hajautettu T&K -organisaatio on myös hyvin vahvasti korreloitunut yritysten ulkopuolisten tietolähteiden käytön kanssa. Käytetyssä edustavassa suomalaisten teollisuusyritysten aineistossa hajautetun innovaatiotoiminnan hyödyt eivät kuitenkaan näyttäydy markkinoille uusien (eikä ainoastaan yritykselle itselleen uusien) tuotteiden kehittämisessä. Hajautetun innovaatiotoiminnan tuotokset ovat siis useimmiten vähittäisiä ja soveltavia.

AVAINSANAT T&K, innovaatiot, organisaatio, hajauttaminen, tietolähteet

1. INTRODUCTION

Technological innovation has become ever more important to business success. No industry is immune. Technological innovation in information technology has revolutionized the retailing industry. Technological innovation in oil exploration and production has vastly increased accessible reserves. Technological advance affects all industries and firms, not just those considered hi-tech. At the same time as technological innovation has become pervasive, companies face an expanding set of choices about where to locate their innovation activity, both within their home countries and abroad. This location choice requires firms to make a simultaneous choice about the organizational structure of innovation activity: almost by definition, multiple locations per firm imply some degree of decentralization.

Until recently, the questions of whether to conduct innovation activity in multiple locations and whether to decentralize technological innovation within the firm remained largely separate. Researchers generally have viewed the location question through the lenses of related literatures on foreign direct investment and the knowledge-based view of the firm. From this perspective, multiple locations of innovation activity are thought to improve the ability of firms to innovate. Recently, researchers have begun to bring organizational issues into the analysis by noting that having multiple locations for R&D (also termed distributed R&D) entails organizational costs (e.g., Chacar and Lieberman, 2003; Singh, 2006; Zhao, 2006). We take this research a step further and directly compare several implications of the literature on organizational form and incentives with implications of the knowledge-based view. The predictions of these two literatures regarding the effect of multiple R&D locations on innovation outcomes sometimes conflict and other times are congruent. Using uniquely detailed firm-level data on innovation output and the location of R&D activity, we shed new light on whether firms that have multiple R&D locations also have greater innovation success. We do not investigate other aspects of organizational form, such as which units have budgetary control over R&D decisions. Although our data contain detailed information about R&D locations within firms, we do not have additional information regarding the structure of command and control inside organizations.

Our analysis makes several contributions. First, we bring together two relatively separate literatures in order to examine the impact of multiple R&D locations on innovation output. Second, we examine R&D location choices within a single country, and find that although some of the theory was developed for an international setting, the theoretical predictions apply within countries as well. Third, we compare the effect of location choice on novel versus non-novel innovation output. Although research on innovation often focuses on patents, which reflect novel inventions, a great deal of business R&D is directed toward imitative innovation. Fourth, we can overcome some well-known problems of using patent data. Patents are meaningful measures only of intermediate innovation output and in a relatively small number of industries (Griliches, 1990). In contrast, our data include sales of commercialized innovations, and the sample is close to representative of the manufacturing sector as a whole. Finally, we test whether any positive impact on innovation output of distributed R&D reflects an accompanying breadth of search for new knowledge, as the knowledge-based view predicts.

In what follows, we first compare and contrast the relevant literatures and develop testable hypotheses. Then we describe our data, empirical methodology, and results.

2. LOCATION AND INNOVATION ACTIVITY

There are two distinct perspectives regarding the location of innovation activity. One general perspective emphasizes the benefits of distributed R&D, and derives primarily from the literatures on foreign direct investment and the knowledge-based view of the firm. The other general perspective emphasizes the costs of decentralization of innovation activity, and derives in part from the economics of transactions costs and agency. To date, research on R&D location choice, while acknowledging the possibility of organizational costs, often emphasizes the benefits. Here we focus on the *net benefits* of distributed R&D and ask under what conditions we might expect the benefits net of costs to be positive or negative. We first examine the implications of the two literatures separately, since they differ in important ways. Then we develop hypotheses that compare and contrast the implications of these literatures.

2.1 Benefits of Multiple Locations

The question of whether firms should utilize multiple locations for technological innovation activity is often framed as one of foreign direct investment (Kuemmerle, 1999; Penner-Hahn and Shaver, 2005). This framing derives from the literature on how firms should best access foreign markets, whether by foreign direct investment or by other means such as exporting. We briefly highlight key aspects of this very large literature relevant to our analysis here.

Early literature argued that firms should use foreign direct investment (FDI) when they sought to exploit idiosyncratic assets in foreign markets. Following Hymer (1970), Buckley and Casson (1976), and Dunning (1981), scholars of the “internalization school,” emphasized the advantages of internalizing economic activity when firms sought to employ special advantages abroad, such as through technology transfer (Teece, 1977). Subsequently, scholars argued that in addition to exploiting special advantages such as technological knowledge, firms often seek to acquire technological knowledge abroad (Kogut, 1991; Chung and Alcacer, 2002), including through R&D (Kuemmerle, 1999). Kuemmerle termed these two motives “home base exploiting FDI” and “home base augmenting FDI,” respectively.

The “internalization” view aligns with the predictions of transaction cost theory, which states that firms should use internal organization rather than markets when transactions run a large risk of ex post opportunism. Since technology transfer, and transfer of proprietary knowledge more generally, has a high risk of ex post opportunism, transaction cost theory implies that firms should use internal organization to transfer proprietary technology abroad (Teece, 1980).

Kogut and Zander (1993) subsequently argued that other factors could explain technology transfer through foreign direct investment, independent of internalization and transactions cost arguments. They noted that a great deal of technological knowledge is tacit or not well codified, and that even the use of codified knowledge requires an understanding of the context in which the knowledge is employed. Under these circumstances, firms are superior to markets in transferring technology and information because organizations consist of communities of practice that are more likely to have shared language and understanding. This in turn greatly reduces the costs of transferring information internally relative to using the market. This general argument regarding the superiority of firms as a means of knowledge transfer is one of the central tenets of what has come to be known as the “knowledge-based view of the firm” (Kogut and Zander, 1992; Grant, 1996).

The knowledge-based view has clear implications for the location of innovation activity. Particularly when firms seek to acquire knowledge specific to other locales, they may need to establish facilities in those locales (Kogut, 1991). Since much technological knowledge is tacit, its transfer requires frequent interactions between the sender and receiver—which proximity facilitates (Kogut and Zander, 1993). For example, if researchers in a particular locale possess relevant tacit knowledge and are unwilling or unable to relocate, or if research activity stands to benefit from local knowledge sharing among a community of researchers, firms may locate their R&D where the researchers reside. Moreover, specialized local knowledge useful in innovation can come from many different sources, including universities, research institutes, suppliers, customers, and competitors. Studies have documented that firms that conduct R&D outside of their home countries or that apply for patents abroad through foreign subsidiaries are better positioned to access and build upon technological expertise located abroad (e.g., Almeida, 1996; Florida, 1997; Kuemmerle, 1999; Frost, 2001). A firm therefore may choose to have multiple locations for its innovation activity (Kuemmerle, 1999; Tripsas, 1997). This reasoning further implies that distributed R&D will lead to greater innovation success because the firm can access a larger number of different sources of technological knowledge than it could in a single location.

The internalization, transactions cost, and knowledge-based view arguments all contain the same implication: when firms seek to transfer or access technology abroad, including through R&D, there are clear advantages to doing so internally by setting up R&D units in multiple locations. Firms that use multiple locations for purposes of “home base augmenting” R&D will be able to access a greater number of knowledge sources than in a single R&D location. This in turn improves the likelihood of innovation success. Moreover, if having multiple locations yields a wider range of knowledge on which to build and recombine, the resulting set of innovations should span a wider range of applications as well (e.g., product and process innovations, rather than just one or the other). These arguments, although developed largely in the context of foreign expansion of firm activity, also apply to locations within a single country if these locations have types of knowledge relevant to innovation that are difficult to obtain from afar (Kuemmerle, 1999; Furman et. al., 2005). The theoretical arguments regarding physical location and knowledge sources do not require country boundaries. The foregoing literatures apply to all location choices, whether foreign or domestic.

2.2 Costs of Multiple Locations

The use of multiple locations for innovation activity necessarily brings with it some degree of organizational decentralization. As a practical matter, a firm cannot operate organizational units from a distance without delegating at least some responsibility for day-to-day operations to its physically distant units. Headquarters may or may not choose to delegate longer-term operational or strategic decisions to organizational units in other locations. But even if headquarters delegates no long-term decisions at all, the firm will need an operational manager on-site in its distant units to carry out the decisions of headquarters. Moreover, research and development requires some independence on the part of the individuals doing the research in order for the individuals to make progress creating or implementing new products and processes. Distance is likely to make it even more difficult to monitor the activity of researchers who already have some independence regardless of location. In short, firms cannot avoid some amount of decentralized authority when they utilize multiple locations for innovation. Therefore, the choice of organizational form and associated incentives to elicit innovation has important implications for the location of innovation activity. Since the literature on organizational form and incentives is very large, we abstract a few key points for our analysis.

