Understanding the Inputs into Innovation:

Do Cities Substitute for Internal Firm Resources?

April 2006

Chris Forman
Tepper School of Business
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
cforman@andrew.cmu.edu

Avi Goldfarb
Rotman School of Management
University of Toronto
105 St George St
Toronto, ON M5S 3E6
CANADA
agoldfarb@rotman.utoronto.ca

Shane Greenstein*
Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston IL 60208
greenstein@kellogg.northwestern.edu

Abstract

We examine whether there is a tradeoff between employing internal (firm) resources and purchased external (local) resources in process innovation. We draw on a rich data set of Internet investments by 86,879 U.S. establishments to examine decisions to invest in advanced Internet technology. We find evidence of localization of substitution. In particular, we show that the marginal contribution of internal resources is greater outside of a major urban area than inside one. Agglomeration is therefore less important for highly capable firms. When firms invest in innovative processes they act as if resources available in cities are partial substitutes for both establishment-level and firm-level internal resources (JEL Classification: R30, O33, L86).

* We thank Ashish Arora, Tom Åstebro, Maryann Feldman, Steven Klepper, Megan MacGarvie, Brian McManus, Sandra Slaughter, Kristina Steffenson McElheran, Scott Stern, Will Strange, and many seminar participants for comments. We also thank Harte Hanks Market Intelligence for supplying data. We received funding from the Kellogg School of Management, SSHRC grant #538-02-1013, and the GM Strategy Center. All opinions and errors are ours alone.

1. Introduction

Innovation does not arise by accident. Firms choose to develop resources and processes to facilitate innovation (e.g., Cohen and Levinthal, 1990; Åstebro, 2002). Geographic location also plays an important role (e.g., Griliches, 1957). Prior work has demonstrated evidence of localization in innovation (Jaffe, Trajtenberg, and Henderson, 1993), suggesting that a propitious location may lower the costs of innovative output (Furman et al., 2005). It is widely assumed that such concerns have motivated firms in information technology (IT) hardware, software, and pharmaceuticals to cluster together (e.g., Saxenian, 1996; Bresnahan and Gambardella, 2004).

There is less understanding of the tradeoff between location and internal resources for investment in process innovation. In particular, there is little empirical evidence on the extent of localization of substitution between internal (firm) resources and purchased external (local) inputs into innovation. This is a surprising gap in understanding. If firms can innovate in their operations by substituting away from internal resources when they locate in areas such as cities, where purchased services are more readily available, then agglomeration will be most important for smaller firms with fewer internal resources.[1] Alternatively, if firms with rich internal resources use them to invest in process innovations, then locating in lower-density areas may have little effect on their ability to innovate.

We examine these trade-offs in a model of process innovation. In this model, establishments face an opportunity cost from redirecting resources to innovation that would otherwise be used for operations. We label these resources "capabilities." Establishments choose what fraction of their existing capabilities to dedicate to a new innovative investment and how many other inputs they should purchase from market suppliers. In this model, location shapes the cost of purchasing services for innovative activity. Firms in

---

[1] Vernon (1963) first articulated this idea in a case study of the New York City area. He argued that agglomeration economies are especially important for small (start-up) firms that lack scale economies in their own organizations. Agglomeration economies have also been cited as a reason for the success of nascent software firms in countries such India, Ireland, and Israel (Arora et al., 2004).

good locations can access local markets when they invest in a process innovation, even when they lack internal expertise. Alternatively, if a firm has rich internal expertise, it can innovate without accessing local markets.

We then develop hypotheses derived from this model. We compare these hypotheses with actual investment in an Internet-based process innovation in a large cross-section of U.S. establishments at different locations.

We focus on the largest investors in IT in the United States, a sample that contains considerable heterogeneity in the dimensions of interest. Specifically, we analyze a survey (conducted by Harte Hanks) of use of advanced Internet technologies at 86,879 establishments that had over 100 employees at the end of 2000.[2] This sample consists of established firms rather than start-ups, which allows us to treat establishment location as determined prior to the decision to invest in Internet technologies. Approximately two-thirds of the U.S. workforce is employed in this type of establishment. Most of the organizations in the sample have *some* experience with basic IT technologies, such as personal computers (PCs), but they differ tremendously in their capacity to manage large IT projects. Only a fraction of these establishments have extensive experience with advanced IT projects. The data contain detailed information about establishment-level IT personnel and infrastructure assets. Because 45,948 establishments come from one of 7,035 multi-establishment organizations, they also vary in their potential to move assets between establishments. Furthermore, establishments come from all over the United States, both major urban areas and isolated rural locations, so they vary in their potential to hire from local labor and service markets.

We show that establishments that are part of firms with greater internal capabilities invest in an Internet-based process innovation more frequently. Furthermore, we provide evidence that establishments in large cities behave as if they draw on local resources to invest in process innovation. Establishments

[2] We use the terms *firm* and *organization* interchangeably. *Multi-establishment firms* can have establishments in

act as if these inputs into innovation are partial substitutes for each other. Establishments also act as if they substitute internal capabilities located at other establishments in the same firm for those available locally. We infer that the marginal contribution of internal resources to innovation is lower in cities than in other areas. We also infer that firms that can draw on rich internal resources can innovate outside of cities. We find no evidence that internal and external resources are complements in process innovation.