Many large, diversified firms utilize a decentralized structure organized according to product-markets, in contrast to earlier structures that organized tasks centrally according to function (Chandler, 1962). A multi-divisional (M-form) organization minimizes coordination among product lines and enables top managers to focus on strategic issues for the company as a whole (Williamson, 1975). Top management also can more easily obtain information about business unit performance and use these data to reward division managers. This creates higher-powered incentives and reduces agency problems and opportunism within the corporation (Williamson, 1991).

This logic, when applied to the organization of R&D, suggests that decentralization has the advantage that by delegating decision making it minimizes coordination of R&D decisions across divisions (Argryes and Silverman, 2004). In addition, the M-form ensures that the division managers, who are closer to customers and markets and therefore have superior information about those markets (relative to top management), have responsibility for product R&D directed towards those customers and markets (Argyres and Silverman, 2004). The literature on "user-based" innovation (Von Hippel, 1998), whereby companies receive valuable ideas for innovation from their customers, also implies that decentralization of innovation activity may benefit firms by bringing them closer to their customers.[1]

The preceding arguments imply that a decentralized organizational structure reduces coordination and information costs of conducting R&D because it delegates at least some decisions to local or divisional managers. These arguments do not, however, take into account that R&D may be subject to economies of scale and scope (e.g., Henderson and Cockburn, 1994, 1996), which requires information transfer. This leaves firms in a quandary. If they put mechanisms in place to transfer knowledge between disparate organizational units, this may conflict with the incentive efficiencies of the M-form firm. In particular, if firms tie the compensation of division managers to the performance of their divisions, as transactions cost theory suggests, division R&D managers may have little incentive to transfer knowledge across divisions (Kay, 1988). Furthermore, research has documented the large difficulty that even motivated personnel have in effectively transferring knowledge between organizational units (Szulanski, 1996). Knowledge sharing among researchers is more effective when researchers are physically located close to one another (Allen, 1977), notwithstanding the use of information technology today (Cummings, 2003). Then, one alternative is to centralize R&D in order to facilitate knowledge transfer across projects. Centralization also enables firms to provide R&D personnel with appropriate corporate-level rather than division-level incentives to perform research that spans a range of businesses (Kay, 1988; Lerner and Wulf, 2006).

How can we resolve the foregoing arguments regarding decentralization versus centralization of R&D? The answer depends on the type of R&D in question. Closer contact with markets and customers through organizational decentralization is particularly appropriate for more applied, market- and customer-oriented R&D. In contrast, “non-specific,” often fundamental, research that potentially can be applied in multiple businesses is more likely to be subject to economies of scope (Argyres and Silverman, 2004). Firms should centralize such R&D, thus utilizing a single location (Chacar and Lieberman, 2003). Moreover, since more centralized R&D is directed toward innovations with broader applicability, Argyres and Silverman (2004) argue that: 1) firms with more centralized R&D activity will search more broadly outside of their organizational boundaries for information relevant to innovation; 2) the resulting set of innovations will have a wider range of applications as well. By implication,

[1] Brown and Eisenhardt (2000) further argue that a modular decentralized organization enables companies to better adapt to market and technological change through “patching,” involving rapid movements into and out of evolving markets.

firms with more decentralized R&D (including through multiple locations) will search less broadly outside of their organizational boundaries and the resulting innovations will have a narrower range of applications.

2.3 Benefits versus Costs of Multiple Locations of Innovation Activity

Based on the foregoing literatures, we next formulate hypotheses regarding the location of R&D activity. The knowledge-based literature suggests that having multiple R&D locations leads to greater innovation success. The literature on organizational structure and incentives points to the costs of distributed R&D if firms seek to achieve knowledge spillovers across organizational units. A reconciliation of these two viewpoints suggests that although firms may benefit from multiple locations, if firms seek to share knowledge obtained in multiple locations, diminishing returns may set in due to costs of coordination. Moreover, holding the size of the R&D budget constant, a firm is likely to face diminishing returns to the number of locations simply due to the "minimum efficient scale" required for cost effective operation of a stand-alone R&D unit:

H1: Firms that have multiple locations of R&D activity experience greater innovation success than firms that have a single R&D unit, ceteris paribus, but there are diminishing returns to the number of R&D locations.

The knowledge-based view puts forward a specific mechanism through which distributed R&D leads to greater innovation success: multiple locations enable the firm to access a larger number of different knowledge sources outside of the organization than it could in a single location. In contrast, the literature on organizational form and incentives suggests that when firms decentralize R&D, they will search less broadly outside of their organizational boundaries for information relevant to innovation. The following hypothesis reflects the arguments of the knowledge-based view. Rejection of the hypothesis would provide evidence consistent with the organizational arguments.

H2: Any positive association between multiple locations of R&D activity and innovation success reflects access to a larger number of different knowledge sources outside of the organization.

As argued earlier, if the knowledge-based view is correct that distributed R&D provides access to a wider range of knowledge on which to build and recombine, then the resulting innovations should span a wider range of applications as well. This conflicts directly with transactions cost arguments that decentralization of R&D will result in innovations that have a narrower range of applications. We test these two competing arguments against one another, as follows:

H3a: Firms that have multiple R&D locations generate a wider range of innovation output than firms that have a single R&D unit.
H3b: Firms that have multiple R&D locations generate a narrower range of innovation output than firms that have a single R&D unit.

Finally, the literature on organizational form suggests that a decentralized R&D organization will generate less fundamental innovations than a centralized organization. In contrast, the knowledge-based view does not contain a clear-cut prediction regarding the type of innovation output. Instead, firms may use multiple R&D locations in order to gain access to tacit knowledge useful in innovation of many types. The following hypothesis reflects the organizational arguments:

H4: Multiple locations of R&D activity are associated with greater innovation success for non-novel (imitative) innovations than for novel innovations.

3. THE SURVEY DATASET, VARIABLES, AND STATISTICAL INFERENCE

The empirical setting is the manufacturing sector in Finland. Like several other European Union countries, the national statistical agency in Finland conducts surveys on innovation activity. These surveys contain uniquely detailed data on innovation outcomes and R&D locations that enable us to test hypotheses such as those regarding positive but diminishing returns to multiple locations and the role of knowledge sources.

Although Finland is relatively small in surface area (similar to the state of California) and population (5.2 million), it is physically dispersed and diverse. The Helsinki metropolitan area is the most important commercial, industrial, and intellectual center, but the regional centers of Turku and Tampere also provide significant markets and high quality universities. Moreover, there are physically dispersed technological concentrations, such as electronics around Oulu University in the North, the forest sector around Lappeenranta University of Technology in the Southeast, and medical research around University of Turku and Åbo Akademi in the city of Turku. Firms thus may consider expanding R&D activities to locate in major markets, in the vicinity of major universities, or close to peers in technology "hot spots."

The data used to test the hypotheses come from Finnish R&D Surveys, administered every other year, and the Finnish Community Innovation Survey (CIS), administered every four years. Statistics Finland, the national statistical agency, administered all of the surveys. The Community Innovation Survey was coordinated with the statistical agency of the European Union, Eurostat, which sponsors CIS surveys in several member countries. Eurostat coordinated the initial development of the survey instrument and the data collection techniques.

Within the manufacturing sector, all firms with more than 100 employees received the R&D Survey, as did a stratified random sample of firms with 10-99 employees. The sample for the smallest firms with fewer than 10 employees is not random; it includes only firms known to perform R&D, based on information such as earlier surveys or firms' public R&D funding applications. This survey collected detailed information about investments in R&D, in-house and external, including the locations of R&D activities. We use the R&D location data to construct our main explanatory variables of interest.

For the Finnish CIS, Statistics Finland surveyed all Finnish manufacturing firms with more than 100 employees, as well as a random sample stratified by size and industry of the remainder of the population of Finnish manufacturing companies. The innovation survey included questions about innovation output, R&D activity, and knowledge sources related to innovation. As the CIS data have become available in several countries, scholars have begun to use them to measure innovation output as a complement to more traditional measures such as patents (e.g., Leiponen, 2000; Leiponen, 2005; Mairesse and Mohnen, 2002; Cassiman and Veugelers, 2002).[2] The CEO or the R&D manager of each firm filled out the survey; the response rate was 50 percent.