Our paper advances the existing literature on the role of capabilities in innovation. First, we provide empirical evidence on the determinants of process innovation related directly to operations. Almost all prior research about the inputs into innovation has employed different measures of innovation inputs or outputs, such as patent citations or patent output (e.g., Singh, 2004). Second, we test our model on a large cross-section of industries and locations. Prior empirical studies of the role of resources in process innovation focus on case studies of a narrower set of industries and locations (e.g., Kelley and Helper, 1999; Henderson, 2003). Third, we develop and test a model of investment in process innovation that rationalizes observed patterns without assuming any unmeasured information spillovers or firm networks.[3] Fourth, our modeling approach links more closely with issues common in estimation of cost functions and production functions rather than the identification of rank, order and stock effects, as is more common in other analyses of new information technology adoption, such as CAD and CAM (e.g., Åstebro, 2002; Karshenas and Stoneman, 1993).

The emphasis in our research and our conclusion contrasts with much of the existing literature on the role of internal expertise in research and development activities (e.g., Arora and Gambardella, 1994;

---

more than one location, more than one establishment in a single area, or both.

[3] This is a common assumption in the empirical literature on innovation. In practice, many authors, such as Furman, Kyle, Cockburn, and Henderson (2005) and Jaffe, Trajtenberg, and Henderson (1993) are agnostic about what drives the spillovers. They focus on showing that the evidence is consistent with agglomeration contributing to a sharing of knowledge. To be precise, we neither reject nor accept any hypothesis about the presence of spillovers. Instead, we show that a plausible model without spillovers generates hypotheses consistent with our empirical data on process innovations.

Cassiman and Veugelers, 2006).[4] This literature argues that internal expertise complements the external stock of knowledge. We begin from the opposite premise, that external (local) expertise substitutes for internal expertise during innovative investments, i.e. there is localization of substitution between internal and external inputs into innovation. We find evidence consistent with our premises. In contrast to prior research on localization of innovation (e.g., Porter 1998), we conclude that highly capable firms do not necessarily need to locate in agglomerated areas or clusters to successfully innovate. Capable firms also may locate in more isolated regions and still, nonetheless, innovate at low cost. Firms without such capabilities, however, will face different concerns when they innovate, and, for purposes of innovating, may benefit from locating near other firms that also innovate.

We next develop the conceptual framework and hypotheses. Sections 3 and 4 present the data and empirical framework. Section 5 presents our results. In the conclusion, we develop implications for the literature on the geography of innovation and the literature on outsourcing.

2. Conceptual Framework and Hypotheses

We will focus on only one broad category of process innovations, those reducing costs of communications internal to the establishment but not outside of the establishment. This type of innovation appeals to establishments that have already found users for basic email and Internet browsing, but involves substantial further investments. Following nomenclature from prior work (Forman, Goldfarb, and Greenstein, 2005a), we label these investments Within-Establishment Internet (WEI). We provide a precise definition below. WEI was one of many new process innovations associated with the Internet. Because we examine investment in WEI early in its diffusion, we focus on an establishment's short-run

[4] Arora and Gambardella (1994) examine whether internal and external knowledge are complements in biotechnology innovation. Cassiman and Veugelers (2006) find that internal R&D and external knowledge acquisition are complements in a sample of 269 Belgian manufacturing firms. In both of these cases, external knowledge is not defined relative to location.

choice between employing internal or market-supplied inputs to implement this process innovation. We consider only this technology and tailor our model to these circumstances.[5] The model can be generalized to other innovative investments, but we leave that task to future work.

We expect that all firms who invest in WEI will use at least some of their internal capabilities. Firms investing in new general purpose technologies (GPTs) such as the Internet engage in co-inventive activity to translate the opportunities created by the diffusion of new GPTs into a useful specific application (Bresnahan and Trajtenberg, 1995). Firms may draw upon internal capabilities to engage in co-inventive activity. Co-inventive activity typically draws on firm-specific inputs, tailored to firm-specific needs.

Internal capabilities arise from prior firm investments in shared human and physical capital. They include intangible firm-specific investments in training employees and building internal knowledge of IT systems and how these investments support firm business processes, workplace organization, and product and service innovation.[6] For example, resources in the organization may already be employed in some IT-intensive task. In the course of such tasks, organizations may develop human and physical capital that reduces the costs of other IT-related innovations (such as developing Internet applications internally). There may be significant fixed costs to developing this capital, particularly if it relies on firm-specific knowledge or experience of the organization's particular technical infrastructure and practices.[7]

---

[5] We do not consider the long-run decision to develop internal capabilities or the decision of where to locate. As has been noted in many contemporary accounts, the rapid diffusion of the Internet took most commercial establishments by surprise. Greenstein (2001), Mowery and Simcoe (2002) and Kenney (2003) discuss many of the elements of this setting that contributed to the surprising boom in investment. These long-run choices were not within the purview of the medium and large firms that comprise our sample when we observed them.

[6] For further details on how IT may be complementary with particular organizational practices, see Bresnahan, Brynjolfsson, and Hitt (2002), Hubbard (2000), and Marschak (2004). In networking technologies, in particular, an establishment may have to alter complementary facets of organizational processes (Brynjolfsson and Hitt, 2000), re-apply technical or business lessons from prior experience in new contexts (Attewell, 1992), or invent new process to adjust older operations and make the most efficient use of legacy investments in installed equipment and operations (Bresnahan and Greenstein, 1996; Forman, 2005).

[7] Prior research in the costs of innovative activity has presumed that experience with prior related projects can lower the costs of innovation (Cohen and Levinthal, 1990). Internal capabilities also may help customize a general purpose technology to the idiosyncratic needs of the establishment (Forman, Goldfarb, and Greenstein, 2002).

Internal capabilities may reside at the same establishments (establishment capabilities) or within other establishments in the same firm (organizational capabilities). Though case evidence suggests these capabilities move effectively across IT projects or establishments within the same firm, it is an open question whether they have a significant effect on innovative outcomes.