We combined data from the CIS survey administered in 2001 with data on R&D location from the R&D survey administered two years earlier, in order to obtain a cross-sectional sample of firms that had data in both surveys. The R&D survey reflects data as of 1998 and the CIS survey covers the pe-

---

[2] Although patents reflect success in creating something new, they do not necessarily result in commercially viable innovations (Griliches, 1990). Moreover, in most industries, firms do not rely heavily on patents (Levin et. al., 1987). The CIS data provide a direct measure of success in commercializing innovations for a broad range of industries. Kleinknecht, Montfort and Brouwer (2002) found that CIS innovation output (measured as the share of sales revenue per employee derived from innovative products) was not correlated with the number of patent applications per employee. This finding suggests that the CIS data provide useful complementary measures of innovation success that more traditional measures may not capture.

riod 1998-2000. We used these cross-sectional data as our primary sample. In order to conduct robustness checks and control for potential unobserved heterogeneity, we also created a short panel that included a subset of firms for which we have observations of R&D locations and innovation success from later R&D surveys. Here we describe the cross-sectional sample. We describe the panel data in the section on statistical inference.

The cross-sectional sample includes 469 manufacturing firms with activity directed toward innovation, regardless of whether the firms succeeded in innovating, and encompasses all of the manufacturing industries in Finland. Some Finnish industrial companies are organized as so-called "business groups." For example, the basic metals company Outokumpu Inc, is a business group that has divisions, in this case Outokumpu Stainless Steel and Outokumpu Technology, which are wholly owned subsidiaries of the parent company. Survey respondents were either independent (not part of a business group), a subsidiary of a business group, or in a few instances, the parent company itself (the business group). As a result, the firms in the sample were not widely diversified. Because the data are confidential, firms in the dataset are not identified by name.

Table 1 compares the industry distribution of firms in our sample with the industry distribution of the Finnish manufacturing sector as a whole. Relative to the manufacturing sector as a whole, the forest sector (wood, pulp, paper, printing and publishing) is underrepresented in our sample, while the chemicals and electronics sectors are overrepresented. This may occur because our sample includes only firms with some activity directed toward technological innovation, and innovation activities are relatively less frequent in the forest-related sector and relatively more frequent in chemicals and electronics. Otherwise, the sample used here is very similar to the actual distribution of Finnish manufacturing industries. Importantly, unlike in many analyses of innovation, the majority of the firms in the sample are not in science-based industries.

When we combined data from the R&D survey with data from the CIS innovation survey, we experienced some sample attrition. Compared with firms active in R&D in the manufacturing sector as a whole, firms in our cross-sectional sample are larger (a mean of 383 employees in 2000 versus 182 employees in the manufacturing sector as a whole) and slightly more R&D intensive (R&D expenditures equal to 3.3% versus 2.9% of sales). But surprisingly, the firms in our sample are less innovative than R&D active firms in the manufacturing sector as a whole (67% are successful product innovators in our sample versus an average of 75% in the entire manufacturing sector). Although the firms in our sample are substantially larger than those in the broader Finnish economy, the sample does not appear to be consistently or strongly biased towards firms with greater innovation success.

Table 1 Industry Representation

NACE class	Industry	Percentage in sample	Percentage in the manufacturing sector as a whole
15-16	Food products, beverages, and tobacco	7.5	8.7
17-19	Textiles, textile products, leather, and leather products	3.4	5.1
20-22	Wood, wood products, pulp, paper, paper products, printing, publishing	10.0	18.7
23-25	Coke, refined petroleum products, nuclear fuel, chemicals, chemical products, manmade fiber, rubber, plastics	15.8	10.1
26	Nonmetallic mineral products	6.0	4.5
27-28	Basic metals, fabricated metal products	13.2	15.8
29	Machinery and equipment	17.5	16.0
30-33	Electrical and optical equipment	17.1	11.1
34-35	Transport equipment	5.5	4.2
36	Furniture, other manufacturing, recycling	4.1	5.9
	Total manufacturing	100.0	100.0

3.1 Variables

Using data from the surveys, we constructed variables that measure innovation success, the extent to which firm use multiple R&D locations, the range and types of external knowledge sources utilized by firms in their innovation activities, and variables that control for many additional factors that could affect innovation success or the propensity of firms to utilize multiple locations for their R&D activity. These variables are next described in turn.

3.1.1 Dependent Variables

The data used to construct the dependent variables regarding innovation success come from the CIS survey, which provided a detailed explanation to respondents of what constituted a technological innovation.[3] To measure innovation success, we utilized binary variables as well as sales revenue data. We constructed several binary (0,1) variables that indicate whether or not the firm introduced techno-

[3] The survey defines a product innovation as either a technologically new product or a technologically significant product improvement. A technologically new product is one whose purpose or technological characteristics are clearly distinct from those of the existing products of the firm. The new product can be based on a new technology, a new application of existing technologies, or application of new knowledge. A technologically significant product improvement significantly improves on the characteristics or performance of an existing product of the firm, and may include improvements in components, materials, or subsystems. The survey defines a process innovation as one that is technologically new or that contains a fundamentally improved method of production or product distribution. A process innovation may include (but is not limited to) improvements based on changes in equipment, instruments, organization of production, or new knowledge.

logical innovations during the 1998-2000 period. The first binary variable indicates whether the firm introduced technological innovations of any type, either product or process innovations or both (relevant to hypotheses 1 and 2). These innovations were new to the firm, and may or may not have been new to the market (novel). The innovation survey also asked respondents whether they introduced any novel product innovations (new to the market). In order to test hypothesis 4, the second binary variable indicates whether or not the firm had any novel product innovations. For purposes of comparison with the latter variable, we also utilized a third binary variable indicating whether the firm had product innovations of any sort, regardless of their novelty. 67% of the firms in the sample introduced product innovations. Approximately three quarters of these firms (and 52% of all firms in the sample) introduced novel (new to the market) product innovations, often in addition to imitative innovations; the remaining firms introduced only imitative product innovations (16% of all firms). Fewer firms introduced process innovations (46% of all firms).

The second type of dependent variable measures sales from innovative products. The CIS contains information about the percentage of total firm sales revenues in 2000 that derived from technologically new products that the firm introduced during the period 1998-2000. We multiplied the percent of sales from new products by total firm sales revenues in 2000 to obtain the value of sales from innovative products in 2000. This variable reflects sales from all new products, novel or not. We also constructed a variable that measures innovative sales revenues only for firms that had novel product innovations. All product sales variables are expressed in natural logarithmic terms to reduce variance and give less weight to the handful of very large firms.

The product sales variables have the advantage that they provide measures of the extent of commercial success, in contrast to the binary innovation variables which provide only a minimum measure of innovation success (commercialization of at least one innovation). Among the firms that innovated, approximately 90 percent introduced product innovations, indicating that use of a product sales variable is appropriate. Since over half of the product-innovating firms also had process innovations, however, product sales do not fully reflect innovation success. By using both the binary innovation and the product sales variables, we obtain a fuller picture of the different dimensions of innovation success.

In addition to the binary and sales variables that measure the type and amount of innovation success, we created two dependent variables that reflect the range of applicability of each firm's innovations, in order to test hypotheses 3a and 3b. The first variable is binary (0,1), indicating whether the firm had both product and process innovations, as opposed to only one or neither type of innovation. Three-quarters of the firms introduced some type of innovation. Of this group, about half introduced *both* product and process innovations. The latter variable is a measure of the range of innovation impact. The innovation survey also asked the firms to assess the impact of their innovations in nine different categories: expanding product range, improving product quality, extending market share or opening up new markets, improving production flexibility, expanding production capacity, reducing labor costs, reducing materials costs, reducing environmental effects, and fulfilling government regulations and standards. The survey asked respondents to evaluate each of these effects on a four-point scale ranging from zero (not applicable or no impact) to three (a very substantial impact). To measure the range innovation impact, we constructed a variable that reflects the number of different effects that firms viewed as important or very important, using the same procedure as that described below for knowledge sources.

3.1.2 Explanatory Variables of Interest

Our key explanatory variables relate to: 1) the number of locations where firms carried out R&D activities and 2) the number of different external sources of knowledge that firms reported were important in their innovation activities. The location variables derive from the R&D survey and reflect 1998 data. R&D locations are identified as separate counties where R&D activity took place within Finland. Most firms centralized their R&D activities: only 7.5% of the firms had two R&D locations and 5.3% had three or more locations. A mere 1% of firms had six or more locations. Given this distribution of the location data, we formed two binary indicators, one indicating whether a firm had two locations and another indicating whether a firm had three or more locations. The reference case (and omitted variable) is one location.

The innovation survey asked respondents to identify the importance of each of seven different external sources of information used in innovation activities: customers, suppliers, competitors, universities, non-profit research institutes, professional meetings and publications, and trade fairs and exhibitions. We use this information to test hypothesis 2 concerning the relationship between multiple R&D locations and the number of external knowledge sources. For each knowledge source listed in the survey questionnaire, each firm was asked to “evaluate the importance of the following sources of information for the innovation activities of your firm" on a four-point Likert scale from zero (not important at all/not used) to three (very important).

To account for the varying importance of different knowledge sources, we adopted the approach introduced by Cohen and Malerba (2001) in their analysis of industry-level innovation activity and subsequently used by Leiponen and Helfat (2005) for firm-level analysis. For each of the seven knowledge sources external to the firm, we first assigned a binary value based on whether or not the survey response indicated that the item was important to the firm. A survey response of either two (important) or three (very important) for a knowledge source received a binary value of one; a survey response of zero (not important at all/not used) or one (some importance) received a binary value of zero. The use of binary values of importance helps to alleviate potential measurement error that might arise from use of a Likert scale in the survey questions (Cohen and Malerba, 2001). When filling out a survey, respondents may have difficulty making fine-grained distinctions between very important and important sources, or between unimportant and not very important sources, which could result in measurement error. Additionally, use of the binary transformation alleviates the problem that an ordinal Likert scale should not be interpreted as an interval scale: a survey response of "3" does not necessarily mean that the item is 3 times more important that an item that received a response of 1.

To construct a variable indicating the number of important external knowledge sources, we summed the binary values for the seven external knowledge sources. Other researchers, such as Mol and Birkenshaw (2006) and Laursen and Salter (2006), have adopted a very similar approach using UK CIS data. This variable has a maximum value of seven. The most important external sources in the sample are customers and suppliers, followed by professional and trade meetings. Interestingly, universities and competitors are considered less important than professional and trade sources. In sensitivity analyses reported later, we assess the role of particular types of external knowledge sources.

As noted earlier, to measure the range of applicability of a firm's innovations, we constructed a variable that reflects the number of different effects that firms viewed as important or very important, using the same procedure as that just described for knowledge sources. For the innovation effects variable, we assigned a binary value of one if an individual innovation effect received a value of two (substantial impact) or three (a very substantial impact), and a zero otherwise. We then summed the binary values for the nine possible innovation effects.

3.1.3 Control Variables

Clearly, factors other than the location of R&D activity and external sourcing of knowledge affect innovation success. We control for many such factors in the analysis. In addition, we control for factors other than those connected with external sourcing of knowledge that might affect the propensity of firms to use multiple locations for their innovation activity. One alternative method would be to use a two-stage model to first predict the number of R&D locations and then use this predicted value in the innovation output equations. We attempted to do this, but were unable to construct valid and statistically significant instruments from the available data. The control variables are described below.

Logarithm of Number of Employees. Firm size is likely to be a particularly important predictor for the binary (0,1) innovation variables. Because larger firms have access to greater financial and human resources, these firms may have a greater ability to achieve at least a single innovation. Larger firms also may derive more sales from a single innovation, since the firms have a larger base of customers on which to build. Additionally, including firm size in the analysis helps to control for potential non-response bias, since the firms that responded to the innovation survey were larger than those in the target population. Finally, since larger firms may have a greater propensity to utilize multiple R&D locations, it is important to control for firm size. Larger firms are more likely to be able to afford the fixed costs of setting up multiple R&D locations, as well as the incremental operating costs. Again, we use the natural logarithmic transformation of the raw data.

Logarithm of R&D Expenditures. Because firms explicitly direct R&D spending toward the development of new products and processes, greater R&D expenditures may increase the probability of, and revenues from, successful innovation. The greater the total amount of R&D spending, the greater is the likelihood of achieving at least one innovation and the greater is the likely total amount of sales revenues from innovation. (Note that it is the total amount of R&D spending that potentially affects our dependent variables, not the ratio of R&D spending to sales.) In addition, since the number of R&D locations may be correlated with the amount of R&D spending, it is important to control for scale of R&D. Although R&D spending does not include all funds devoted to innovation activity, which can take other forms (e.g., process engineering), it should capture most direct spending on innovation. R&D expenditures are also measured in natural logarithmic form.

Business Group Subsidiary and *Business Group Parent.* These binary (0,1) variables come from the innovation survey and measure whether a firm was either a subsidiary of a larger firm or a business group parent company at the time of the survey. These variables are mutually exclusive. The excluded dummy variable is that for independent firms. The business group variables control for whether the firm was part of a diversified company, which can affect innovation success as well as the propensity to have multiple R&D locations. For example, firms in a business group might have access to internal knowledge sources of other firms in the group that could influence their ability to introduce innovations. The propensity of firms in business groups to set up multiple R&D labs also might be affected by corporate parent decisions regarding the conduct of R&D for all firms in the business group. In addition, the business group parent variable controls for the possibility that these firms might report a larger number of R&D locations due to the inclusion of their subsidiaries in their survey responses.

Logarithm of Export Revenues. The potential for greater sales outside of Finland may increase the incentive to innovate. Additionally, if firms with a high volume of exports face more intense international competition, they may have stronger motivation to introduce new innovations than domestically oriented firms. We therefore control for export revenues, measured in natural logarithmic form.

Sales Growth Due to Mergers & Acquisitions and *Sales Reduction Due to Divestitures.* These binary (0,1) variables come from the innovation survey. The first variable measures whether or not mergers or acquisitions increased firm sales revenues by 10% or more during the 1998-2000 period covered by the CIS survey. The second variable measures whether divestitures decreased firm sales revenues by 10% or more during the same period. These variables control for the possibility that both innovation output and R&D locations could reflect recent large acquisitions or divestitures. The excluded dummy variable is that for no (or less than 10%) change in sales from mergers, acquisitions, and divestitures.

Industry of Operation. Industry level factors such as technological opportunity, appropriability of the returns to innovation, and customer demand for new products may affect the incentives of firms to innovate as well as the likelihood and extent of innovation success. We include a (0,1) dummy variable for each two-digit level NACE industry in the sample (excluding the non-metallic minerals industry).[4]

Table 2 presents descriptive statistics for the variables used in the analyses, as well as for additional variables describing the characteristics of the sample. In 1998, the average firm had approximately 350 em-

[4] The data make it difficult to utilize a more fine-grained classification of industry affiliation, which would result in a significant number of industries with just one or a few firms. NACE (Nomenclature Actuariel dans la Communauté Européenne) is the European standardized industry classification system that is similar to the North American Standard Industrial Classification (SIC) system.

ployees and sales of 64 million euros, and spent 4.3% of sales on R&D activities. Moreover, Finnish manufacturing firms were highly internationalized. Average sales revenues from exports were 35 million euros, or 42% of average total revenues per firm. Approximately half of the sampled firms were part of a business group and 5% were business group parents. Finally, 10 percent of the firms experienced sales growth from mergers and acquisitions, and 4 percent experienced sales declines from divestitures.

Table 2 Descriptive statistics for the cross-sectional sample (N = 469)

Variable	Mean	Standard Deviation	Minimum	Maximum	Mean for firms with 1 R&D location	Mean for firms with multiple R&D locations
Employees	354.122	1553.888	5	22000	209.086	1342.783
Log(employees)	4.443	1.394	1.609	9.461	4.272	5.612
Export revenues (thousands of euros)	35797.550	190147.300	0	3410100	17821.35	158335.3
Log(export revenues)	7.949	2.309	0	15.042	7.697	9.668
Sales growth due to mergers or acquisitions (M&A)	0.102	0.303	0	1	0.090	0.183
Sales reduction due to divestitures	0.038	0.192	0	1	0.034	0.067
Business group subsidiary	0.510	0.500	0	1	0.482	0.700
Business group parent	0.051	0.221	0	1	0.037	0.150
R&D expenditures	5800.192	20733.340	0	311149	3206.785	23478.580
R&D intensity (expenditures/sales)	0.043	0.095	0	0.933	0.043	0.045
Log(R&D expenditures)	6.361	2.831	0	12.648	6.025	8.655
Product innovation	0.674	0.469	0	1	0.648	0.850
Process innovation	0.458	0.499	0	1	0.425	0.683
Novel innovation	0.546	0.498	0	1	0.521	0.717
Any innovation	0.742	0.438	0	1	0.719	0.900
Both product and process innovation	0.390	0.488	0	1	0.355	0.633
Sales share from newly introduced products	0.151	0.212	0	1	0.153	0.140
Log(product innovation sales revenue)	5.192	3.996	0	13.038	4.832	7.646
Sales share from newly introduced novel products	0.103	0.185	0	1	0.106	0.085
Log(novel innovation sales revenue)	3.950	4.043	0	13.038	3.677	5.812
Number of R&D locations	1.2708	1.026	1	13	1	3.117
Two R&D locations	0.075	0.263	0	1	0	0.583
Three or more R&D locations	0.053	0.225	0	1	0	0.417
Sum of important innovation effects	3.188	2.730	0	9	3.046	4.150
Sum of important external knowledge sources	2.525	2.015	0	7	2.320	3.900

In order to provide a sense of how firms with multiple R&D locations differ from firms with only a single location, the last two columns in table 2 display descriptive statistics for these two groups separately. Clearly, multiple R&D locations correlate with firm size, which also translates into higher probability of exports, mergers or acquisitions, divestitures, group structure, and higher R&D expenditures. Firms with multiple R&D locations, however, do not have greater R&D intensity (expenditures relative to size) than their single-location counterparts. Nevertheless, firms with multiple locations have a higher probability of innovation, greater sales from innovative products, and broader range of innovation impact. Interestingly, multi-locating firms also seek external information from a larger number of sources. Additionally, table A1 reports pairwise correlation coefficients for the variables included in the regressions.