Use of internal capabilities will typically involve opportunity costs arising from the utilization of scarce resources to engage in co-inventive activity. These include altering software, procedures, business routines, database fields and other firm-specific assets. Human and physical assets may have a market value different from their opportunity costs of use within the firm. Thus, we expect that increases in the stock of total internal capabilities will decrease the costs of using any fixed quantity of capabilities on any one process innovation, ceteris paribus. We state this hypothesis with greater precision below.

We also expect that firms will incur monetary costs when implementing a particular IT project. For example, new projects will involve monetary costs arising from wages, costs of IT hardware, as well as software licenses and fees. New process innovation draws on purchased market inputs, obtainable either through third party outsourcing firms or through independent contract programmers. External input suppliers may be able to apply lessons learned from heterogeneous projects in different contexts, thereby lowering the costs of new process innovation. For example, one common reason for utilizing systems integration and design firms is that they bring to bear experience with similar projects with other firms within the same industry.

We expect that the attractiveness of purchased market inputs will depend upon local supply conditions, which will drive costs up or down. We expect the availability of such "general specialties" to be increasing in market scale (Stigler, 1951; Bresnahan and Gambardella, 1998). Once again, we state this hypothesis with greater precision below.

In summary, existing case literature frames the basic features of a model of investment in a new process innovation using information technology, such as WEI. In our setting, the opportunity to invest in WEI became unexpectedly available and in the short run firms had to decide whether to invest in the innovative opportunity. We hypothesize that firms draw on potential inputs from market suppliers and from inside their establishment, and if they can, from other locations inside their organization. When firms purchase resources from a market they pay a monetary cost and when firms draw on internal capabilities they pay an opportunity cost. We also hypothesize that firms plausibly *might* use inputs from a variety of sources, but there is little evidence on the extent of substitution. We expect firms in good locations will be more inclined to invest than others and so will firms with greater internal capabilities. How often will firms with comparatively poor locations and good internal capabilities adopt? What about firms with comparatively good locations and poor internal capabilities? The answers to those open questions will provide evidence about the extent of substitution between inputs, if any.

2.1. A Model of WEI co-invention and investment

We consider a simple model of investment in an innovation. The firm has an existing production process it intends to improve with a one-time project. We use this model to develop empirically relevant hypotheses.

The decision to undertake such investment depends on anticipated benefits and costs, as does the size of the investment. We write the model of costs and benefits in anticipation of our empirical implementation. We do not expect to observe the size of benefits, except rather crudely through measurement of factors that move their level up or down, such as industry, establishment size and multi-establishment status. We also do not expect to observe the level of costs, but we do expect to observe general features that move the costs of co-invention up or down, such as the thickness of the labor market in the location of the establishment, historical software use, number of programmers under employment, and type of software languages in use, both at the establishment and at other establishments within the organization. Therefore we develop a model tailored to identify how different factors alter the costs and benefits of co-invention activity in comparison to some baseline level.

We posit a process determining the value of undertaking investment in WEI technologies, where this value is observed by the decision makers in the establishment and not by the researcher. We define

(1) $$Y = \overline{B} - \overline{C}$$

where Y is a latent variable for net benefits, $\overline{B}$ is the gross benefit of undertaking the investment, and $\overline{C}$ is the total cost. We let

$$\overline{B} = B(ec, oc, ps, me, x, t) + u^B$$

$$\overline{C} = C(\frac{ec}{EC}, \frac{oc}{OC}, me, x, t) + ps * c(z) + u^C \,.$$

Where

- *B()* is the gross benefit of the investment without the establishment-specific random benefit variable, u^B.[8]
- *C()* is the total cost of the investment without the establishment-specific random cost variable, u^C.
- *ec* represents the establishment's internal capabilities invested in the project.
- *oc* represents the capabilities at *other* establishments within the same firm invested in the project.
- *ps* represents purchased services from local markets used for the project. By purchased services, we mean any assets purchased in markets – either local or national – for co-investment activity aimed to implement the process improvement. This can include contract programming, consulting for installation design, contract maintenance, or software and capital equipment, for example.
- *EC* represents the total available establishment capabilities. For brevity, we will hereafter refer to these as establishment capabilities.
- *OC* represents the total available capabilities of the organization at other establishments within the same firm. For brevity, we will hereafter refer to these as organizational capabilities.

---

[8] Case evidence suggests that other IT projects at the same or other establishments within the same organization affect both the benefits and costs of investment in process innovations. For examples, see Austin et al. (1999), Bendolay and Jacobs (2005), and Nolan (2001).

- $c(z)$ represents monetary costs for purchased services, which are a function of local factors such as population size or population density, z.
- me indicates multi-establishment status.[9]
- x indicates establishment characteristics such as number of employees and industry.[10]
- t is time.
- u^B is an establishment-specific random benefit variable.
- u^C is an establishment-specific random cost variable.

Table 1 provides a glossary of all symbols and acronyms used in this paper.

We define $u=u^B-u^C$, assume that $u \sim N(0,\sigma^2)$, and rewrite equation (1) as a probit model (David, 1969; Karshenas and Stoneman, 1993):

$$(2) \qquad Y(ec,oc,ps,EC,OC,z,me,x,t) = B(ec,oc,ps,me,x,t) - C(\frac{ec}{EC},\frac{oc}{OC},me,x,t) - ps*c(z) + u$$

The main questions of interest relate to how an establishment's characteristics and environment affect its ability to undertake process innovation. It is not the actual *ec*, *oc*, and *ps* used that are of interest, but, instead, how the costs of using these inputs change across establishments and locations. In particular, we are interested in how establishment capabilities (*EC*), organizational capabilities (*OC*), and the local population size (*z*) interact to drive net benefits from process innovation.