3.2 Statistical Inference

Surveys can pose issues related to non-response bias and common method variance. Controlling for firm size in the regressions alleviates the issue that the firms that responded to both the innovation survey and the R&D surveys were twice the size of the set of firms that received the survey. We also noted earlier that despite the difference in average firm size of respondents and non-respondents, respondents had less innovation success than the targeted population as a whole. This suggests that any non-response bias does not inflate the survey measures of innovation success.

Both the innovation and the R&D surveys had a single respondent per firm, suggesting the need to check for common method variance, which could inflate any observed correlations between the dependent and independent variables. Since the variables come from two different surveys, and the same person may not have filled out all forms, this reduces the potential for common method variance. Nevertheless, as insurance we conducted a standard check for common method variance, using Harmon's one-factor test to assess common method bias. If common method variance is a serious problem, a factor analysis would produce a single factor that accounts for most of the correlation between the dependent and independent variables (see Podsakoff and Organ, 1986). The results of this analysis, described in the footnote below, indicate that our regressions are not subject to an inherent common method bias in the survey responses.[5]

Most of our explanatory and control variables reflect data for 1998 and the dependent variables reflect data for 2000 (sales variables) and 1998-2000 (binary variables), thus alleviating simultaneity issues. In particular, our main explanatory variables for number of locations date from 1998, as do R&D expenditures, number of employees, and export sales. The knowledge source variable reflects

[5] We performed a principal components analysis for the cross sectional sample that included all explanatory and control variables and the dependent product innovation sales variable. We used the sales dependent variable rather than the binary dependent variable, because principal component analyses tend not to work as well for binary variables. The analysis retained nine factors with eigenvalues greater than 1.00, and no factor explained more than 20 percent of the variance. Moreover, the dependent variable did not load most strongly on the same component as did the explanatory variables of location. Excluding the industry dummy variables, the principal component analysis still retained five components with eigenvalues above 1.00, and the first component explained only 25% of the variance. R&D location variables continued to load most strongly on different components than the dependent variable.

the years 1998-2000, which suggests greater potential for simultaneity. A common solution to this form of potential endogeneity utilizes instrumental variables. Although we do not have data with which to construct appropriate instruments, the use of product sales as a dependent variable mitigates this concern to some extent, since the variable includes only sales in 2000.

The use of cross-sectional data also poses the concern that despite the many control variables in our analysis, unobserved firm characteristics might explain both R&D location choices and innovation outcomes. We attempted to examine this possibility with panel data derived from the R&D surveys, which are conducted every two years. (Because the CIS survey is conducted only every four years and data subsequent to 2000 are not yet available, we could not use multiple years of CIS data to form a panel.) The R&D surveys contain data on R&D locations, expenditures, and employees, as well as some of the binary and sales innovation measures, but lack data on breadth of innovation impact, knowledge sources, and exports.

We constructed two different panels, each consisting of two periods. The first time period in both panels reflected the same time period as the original cross section, with data on innovation output data for 2000 and R&D locations for 1998, and included only firms in the original cross section. To construct the first panel, we merged data from the first time period with innovation output data for 2002 and location data for 2000. Differences in the sample of firms in the R&D surveys used to construct the panel reduced the sample size to a subset of 200 firms that were in the original cross section. Because this first panel revealed relatively little change from 1998 to 2000 in the number of R&D locations within firms, potentially making fixed effects difficult to estimate, we constructed a second panel that would enable us to assess changes in location over a longer time period from 1998 to 2002. For this second panel, we merged data in the first period with innovation output data for 2004 and location data for 2002. The additional data requirements for the second panel resulted in an even smaller sample of 120 firms.

We used a Chamberlain (1980) approach to obtain fixed effects type estimates for probit and tobit models that allow for correlation between the included explanatory variables and unobserved firm characteristics (for more detailed explanation, see Wooldridge, 2002). In addition, we estimated conditional maximum likelihood logit models for the binary innovation indicators.

Unfortunately, inference from these longitudinal data is hampered for two reasons. First, the panels contained too few firms to enable us to identify location effects, even using cross-sectional probit and tobit estimation without accounting for fixed effects. As in the original cross section, a large proportion of firms had only one location. The reduced sample size made location effects difficult to estimate with precision. Second, under these conditions, the Chamberlain-Wooldridge estimation procedure to account for fixed effects did not permit identification of location effects either. Instead, we found that few firms changed the number of their R&D locations over the period of study. This result is perhaps not surprising, given that it may be very costly to add or close down R&D locations. It also suggests that the cross-sectional analysis may depict a steady state situation for many firms. If firms have made efficient R&D location choices, we would not necessarily expect changes over time.

4. RESULTS AND DISCUSSION

Tables 3 and 4 contain the results relevant to hypotheses 1 and 4 regarding the amount of innovation success and the novelty of product introductions. Table 3 displays probit estimates for the binary dependent variables. We find that having two R&D locations is positively and statistically significantly associated with the probability that firms introduced innovations of any type or introduced at least some product innovations; the coefficient for three or more R&D locations does not differ statistically significantly from zero. As expected, R&D expenditures are positively and statistically significantly associated with the probability of innovation success. The only other significant variables are the industry dummy variables, particularly in the product innovation regression, and the divestment variable (although this reflects a relatively small number of firms).

In table 4, continuous dependent variables measuring sales from innovative products focus our attention on the economic impact of innovation on the firm. Largely the same variables are significant in the main product sales regression as for the binary innovation indicators, particularly the R&D location variables. The logarithm of the number of employees is positive and significant as well, and the M&A variable rather than the divestment variable is significant. These results for both the binary and product sales variables lend support for hypothesis 1 that multiple R&D locations, up to a point, are associated with greater innovation success. Consistent with hypothesis 1, we also find evidence of diminishing returns to multiple R&D locations: the coefficient for two locations is positive and significant, but that for three of more locations is insignificant.

Table 3 Estimation results for binary innovation indicators

Dependent variable:	Product innovation			Any innovation			Novel innovation		
Explanatory variables	Coeff.	S.E.	ME	Coeff.	S.E.	ME	Coeff.	S.E.	ME
Constant	-1.055***	0.356		-0.567	0.375		-1.247**	0.344	
Log(employees)	0.054	0.076	0.019	0.095	0.079	0.029	-0.008	0.073	-0.003
Log(R&D)	0.108***	0.027	0.038***	0.115***	0.027	0.035***	0.117**	0.027	0.046***
Log(exports)	0.023	0.041	0.008	-0.014	0.043	-0.004	0.031	0.041	0.012
Business group	-0.309	0.332	-0.114	-0.093	0.356	-0.029	0.076	0.316	0.030
Business group subsidiary	0.103	0.149	0.036	0.050	0.154	0.015	0.109	0.143	0.043
M&A	0.217	0.231	0.072	0.286	0.251	0.079	-0.003	0.216	-0.001
Divest	-0.578*	0.327	-0.220*	-0.581*	0.339	-0.205*	-0.243	0.321	-0.097
Two locations	0.726**	0.321	0.204**	0.917**	0.394	0.195**	0.277	0.249	0.107
Three or more locations	-0.007	0.328	-0.003	-0.163	0.348	-0.052	0.137	0.307	0.054
Food	0.483	0.347	0.147	0.271	0.370	0.075	0.385	0.330	0.146
Textiles	0.214	0.405	0.070	-0.027	0.422	-0.008	-0.046	0.411	-0.018
Forest	-0.278	0.315	-0.102	-0.136	0.331	-0.043	-0.222	0.315	-0.088
Chemical	0.235	0.290	0.078	0.104	0.311	0.031	0.288	0.285	0.111
Metals	0.506	0.302	0.156*	0.248	0.320	0.070	0.544	0.298	0.203*
Machine	0.732***	0.297	0.217**	0.389	0.315	0.106	0.719*	0.287	0.264**
Electronics	0.567*	0.295	0.175*	0.252	0.313	0.072	0.367	0.285	0.141
Vehicles	-0.188	0.362	-0.068	-0.260	0.379	-0.085	-0.040	0.358	-0.016
Other, furniture	0.921**	0.420	0.234**	0.733	0.456	0.164	0.723	0.391	0.255*
Log likelihood	-258.821			-239.559			-290.445		

S.E. = Standard error. ME = Marginal effect. Probit ML estimation method.
*** implies significance at the 99% level, ** at the 95% level, and * at the 90% level.