We assume that the gross benefit, *B*, is increasing and concave in *ec*, *oc*, and *ps*, and independent

---

9 By definition, $me = 1$ for establishments with $OC > 0$.

10 We include the multi-establishment firm dummy (*me*) and the number of employees (in *x*) to control for the well-known result that net benefits to investing in new technologies are increasing in firm and establishment size. While we do not focus on how local competition influences the net benefits of investing in WEI, we later present robustness checks that control for the extent of local competition.

of location.[11] Furthermore, we assume that the gross benefit (given *ec* and *oc*) is independent of the total capabilities of the establishment (*EC*) or organization (*OC*). Therefore the total capabilities directly affect only the cost of implementation, $\overline{C}$.[12]

We begin by discussing the choice of *ec* given *EC*. Then we will summarize the similar arguments that apply to *oc* given *OC* and *ps* given *z*. Establishments take the capabilities of their establishment, *EC*, as fixed. If they undertake the investment project (in WEI), they choose the amount of these capabilities that will go to implementing the project. We assume that costs increase in the fraction of the establishment capabilities used in a project, $\frac{ec}{EC}$. Therefore, the costs of employing a given level of *ec* increase for a fixed *EC:* $\frac{\partial C}{\partial ec} > 0$ given EC. Furthermore, we assume the cost of using capabilities increases at an increasing rate: $\frac{\partial^2 C}{\partial (\frac{ec}{EC})^2} > 0$. We assume no economies or diseconomies of scope in resources used. In particular, the costs of using additional units of establishment capabilities are independent of the total organizational capabilities and of the local resources available: $\frac{\partial^2 C}{\partial \frac{ec}{EC} \partial OC} = 0$ and $\frac{\partial^2 C}{\partial \frac{ec}{EC} \partial z} = 0$. In summary, the costs of using establishment capabilities to invest in process innovation increase (and at an increasing rate) as the establishment capabilities used approach the establishment capabilities available. We make a parallel set of assumptions for organizational

---

[11] We explicitly rule out "spillovers" from locating in denser areas in this benefit function. We are comfortable with this assumption because we examine an innovation whose benefits are realized only inside the establishment.

[12] This assumption is used to better provide intuition regarding how total capabilities affect the net benefit of using capabilities in a given project. The theoretical implications can generalize to allow total capabilities to affect gross benefits.

capabilities.[13]

Furthermore, in terms of local services, we assume $\frac{\partial c}{\partial z} < 0$ and $\frac{\partial^2 c}{\partial^2 z} > 0$. The costs of purchasing local services increase as more of the available local services are purchased.[14] Altogether, the costs of advanced IT projects are lower for firms with access to rich local resources in major urban areas (e.g., Columbo and Masconi, 1995). Increases in each of these factors may decrease the costs of co-invention activities in cities, other things being equal.[15]

Establishments then choose *ec**, *oc**, and *ps** such that

(3) $$ec^*, oc^*, ps^* \equiv \underset{ec,oc,ps}{\arg\max}\, Y(ec, oc, ps, EC, OC, z, me, x, t)$$

Throughout, we assume that we are at an interior optimum.[16] In particular, we assume that all establishments purchase at least some local services and use at least some internal resources.[17] The above assumptions imply that $\frac{\partial ec^*}{\partial EC} > 0$, $\frac{\partial oc^*}{\partial OC} > 0$, and $\frac{\partial ps^*}{\partial z} > 0$.

We will not observe how much investment an establishment will make, but we will observe the decision to make any investment in the process innovation, WEI. An establishment will do so if and only if $Y \geq 0$ and will not do so if $Y < 0$. Since investment in WEI technology usually involves substantial changes to operations, it is rarely reversed. As in the conventional "probit model" of adoption (e.g.,

---

[13] In particular, $\frac{\partial C}{\partial oc} > 0$ given *OC*, $\frac{\partial^2 C}{\partial \left(\frac{oc}{OC}\right)^2} > 0$, $\frac{\partial^2 C}{\partial \frac{oc}{OC} \partial EC} = 0$, and $\frac{\partial^2 C}{\partial \frac{oc}{OC} \partial z} = 0$.

[14] This assumption will be particularly relevant at high levels of service. At low levels of service, it is possible that marginal costs will decrease due to bundling. We treat this as an empirical question.

[15] This is consistent with prior theory work arguing that firms locate administrative and support functions strategically. Duranton and Puga (2002) argue that a firm may find it advantageous to locate administrative and support services in large areas because of better availability and a larger variety of complementary services. In addition, external services may require repeated face-to-face interactions (Kolko, 1999). There is some evidence that locating such services in large metropolitan areas may improve firm productivity (Davis and Henderson, 2002).

[16] We do this for expositional convenience. Accounting for zero investment thresholds yields no additional benefit for formulating the empirical hypotheses.

[17] Empirically (for obvious reasons) we do not assume that *oc** is positive for single-establishment firms.

David, 1969; Karshenas and Stoneman, 1993), we assume costs decline over time for all potential decision makers, i.e., $C = g(t)[C(\frac{ec}{EC}, \frac{oc}{OC}, me, x) + ps * c(z)]$, where $g' < 0$, and benefits (weakly) increase over time, $B = h(t)B(ec,oc,ps,me,x)$ where $h' \geq 0$. In this model $Y= 0$ defines a marginal decision maker in a cross section for whom the costs and benefits of investment are equal. We focus on analyzing this marginal decision maker. Hence, for the remainder of the paper we will suppress the time dimension in our model.