Table 4 Estimation results for sales from innovative products

Dependent variable:	Log(product innovation sales) Coeff.		S.E.	Log(novel innovation sales) Coeff.		S.E.
Constant	-4.748	***	1.365	-7.306	***	1.868
Log(employees)	0.639	**	0.287	0.338		0.388
Log(R&D)	0.417	***	0.107	0.633	***	0.149
Log(exports)	0.197		0.163	0.153		0.222
Business group	-0.510		1.183	-0.397		1.580
Business group subsidiary	0.416		0.566	0.614		0.771
M&A	1.488	*	0.798	0.732		1.087
Divest	-2.026		1.320	-1.015		1.766
Two locations	1.797	**	0.913	1.275		1.219
Three or more locations	0.674		1.157	1.020		1.535
Food	2.021		1.328	2.044		1.807
Textiles	0.486		1.636	-1.610		2.353
Forest	-1.192		1.280	-0.916		1.757
Chemical	1.061		1.154	1.429		1.575
Metals	1.722		1.197	2.410		1.634
Machine	2.817	**	1.141	3.702	**	1.551
Electronics	2.135	*	1.146	1.698		1.566
Vehicles	-0.175		1.453	-0.060		1.988
Other, furniture	3.042	**	1.524	3.854	*	2.064
Sigma	4.888		0.216	6.319		0.328
Log likelihood	-1081.801			-958.053		

S.E. = Standard error. ME = Marginal effect. Tobit ML estimation method.
*** implies significance at the 99% level, ** at the 95% level, and * at the 90% level.

In addition to the foregoing results, tables 3 and 4 show that introduction of novel (new-to-the-market) product innovations and sales from these innovations do not appear to benefit from multiple locations of R&D. The variable for two R&D locations is not significant in either of the novel product innovation regressions, but they are significant in the regressions that include imitative innovations as well. This suggests that imitative imitations in particular benefit from multiple R&D locations. These results support hypothesis 4, in that multiple locations of innovation activity are associated with greater innovation success for non-novel than for novel innovations.

Next we examine how patterns of knowledge sourcing relate to R&D location choice, in order to test hypothesis 2. Table 5 presents a subset of earlier specifications with the sum of important knowledge sources as an additional explanatory variable. This variable includes all knowledge sources covered in the CIS survey, including the public sources of meetings, trade fairs, and publications. Location may affect utilization of these sources, since R&D personnel may be more likely to attend nearby meetings and trade fairs. Furman et al. (2005) also found that firms obtained greater benefit from the publications of university researchers who were located near their R&D labs. The regressions focus on the dependent variables that included non-novel (imitative) innovations, since multiple locations of R&D were not significant in the novel innovation models. (Table 5 includes the novel product sales regression for completeness.)

The models in table 5 demonstrate the strong correlation between the extent of knowledge sourcing and multiple locations of R&D: adding the knowledge source variable to the regressions wipes out the previously significant positive effects of multiple R&D locations on innovation outcomes. We note that the knowledge source variable dates from the same survey as the innovation output variables, reflecting knowledge sources that firms utilized for 1998-2000. This implies that multiple locations in 1998 might have enabled firms to access a greater number of knowledge sources during 1998-2000, which is completely consistent with the literature that motivated hypothesis 2. Overall, these results support hypothesis 2 by suggesting that R&D location decisions and knowledge sourcing decisions are strongly related. The positive association between the number of R&D locations and innovation success appears to reflect the number of knowledge sources that firms access outside of their boundaries.

We conducted several sensitivity tests to probe the results regarding access to external knowledge sources through location of R&D. (Due to space constraints, these results are available on request from the authors.) First, we assessed whether a few key knowledge sources were particularly important in producing these results. We therefore created (0,1) dummy variables for each of the non-public knowledge sources (suppliers, customers, competitors, universities, and research institutes), where a binary value of 1 indicates that the knowledge source was important or very important to the firm. When these variables are entered together in the regressions for the binary and product sales variables (for novel and imitative innovations combined), they all are significant except access to research institutes. Thus, no single knowledge source dominates the results. In addition, access to each of these knowledge sources individually does not completely reflect the breadth of knowledge sources that firms utilize. Unlike the results for the knowledge breadth variable, the location variable retains its significance in the regression for the binary innovation variable (although not in the regression for product sales).

We also assessed whether multiple locations of R&D might facilitate knowledge acquisition through outsourcing or R&D collaboration (Cassiman and Veugelers, 2002). First, we added a variable for expenditures on contract R&D to the regressions in tables 3 and 4. Then we added a variable for expenditures on outsourced knowledge, including patents, licenses, copyrights, and software. The coefficients on these outsourcing variables were positive and significant in the binary and product sales regressions (for novel and non-novel innovations combined). The coefficient on two R&D locations remained significant in the binary innovation but not the sales regression.

Next we assessed whether R&D collaboration in Finland with suppliers, customers, competitors, or universities might explain the location effects. First, we included four (0,1) binary variables in the regressions that indicated whether or not firms had any explicit R&D collaboration with each of these types of entities. When entered together, all of the collaboration variables except collaboration with competitors were significant in both the binary and product sales regressions. Like the results for outsourcing, the coefficient on two R&D locations retained its significance in the binary innovation regression. Thus, while a distributed R&D organization may facilitate R&D collaboration and outsourcing, these individual forms of knowledge acquisition do not appear to fully reflect the breadth of knowledge sources that firms utilize. Next we constructed a breadth of collaboration variable similar to that for knowledge sources. We summed the binary variables for each of type of collaborator. Then in table 5 we replaced the knowledge source variable with the variable for the number of important collaboration sources. Again, the collaboration variable wipes out the effect of R&D locations. As for knowledge sources, the breadth variable for collaboration appears to fully account for the location effects, but the dummy variables for individual types of collaborators do not. Taken together, these results suggest that the benefits of multiple R&D locations reflect both breadth of knowledge sourcing and closely associated breadth of R&D collaboration.

Table 5 **Selected models with important knowledge sources as an additional explanatory variable**

	Product innovation			Any innovation			Log(product innovation sales)		Log(novel innovation sales)	
	Coef.	S.E.	ME	Coeff.	S.E.	ME	Coeff.	S.E.	Coeff.	S.E.
Constant	-1.094 ***	0.378		-0.469	0.430		-4.460 ***	1.249	-7.087 ***	1.772
Log(employees)	0.003	0.080	0.001	0.023	0.089	0.006	0.407	0.264	0.107	0.369
Log(R&D)	0.059 **	0.029	0.020 **	0.066 **	0.031	0.016 **	0.206 **	0.100	0.404 ***	0.143
Log(exports)	0.010	0.044	0.003	-0.042	0.048	-0.010	0.155	0.150	0.122	0.211
Business group	-0.234	0.356	-0.083	0.078	0.422	0.018	-0.341	1.082	-0.308	1.497
Business group subsidiary	0.134	0.161	0.045	0.062	0.179	0.015	0.459	0.518	0.585	0.730
M&A	0.101	0.246	0.033	0.246	0.291	0.054	1.071	0.730	0.333	1.029
Divest	-0.437	0.336	-0.160	-0.407	0.369	-0.117	-0.985	1.207	0.152	1.668
Sum of important knowledge sources	0.326 ***	0.040	0.109 ***	0.475 ***	0.051	0.116 ***	1.126 ***	0.123	1.198 ***	0.175
Two locations	0.410	0.331	0.121	0.637	0.449	0.116	0.801	0.835	0.254	1.153
Three or more locations	-0.241	0.344	-0.085	-0.706 *	0.385	-0.222 *	0.235	1.053	0.517	1.450
Food	0.354	0.366	0.107	0.067	0.419	0.016	1.287	1.215	1.188	1.715
Textiles	-0.039	0.410	-0.013	-0.654	0.443	-0.204	-0.022	1.488	-1.928	2.199
Forest	-0.208	0.336	-0.073	-0.066	0.375	-0.017	-0.946	1.175	-0.659	1.668
Chemical	0.129	0.307	0.042	-0.126	0.354	-0.032	0.547	1.058	0.898	1.493
Metals	0.510	0.322	0.149	0.164	0.364	0.038	1.358	1.097	2.007	1.549
Machine	0.736 **	0.313	0.206 **	0.232	0.354	0.052	2.570 **	1.043	3.423 **	1.467
Electronics	0.490	0.314	0.146	-0.039	0.359	-0.010	1.509	1.050	1.038	1.485
Vehicles	-0.228	0.389	-0.080	-0.399	0.438	-0.114	-0.335	1.334	-0.319	1.893
Other, furniture	0.910 **	0.452	0.217 **	0.681	0.525	0.118	2.397	1.394	3.202	1.952
Sigma							4.445	0.196	5.939	0.306
Log likelihood	-222.100			-183.314			-1040.932		-934.115	

Estimation methods: Probit ML for product innovation and any innovation, tobit ML for log(innovation sales). S.E. = Standard error. ME = Marginal effect.
*** implies significance at the 99% level, ** at the 95% level, and * at the 90% level.

Table 6 Effects of location on the breadth of innovation output

	Product and process innovation			Sum of important innovation effects		Product and process innovation			Sum of important innovation effects	
	Coeff.	S.E.	ME	Coeff.	S.E.	Coeff.	S.E.	ME	Coeff.	S.E.
Constant	-1.178 ***	0.342		0.232	0.246	-1.202 ***	0.352		0.249	0.218
Log(employees)	0.189 **	0.074	0.072 **	0.022	0.052	0.162 **	0.077	0.061 **	-0.008	0.047
Log(R&D)	0.045 *	0.027	0.017 *	0.106 ***	0.024	0.011	0.029	0.004	0.035 *	0.020
Log(exports)	0.001	0.042	0.001	0.018	0.031	-0.009	0.043	-0.003	0.005	0.029
Business group	-0.142	0.313	-0.053	-0.153	0.197	-0.120	0.323	-0.045	-0.199	0.184
Business group subsidiary	-0.155	0.146	-0.059	-0.154	0.105	-0.157	0.149	-0.059	-0.143	0.091
M&A	0.382 *	0.206	0.150 *	0.345 ***	0.128	0.333	0.213	0.130	0.203 *	0.119
Divest	0.056	0.323	0.021	-0.366	0.287	0.212	0.323	0.083	-0.103	0.252
Two locations	0.431 *	0.238	0.169 *	0.291 **	0.142	0.254	0.241	0.099	0.044	0.128
Three or more locations	0.271	0.304	0.106	-0.154	0.195	0.178	0.309	0.069	-0.158	0.176
Sum of important knowledge sources						0.205 ***	0.035	0.078 ***	0.264 ***	0.021
Food	0.056	0.332	0.022	0.128	0.237	-0.055	0.343	-0.021	-0.161	0.212
Textiles	-0.593	0.419	-0.199	-0.177	0.317	-0.708	0.420	-0.226 *	-0.491	0.299
Forest	-0.686 **	0.324	-0.229 **	-0.155	0.236	-0.717 **	0.338	-0.235 **	-0.072	0.206
Chemical	-0.258	0.287	-0.095	-0.012	0.209	-0.369	0.296	-0.133	-0.078	0.183
Metals	-0.022	0.297	-0.008	0.093	0.218	-0.092	0.306	-0.035	0.036	0.191
Machine	-0.282	0.285	-0.104	-0.019	0.208	-0.344	0.294	-0.124	-0.059	0.183
Electronics	-0.335	0.287	-0.122	0.062	0.206	-0.469	0.297	-0.166	-0.040	0.182
Vehicles	-0.481	0.362	-0.167	-0.259	0.277	-0.548	0.374	-0.185	-0.333	0.246
Other, furniture	0.058	0.384	0.022	0.209	0.272	-0.043	0.396	-0.016	0.060	0.242
Delta				2.126	0.270				1.376	0.196
Log likelihood	-287.781			-1043.848		-270.281			-969.088	

Estimation methods: Probit ML for product and process innovation, negative binomial with constant dispersion for sum of innovation effects. S.E. = Standard error. ME = Marginal effect. *** implies significance at the 99% level, ** at the 95% level, and * at the 90% level

Interestingly, the results in table 5 show that novel innovators also benefit from breadth of external knowledge sources, but they do not appear to use multiple locations to access these sources. In addition, relatively few individual knowledge sources stand out as critical for novel innovators. When we utilize dummy variables for individual knowledge sources, only customers and suppliers are significant in both the binary and sales regressions for novel innovations. Of the R&D collaboration variables, only collaboration with customers is significant in both regressions. Although other studies have found that firms, particularly in the pharmaceutical industry, benefit from R&D locations near universities (Jaffe, 1989; Zucker, Darby, and Armstrong, 1998; Chacar and Lieberman, 2003), we do not obtain this result. Perhaps, on average, Finnish companies that have novel innovations do not rely heavily on universities; for example, Finland does not have a large pharmaceutical sector. Instead, the majority of the sample consists of firms in non-science based industries.

Finally, in order to test hypotheses 3a and 3b, we assessed whether multiple R&D locations are associated with wider versus narrower applicability of innovation success. Table 6 reports the results for two dependent variables that reflect the range of innovation impact: a binary variable indicating whether the firm made both product and process innovations (as opposed to one or the other or neither) and the sum of important innovation effects. We first report the base regressions and then add the knowledge source variable. In the base regressions for both the binary indicator and the innovation effects variable, having two R&D locations (but not three or more) is positively and statistically significantly associated with wider applicability of innovation success. This result rejects H3b in favor of H3a: multiple locations of innovation activity are associated with wider, not narrower, applicability of innovation output. We observe diminishing returns to multiple locations here as well. In addition, consistent with hypothesis 2, once we add the knowledge source variable, the variable for two R&D locations becomes statistically insignificant.

To summarize, the cross-sectional results suggest that, first, the extent and the breadth of applicability of innovation success is positively associated with multiple R&D locations, but is subject to diminishing returns; second, this relationship holds for non-novel but not for novel innovations; and, third, access to a wider range of knowledge sources, particularly through R&D collaboration, accounts for the positive effect of multiple locations of R&D. These regressions included a number of important controls for other factors that could have affected innovation success or the propensity to have multiple R&D locations.

5. CONCLUSION

This empirical study introduces a new dataset containing uniquely detailed data on firms' R&D locations to examine the effects of distributed R&D activity on innovation outcomes. We compare and contrast hypotheses that derive from two relatively separate strands of literature, one on multinational firm location decisions and another on organizational form and incentives. The results indicate that both sets of literatures help to explain patterns of R&D location.

Our results indicate that firms with multiple R&D locations generally benefit in terms of the extent and breadth of innovation success, consistent with the knowledge-based view. We also find, however, that these benefits are subject to diminishing returns, as the organizational and incentive literatures predict. Moreover, the benefits of multiple R&D locations are strongly related to the knowledge sourcing strategies that firms employ. These results are consistent with the interpretation that R&D location decisions are driven by the desire of firms to access external sources of knowledge for innovation activities, per the knowledge-based view. As a result, and consistent with depictions in the popular press, firms may locate R&D activities in major markets in order to obtain knowledge from customers and suppliers.

We also find that these results do not hold for novel innovations, consistent with organizational economic theory. By definition, patents represent novel technologies. Our findings regarding novel innovation and location of R&D are consistent with two recent studies on patenting. Singh (2006) found that international dispersion of patenting within firms (used as a proxy for location of R&D) did not have a positive correlation with the extent of forward patent citations (used to measure innovation success). Argyres and Silverman (2004) also found that organizational centralization rather than decentralization of R&D was positively associated with the extent and breadth of firm patenting activity. More generally, our results have implications for interpreting the results of the many studies that use patent data. When analyzing technological innovation, it appears important to distinguish between novel and non-novel innovations, since the results may differ for the two types of innovation.

Our findings overall suggest that although having multiple R&D locations is associated with less novel innovations, they result in increased revenues from innovation. In other words, locating near relevant knowledge sources such as customers and suppliers is associated with more productive but less technologically advanced innovation output. Future research could build on these results to investigate how R&D location affects search for knowledge from different sources, through different mechanisms, in different industries. Additionally, since these results are limited to a single country and relatively small, generally non-science based firms, more analysis is needed to assess their applicability elsewhere. Other data sources also are necessary to assess whether distributed R&D is welfare improving, but our results suggest that it may be profit maximizing (or at least profit-improving) for many firms, since multiple locations are associated with greater sales from innovative products—which in many markets is crucial for sustaining a competitive edge.

REFERENCES

Allen TJ. 1977. *Managing the Flow of Technology: Technology Transfer and the Dissemination of Technological Information with the R&D Organization.* MIT Press: Cambridge, MA.

Almeida P. 1996. Knowledge Sourcing by Foreign Multinationals: Patent Citation Analysis in the U.S. Semiconductor Industry. *Strategic Management Journal* 17: 155-165.

Argyres NS, Silverman BS. 2004. R&D, Organization Structure, and the Development of Corporate Technological Knowledge. *Strategic Management Journal* **25**: 929-958.

Brown S, Eisenhardt KM. 2000. *Competing on the Edge of Chaos.* Boston: Harvard Business School Press.

Buckley PJ, Casson MC. 1976. *The Economic Theory of the Multinational Enterprise.* Macmillan: London.

Cassiman B, Veugelers R. 2002. R&D cooperation and spillovers: Some empirical evidence from Belgium. *American Economic Review* 92(4): 1169-1184.

Chacar A, Lieberman M. 2003. Organizing for Technological Innovation in the U.S. Pharmaceutical Industry. In *Advances in Strategic Management.* Baum JAC, Sorenson O. (eds.) JAI/Elsevier: Oxford, UK 20: 317-340.

Chamberlain G. 1980. Analysis of Covariance with Qualitative Data. *Review of Economic Studies* 47: 225-238.

Chandler AD. 1962. *Strategy and Structure: Chapters in the History of the Industrial Enterprise.* MIT Press, Cambridge, MA.