Our first hypothesis reflects a basic intuition behind the model. We expect that an establishment in a location without access to a good supply of local contract programmers and outsourcing firms will behave differently from one in a location with access to a more elastic supply of these services. The establishment with good access will tend to use purchased services more frequently than an establishment in a thinner market. We now formalize this intuition in hypotheses. In this model, Y is increasing in population density.[18] Location will not shape investment if marginal investors do not use purchased services or if local areas do not differ in their supply of purchased services. We treat these alternative predictions, and any speculation about the magnitude of the contribution, as an empirical question. Following nomenclature from earlier work, we label this first prediction the "urban leadership hypothesis."[19]

Hypothesis 1 (from Forman, Goldfarb, and Greenstein 2005a): Investment of WEI will be increasing as location size and density increase (i.e., $\partial Y / \partial z > 0$).

---

[18] By the envelope theorem, $\frac{\partial Y(ps*)}{\partial z} = -\frac{\partial c}{\partial z} > 0$.

[19] Note minor similarities and differences with prior work. Forman, Goldfarb, and Greenstein (2005a) inferred that the geographic variation in WEI investment was consistent with the "urban leadership hypothesis." However, that prior inference did not control for internal capabilities, nor suggest an equilibrium framework, as we do here. Nonetheless, we believe it would be surprising if the inference of the prior research was sensitive to the omission of *EC* and *OC* variables, so we retain the label.

2.3. Hypotheses about capabilities

What role do internal firm resources play? As noted above, this is an open question in the case literature. We first discuss the role of *EC* and then discuss the role of *OC*.[20]

The intuition is similar to that stated in the previous discussion. An establishment with access to abundant internal resources will use them as long as the opportunity costs are not too high, while an establishment without access to abundant resources will not. More formally, establishments choose ec* to maximize the net benefit of investing in the process innovation, *Y,* given *EC*. Since $\frac{\partial B}{\partial ec} > 0$, $\frac{\partial ec*}{\partial EC} > 0$, and $\frac{\partial C}{\partial EC} < 0$, Y will be increasing in EC. A similar argument implies that Y will be increasing in OC.

Hypothesis 2a: Firms with greater organizational capabilities, ceteris paribus, will be more likely to invest in WEI technology at any one of their establishments than firms with fewer organizational capabilities (i.e., $\partial Y / \partial OC > 0$).

Hypothesis 2b: Greater establishment capabilities, ceteris paribus, will increase investment in WEI technology (i.e., $\partial Y / \partial EC > 0$).

2.4. Substitution

Now we consider substitution between *oc* and *ps.* As with the intuition in the three previous hypotheses, the substitution between different inputs into innovation, if there is any substitution at all, depends on the abundance of internal resources and the supply conditions experienced by the establishment in the local market. More formally, we think it is plausible to assume substitution between internal and external inputs in the benefit function: $\frac{\partial^2 B}{\partial ps \partial ec} < 0$ and $\frac{\partial^2 B}{\partial ps \partial oc} < 0$, i.e., the marginal benefits from additional investment in one input declines as the use of the other input rises. This assumption implies $\frac{\partial^2 Y}{\partial z \partial EC} < 0$ and $\frac{\partial^2 Y}{\partial z \partial OC} < 0$. That is, a rising level of capabilities decreases the relative cost of using internal capabilities to facilitate the investment.

Hypothesis 3a: The sensitivity of WEI investment to increases in location size will be declining as the internal organizational capabilities found in other establishments within the same firm increase (i.e., $\partial^2 Y / \partial z \partial OC < 0$).

Hypothesis 3b: The sensitivity of WEI investment to increases in location size will be declining as the internal establishment capabilities increase (i.e., $\partial^2 Y / \partial z \partial EC < 0$).

These are the core hypotheses of the paper, and are therefore the focus of most of the empirical work and discussion. Note that these hypotheses presume that the marginal investment will be decided based on human capital issues instead of infrastructure issues. We are comfortable with this assumption because access to the relevant infrastructure, broadband, is available for most of the medium and large establishments we observe.[21] Nonetheless, as with other local services, the competitiveness of local broadband markets varies widely. We note the possibility that broadband may be unavailable in rural areas, in which case internal capabilities will not be a substitute for a lack of local infrastructure. We treat this possibility as an empirical question.[22] Similarly, it is theoretically possible that local resources complement internal capabilities by allowing establishments to better apply local general expertise to firm-specific problems. Again, we treat this possibility as an empirical question. Finally, if establishments

---

[20] Formal identification conditions are discussed in Section 3.3.

[21] Broadband was not available as widely for households in this time period. Because business users were the largest users and tended to be concentrated in clustered areas, they were generally the first to receive new sources of broadband during this time period.

[22] If differences in broadband availability do constrain the value of an establishment's Internet investments in enough cases, then internal capabilities and agglomeration would appear to be complements, leading to a coefficient estimate that rejects hypotheses 3a and 3b.

specialize in different types of capabilities than contract suppliers do, then establishment programmers and market contractors could be complements in the benefit function. If this is so, then we will reject the null hypothesis. Once again, we treat this alternative hypothesis as an empirical question.

For our last hypothesis, we think it is plausible to assume $\frac{\partial^2 B}{\partial ec \partial oc} < 0$: there is substitution between two sources of internal capabilities. This is plausible because both sources work through similar mechanisms and the two sources of capabilities may be fungible. This implies our final hypothesis:

Hypothesis 4: Establishment capability and organizational capability are substitutes. (i.e., $\partial^2 Y / \partial EC \partial OC < 0$)

3. Data

The data used in this study come from the Harte Hanks Market Intelligence Computer Intelligence Technology database (hereafter CI database).[23] This database contains establishment- and firm-level data on characteristics such as number of employees, number of programmers, and use of Internet applications. Harte Hanks collects this information to resell as a tool for the marketing divisions of technology companies. Interview teams survey establishments throughout the calendar year; our sample contains the most current information as of December 2000.

Harte Hanks tracks over 300,000 establishments in the United States. Because we focus on commercial Internet use, we exclude government, military, and nonprofit establishments (mostly in higher education). Our sample from the CI database contains all commercial establishments with over 100 employees—in total 115,671 establishments. Harte Hanks provides one observation per establishment.

We use the 86,879 clean observations with complete data generated between June 1998 and December 2000.[24] Harte Hanks also tracks whether an establishment is affiliated with a larger organization. In total, there are 47,966 distinct organizations, and 7,035 of these have more than one establishment. We employ a strategy of utilizing as many observations as possible for establishments in thinly populated areas. This necessitates routine adjustments for the timing and type of the survey given by Harte Hanks.

3.1. Endogenous Variables

Our analysis focuses on investment in advanced Internet technology that either changes existing internal operations or implements new services. As noted above, we label this investment Within-Establishment Internet, or WEI. We look for indications that an establishment has made investments that involved frontier technologies or substantial co-invention. The threshold for defining *substantial* is necessarily arbitrary within a range.[25] It usually arises as part of other intermediate goods, such as software, computing, or networking equipment. Investment in WEI involves the use of Internet protocols in the input and output of data to and from business applications software. Examples include (1) intranet applications that enable Web access to information stored in business applications software, such as inventory or accounting data and (2) applications that have functionality involving integration with back-end databases (e.g., Web access to a data warehouse).[26]

Firms investing in complex applications such as WEI may do so at only a subset of establishments. As we subsequently discuss in further detail, such activity may cloud the relationship

---

[23] This section provides an overview. For more detail, see Forman, Goldfarb, and Greenstein (2002).

[24] We dropped establishments that did not indicate when they were surveyed and establishments that were not surveyed on information technology. There is a small bias in the dropped observations toward locations where WEI investment is high. The weighting scheme controls for any location and industry bias in the sample.

[25] We tested a number of slightly different definitions and did not find any significant changes to our findings.

[26] To be specific, an establishment is counted as investing in WEI if it invests in one of the following: (1) business application software that involves intensive use of database management systems, such as accounting, sales and marketing, payroll, ERP and MRP, inventory, order processing, and data warehousing; (2) science and research applications used for financial analysis and modeling, CAD/CAM/CAE, data analysis, and engineering; or (3) office applications, such as personnel management, project management, and groupware. See Forman, Goldfarb, and Greenstein (2005a) for more details.

between geography, internal capabilities, and technology use in our establishment-level regressions.

Hence, as a robustness check, we also examine the determinants of two additional measures of innovative behavior. One such measure describes whether the IT personnel at the establishment are using languages that are commonly employed in the building of Internet applications. This variable is equal to one when the establishment indicates the use of any of the following: XML, Visual Basic Script, Perl, Java, CGI, and Active-X. The second variable describes when an establishment has made necessary hardware investments to host Internet applications, and it is equal to one when the establishment has installed a PC server. Both of the measures represent investment in intermediate inputs that are used in the building and hosting of Internet applications and may be duplicated at multiple establishments.

3.2. Exogenous Variables

In Table 2, we provide descriptive statistics. Our measure of location size, z, is a dummy variable that equals one when the establishment is located in a metropolitan statistical area (MSA) with a population over 500,000, which we term *sizeable MSA*. This is the simplest way to represent differences between large cities and small cities and rural locations. In our earlier research, we considered a wide variety of different specifications for the effect of increasing density, and these specifications largely did not affect the results. Here, we employ a dummy for "city" to keep the results stark and easy to interpret. Later we demonstrate that variants on this definition do not affect our qualitative inferences.

We measure mobile total capabilities that can be deployed to build new Internet applications. Our first measure of capabilities is based on the number of programmers. For establishment capabilities (*EC*), we measure the number of programmers located within the establishment. For organizational capabilities (*OC*), we examine only the multi-establishment firms within our sample. We compute the total number of programmers from other establishments within the same firm.[27] These measures quantify the total number

---

[27] In our database, the *programmers* variable is constructed using the following cells: 1-4, 5-9, 10-24, 25-49, 50-99, 100-249, 250-499, and 500 or more. To convert this measure into a continuous variable, we take the midpoint of

of programmers instead of the total quality of programmers or the total cost of programmers. If major urban areas have different average programmer quality than other areas, this may have (ambiguous) implications on the productivity implications of our results for establishment capabilities.[28]

Our second measure of establishment capabilities utilizes principal components factor analysis on three variables—employment, programmers, and a measure of non-Internet software use—to compute a composite variable. This variable provides a more complete description of internal capabilities. The variable on software use is a dummy that is equal to one when the establishment uses internal application development software or when it has familiarity with current object-oriented programming languages such as C or C++.[29] Only one factor has an eigenvalue above one and was therefore the only one retained. This factor explains 50% of the variance. We label this factor the *establishment capability factor.*

For organizational capability, we again conducted principal components factor analysis on five variables: (1) total employment in the organization outside the establishment, (2) number of programmers in the organization outside the establishment, (3) a dummy for whether at least one other establishment in the organization used development software, (4) a dummy for whether at least five other establishments in the organization used development software, and (5) a dummy for whether at least ten other establishments in the organization used development software. Again only one factor has an eigenvalue above one. It explains 59% of the variance. We label this factor the *organizational capability factor.*[30]

Table 2 shows that there are few systematic differences between sizeable MSAs and other areas in the distributions of either organizational capability measure. For the establishment capability measures,

---

each interval and use 500 as the value for the right-censored observations. In our sample, less than 1% (85) of the establishments have a right-censored value for programmers. Qualitative results do not change if a dummy for 500 or more is included.