Chung W, Alcacer J. 2002. Knowledge seeking and location choice of foreign direct investment in the United States. *Management Science* 48(12): 1534-1554.

Cohen WM, Malerba F. 2001. Is the Tendency to Variation the Chief Cause of Progress? *Industrial and Corporate Change* 10:587-608.

Chandler AD. 1962. *Strategy and Structure*. MIT Press: Cambridge, MA.

Cummings J. 2003. Work groups, structural diversity, and knowledge sharing in a global organization. *Management Science* 50(3): 352-364

Dunning JH. 1981. *International Production and the Multinational Enterprise*. Allen and Unwin: London.

Florida R. 1997. The Globalization of R&D: Results of a Survey of Foreign-Affiliated R&D Laboratories in the USA. *Research Policy* 26: 85-103.

Frost TS. 2001. The Geographical Sources of Foreign Subsidiaries' Innovations. *Strategic Management Journal* 22: 101-123.

Furman JL, Kyle MK, Cockburn I, Henderson RM. 2005. Public & Private Spillovers: Location and the Productivity of Pharmaceutical Research. Forthcoming, *Annales d'Economie de Statistique.*

Grant RM. 1996. Toward a Knowledge-Based Theory of the Firm. *Strategic Management Journal* **17**(Winter Special Issue): 109-122.

Griliches Z. 1990. Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature* 28(Dec.): 1661-1707.

Henderson R, Cockburn I. 1994. Measuring Competence? Exploring Firm Effects in Pharmaceutical Research. *Strategic Management Journal* **15**(Special Issue): 63-84.

Henderson R, Cockburn I. 1996. Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery. *RAND Journal of Economics* **27**(1): 32-59.

Hymer S. 1970. The Efficiency (Contradictions) of Multinational Corporations. *American Economic Review* 60: 441-448.

Jaffe AB. 1989. Real Effects of Academic Research. *American Economic Review* 79: 957-970.

Kay N. 1988. The R&D Function: Corporate Strategy and Structure. In *Technical Change and Economic Theory.* Dosi G., Freeman C., Nelson R., Silverberg G., Soete C. (eds.) Pinter: London: 283-294.

Kleinknecht A, Van Montfort K, Brouwer E. 2002. The Non-Trivial Choice Between Innovation Indicators. *Economics of Innovation and New Technology* 11(2): 109-121.

Kogut B. 1991. Country Capabilities and the Permeability of Borders. *Strategic Management Journal*, Summer special issue 12: 33-47.

Kogut B, Zander U. 1992. Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organization Science* 3: 383-397.

Kogut B, Zander U. 1993. Knowledge of the Firm and the Evolutionary Theory of the Multinational Corporation. *Journal of International Business Studies* **24**: 625-646.

Kuemmerle W. 1999. Foreign Direct Investment in Industrial Research in the Pharmaceutical and Electronics Industries--Results From a Survey of Multinational Firms, *Research Policy* 28: 179-193.

Laursen K, Salter A. 2006. Open for Innovation: The Role of Openness in Explaining Innovative Performance among UK Manufacturing Firms. *Strategic Management Journal* 131-150.

Leiponen A. 2000. Competencies, Innovation, and Profitability of Firms. *Economics of Innovation and New Technology* 9(1):1-24.

Leiponen A. 2005. Skills and Innovation. *International Journal of Industrial Organization* 23(5-6): 303-323.

Leiponen A, Helfat CE. 2005. *Innovation Objectives, Knowledge Sources, and the Benefits of Breadth.* Ithaca NY and Hanover, NH. Unpublished manuscript.

Lerner J, Wulf J. 2006. Innovation and Incentives: Evidence from Corporate R&D. Working Paper: Harvard Business School.

Levin R C, Klevorick AK, Nelson RR, Winter SG. 1987. Appropriating the Returns from Industrial Research and Development. *Brookings Papers on Economic Activity* 3: 783-820.

Mol MJ, Birkenshaw J. 2006. The Antecedents and Performance Consequences of Management Innovation. Working Paper: London Business School

Penner-Hahn J, Shaver JM. 2005. Does international research and development affect patent output? An analysis of Japanese pharmaceutical firms. *Strategic Management Journal* **26**: 121–140.

Podsakoff PM, Organ DW. 1986. Self-reports in Organizational Research: Problems and Prospects. *Journal of Management* 12:531-544.

Singh J. 2006. Distributed R&D, Cross-regional Knowledge Integration and Quality of Innovative Output. Working Paper, Insead.

Teece DJ. 1977. Technology Transfer by Multinational Firms: The Resource Cost of Transferring Technological Know-How. *Economic Journal* **87**(June): 242-261.

Teece DJ. 1980. Economies of Scope and the Scope of the Enterprise. *Journal of Economic Behavior and Organization* **1**: 223-247.

Tripsas M. 1997. Surviving Radical Technological Change through Dynamic Capability: Evidence from the Typesetter Industry. *Industrial and Corporate Change* 6(2): 341-377.

Von Hippel E. 1988. *The Sources of Innovation*. Oxford University Press: Oxford.

Williamson OE. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. Free Press: New York.

Williamson OE. 1991. Comparative Economic Organization: The Analysis of Discrete Structural Alternatives. *Administrative Science Quarterly* **36**(June): 269-296.

Wooldridge JM. 2002. *Econometric Analysis of Cross Section and Panel Data.* The MIT Press: Cambridge, MA.

Zhao M. 2006. Cross-regional Ties Within Firms: Promoting Knowledge Flow of Discouraging Knowledge Spillovers? Working Paper: University of Michigan.

Zucker LG, Darby MR, Armstrong J. 1998. Geographically Localized Knowledge: Spillovers or Markets? *Economic Inquiry* 36(1): 65-86.

Appendix

Table A1 Correlations among estimation variables

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	Log(employees)	1																	
2	Log(export revenues)	0.7265*	1																
3	Sales growth due to mergers or acquisitions (M&A)	0.1493*	0.1527*	1															
4	Sales reduction due to dives-titures	-0.0496	0.0162	-0.0675	1														
5	Business group subsidiary	0.4131*	0.3380*	0.0779	0.0628	1													
6	Business group parent	0.2135*	0.2322*	0.1770*	0.004	-0.1206*	1												
7	Log(R&D expenditures)	0.4292*	0.4753*	0.038	-0.0937	0.2315*	0.1584*	1											
8	Product innovation	0.1434*	0.1665*	0.0699	-0.074	0.118	0.0171	0.2753*	1										
9	Process innovation	0.2269*	0.1483*	0.1129	-0.0056	0.038	0.0971	0.1470*	0.3481*	1									
10	Novel innovation	0.1290*	0.1705*	0.0254	-0.0407	0.1075	0.0564	0.2823*	0.7628*	0.3579*	1								
11	Any innovation	0.1476*	0.1398*	0.0705	-0.0851	0.0942	0.0485	0.2710*	0.8474*	0.5425*	0.6464*	1							
12	Both product and process innovation	0.2372*	0.1861*	0.1193*	-0.0005	0.0677	0.0721	0.1717*	0.5566*	0.8694*	0.5189*	0.4717*	1						
13	Log(product innovation sales revenue)	0.3743*	0.3696*	0.1546*	-0.0776	0.2157*	0.1061	0.3651*	0.9051*	0.3896*	0.7303*	0.7670*	0.5799*	1					
14	Log(novel innovation sales revenue)	0.2866*	0.2945*	0.0945	-0.0473	0.1780*	0.0797	0.3352*	0.6806*	0.3566*	0.8922*	0.5768*	0.5011*	0.7796*	1				
15	Number of R&D locations	0.3926*	0.2956*	0.1647*	-0.0095	0.0844	0.2313*	0.3127*	0.104	0.1369*	0.1198*	0.0892	0.1597*	0.2218*	0.1989*	1			
16	Two R&D locations	0.0812	0.1542*	0.0647	0.1122	0.1	0.1182	0.1589*	0.1284*	0.1133	0.0961	0.1304*	0.1221*	0.1473*	0.0957	0.2021*	1		
17	Three or more R&D locations	0.3830*	0.2442*	0.0765	-0.0474	0.0999	0.1172	0.2761*	0.0639	0.1246*	0.083	0.0531	0.1410*	0.1777*	0.1507*	0.7893*	-0.0674	1	
18	Sum of important innovation effects	0.1357*	0.1604*	0.1496*	-0.0544	0.0096	0.0763	0.2390*	0.5932*	0.5613*	0.5512*	0.6786*	0.5348*	0.5989*	0.5366*	0.0825	0.1322*	0.0463	1
19	Sum of important external knowledge sources	0.3083*	0.2876*	0.1042	-0.0906	0.1475*	0.0885	0.3766*	0.4704*	0.3533*	0.3844*	0.5191*	0.3474*	0.5172*	0.4198*	0.2009*	0.1759*	0.1833*	0.5697*

Note: Correlation coefficients are marked with an asterisk * if they are significant at the 99% level of confidence.