[28] If urban areas have thicker labor markets for higher quality programmers at the same wage rates as rural areas then that would bias our estimates away from hypothesis 3a and 3b. If urban areas have higher wage rates for higher or same quality programmers, then the bias in our estimate is ambiguous.

[29] We experimented with a variety of alternative measures of software use, including additional measures of development applications as well as measures of complex client/server applications such as enterprise resource planning. These alternative measures gave qualitatively similar results.

[30] We also experimented with higher-order terms for establishment and organizational capabilities, but these had no

both the means and the standard deviations are slightly higher in cities. The differences are small enough not to affect our interpretation of the marginal effects across major cities and other areas.

Our regressions also include controls for establishment employment, controls for whether the establishment is part of a multi-establishment firm, three-digit NAICS dummies, and dummies for the month the survey was conducted.[31] These variables control for many other unmeasured determinants of demand and supply.

4. Statistical Method and Identification

We estimate a probit model of investment in the process innovation, WEI. Our endogenous variable is Y_i, the net value to establishment i of investing in co-invention activity related to use of WEI. The variable Y_i is latent. We observe only discrete choices, namely, whether or not the establishment chooses to invest. The observed decision takes on a value of one or zero.

4.1. Specifying the model

We specify the net benefit function as a linear function of all its parameters except the investment inputs (z_i, EC_i, OC_i), where we allow for interaction terms.[32] Our base specification for the net value from investing in WEI technology is

(2) $Y_i=\alpha_0+\alpha z_i+\beta OC_i+\gamma EC_i+\delta OC_i z_i+\phi EC_i z_i+\lambda OC_i EC_i+\theta_1 me_i+\theta_2 x_i+u_i,$

where α_0 is a constant, α captures the influence of location on investment behavior, β captures the influence of organizational capability, γ captures the influence of establishment capability, δ captures substitution between organization capability and location, ϕ captures substitution between establishment capability and urban location, λ captures substitution between organizational capabilities and

effect.

[31] Establishments were interviewed over a two-year period. Those interviewed toward the end of the period are more likely to have invested. Therefore, we control for the month surveyed.

[32] We also experimented with adding quadratic and other higher-order terms for establishment and organizational capabilities. These terms had no effect, so we show only a parsimonious specification.

establishment capabilities, θ_1 controls for whether the establishment is part of a multi-establishment firm and θ_2 captures the influence of a vector of controls including three-digit NAICS industries, employment, and time of survey. Each parameter is the composite of the difference in net benefits and costs, and some of these are non-linear.

The signs of the coefficients in a probit do not necessarily indicate the sign of the marginal effect (Ai and Norton, 2003). Thus, to identify each of our hypotheses, we focus directly on the signs and magnitudes of the marginal effects, calculated at mean values and using the proper formulas as in Ai and Norton.[33] In all cases, the signs of the coefficients hold for the marginal effects. As a robustness check, we also present evidence on the signs and significance of our results evaluated at other points in the distribution of the variables.

Our base specification treats our variables measuring capabilities as statistically exogenous, and then we will later test for endogeneity. We assume that u is distributed as a normal i.i.d. variable. We weight models by the actual geographic distribution of establishments for industry and size, according to Census County Business Patterns data. If our data under-sample a given two-digit NAICS at a location relative to the Census, then each observation in that NAICS-location is given more importance (for details, see Forman, Goldfarb, and Greenstein, 2002).

4.2 Identification

---

[33] The formulas are: $\frac{\partial Y_i}{\partial z_i} = (\alpha + \delta OC_i + \phi EC_i)\Phi'(.)$, $\frac{\partial Y_i}{\partial OC_i} = (\beta + \delta z_i + \lambda EC_i)\Phi'(.)$, $\frac{\partial Y_i}{\partial EC_i} = (\gamma + \phi z_i + \lambda OC_i)\Phi'(.)$, $\frac{\partial^2 Y_i}{\partial z_i \partial OC_i} = \phi\Phi'(.) + (\alpha + \delta OC_i + \phi EC_i)(\beta + \delta z_i + \lambda EC_i)\Phi''(.)$, $\frac{\partial^2 Y_i}{\partial z_i \partial EC_i} = \delta\Phi'(.) + (\alpha + \delta OC_i + \phi EC_i)(\gamma + \phi z_i + \lambda OC_i)\Phi''(.)$, and $\frac{\partial^2 Y_i}{\partial EC_i \partial OC_i} = \lambda\Phi'(.) + (\beta + \delta z_i + \lambda EC_i)(\gamma + \phi z_i + \lambda OC_i)\Phi''(.)$. It is not straightforward to interpret the standard errors on these marginal effects. They do not combine to give an overall test statistic. The test statistics need to be calculated observation-by-observation. We present the marginal effects and standard errors at mean values for the sample. Forman, Goldfarb, and Greenstein (2005b) present coefficient significance for all the regressions in this paper. In each case the sign of the marginal effect at the mean presented here is identical to the sign of the coefficient on the interaction term.

Identification relies on several assumptions. First, because of our use of industry dummies to control for similarity in benefits and other unobservables, the estimates on capabilities and location are identified from the variation between establishments within the same industry and different type of city, as well as variation between establishments in the same type of city and different industries. Hence, the model is not identified without sufficient variability in supply conditions and innovative activity.

Second, we assume that the location of an establishment is predetermined. This assumption is supported by the unexpectedly rapid diffusion of the Internet, as previously noted. Also, the establishments in our sample are large and from firms with long histories, so, as noted, they did not suddenly relocate when the Internet became available. In other contexts one might want to test for the sensitivity of the inferences to the endogeneity of location choice, but the historical record is so overwhelming we think there are other and more salient statistical issues in this data.

Third, though i.i.d. errors are a routine assumption, they are less innocuous than they seem. They imply that our estimates are not influenced by firm-level simultaneity bias. In other words, we assume that the error in measuring the investment decision of one establishment is independent of the error in every other establishment's decision, including other establishments in the same firm. This assumption is questionable for multi-establishment firms in which a central executive decision maker (e.g., Chief Information Officer) possibly coordinates the choice for each establishment under his domain and allocates mobile internal capabilities across establishments within the firm. If IT investment decisions are centralized and these firms have greater capabilities, then the coefficient estimates for capabilities for multi-establishment firms will be biased. To look for such bias, we estimate the coefficients both with all the data and with a sub-sample of establishments with autonomy to make their own decisions. Furthermore, while we present results without clustering of standard errors, all of our significance results are robust to clustering the standard errors by firm.

Fourth, our base econometric model assumes that establishment capabilities (*EC*) are statistically exogenous. In support of this assumption we note that many of the establishments in our sample maintain large Information Systems groups that support many internal IT services, so that the WEI technologies will be only one of many projects undertaken by such groups. Yet, we recognize that marginal changes in *EC* could make the variable statistically endogenous for the marginal investor. We believe that statistical endogeneity upwardly biases the coefficients on capabilities, that is, in the event that human capital for establishment personnel and IT capital decisions are co-determined. If this arises, it biases our establishment capability estimates in favor of Hypothesis 2b. We believe it is less likely to bias our organizational capability estimates.

We take several actions to control for this form of endogeneity. As was previously noted, we re-estimate our regressions over a subsample in which the establishments have autonomy to make IT investment decisions. If the bias on Hypothesis 2b is large, then this subsample should alter the estimate on organizational capabilities, but not establishment capabilities. In other words, if the bias is large, the result is more likely to hold for Hypothesis 2b in this subsample, but it is not more likely for Hypothesis 2a. We also estimate capabilities in different ways to check for robustness.

We next present results of instrumental variables regressions that use capabilities of other establishments and organizations in the same industry as instruments for establishment and organizational capability. We also instrument for endogeneity of the location-capability interaction variable with capability of other establishments in other industries in the same location. While we do present results of instrumental variables regressions that control for potential endogeneity between establishment and organization capability levels, we have no reason to believe the endogeneity affects the interaction between capabilities and location. As we show below, our results support this belief because we find no qualitative differences in the core results of a number of different specifications.

Fifth, we are interested in the sign of our substitution marginal effects. Our statistical null hypothesis is that these are equal to zero and we search for evidence to reject zero in favor of our alternative hypothesis of substitution. To understand the importance of this null, consider one alternative null that we regard as *ex ante* less plausible, i.e., there is (a) *complementarity* between *EC* and *OC*, and (b) the level for *EC* and *OC* are endogenously chosen to maximize *only* the return from WEI investment. The combination of (a) and (b) generates a scenario that fits a type of model analyzed by Arora and

Gambardella (1990), Arora (1996), and Athey and Stern (2003) in the context of endogenous organizational design decisions with observable outcomes. They show that unobserved heterogeneity in the endogenous complementary inputs creates challenges for statistical inference under the null hypothesis of complementarity, particularly in the face of common forms of measurement error. In such a setting, it is difficult to articulate both *necessary* and *sufficient* conditions to infer complementarity. However, all authors articulate *necessary* conditions for maintaining the null of complementarity.

We examine substitutability, but some of the reasoning for this different situation informs our approach. As we show below, most of our data will reject the *necessary* conditions for complementarity between *EC* and *OC*, even controlling for endogeneity.[34] Relatedly, and as argued by these authors, at most we can make inferences on necessary conditions for substitutability, not sufficient conditions.

Generally, the probit model limits inference. As in any probit model, we do not observe the variance of u_i, so we only infer the sign of coefficients, the relative size compared to each other within the relevant range, and their partial elasticity with respect to the probability of innovating at the threshold value. Unlike a conventional production or cost function estimate, we cannot infer the level of costs or benefits except in terms of the contribution to the probability of an establishment choosing to invest in innovative activity for a marginal investor. At most, we can infer whether the estimated direction of the net benefit function with respect to variables is consistent with predictions from theory. We can also infer whether the estimated direction is consistent with substitution/complementarity among inputs under the null. Though such findings are necessary but not sufficient for inference, given the novelty of the question and setting, we believe these partial inferences still are interesting.

Sixth, and finally, our statistical approach relies on the accumulated weight of many different tests, not any specific estimate. Any specific estimate is vulnerable to concerns that unmeasured demand correlates with our measures of costs, especially those in Hypothesis 2b, as was noted above.

---

[34] As the tables below show, the rejection is strong under most specifications. Occasionally, the rejection is weak

Nevertheless, we gain confidence from estimating the sensitivity of the statistical inference to a variety of (1) different estimation corrections, as described above, (2) different implementations for the endogenous and exogenous variables, as we describe below, and (3) different samples, as we describe below.

5. Results

5.1. How Does Location Affect the Contribution Of Internal Capabilities?

In this section, we first show the impact of changes in location size and internal capabilities on IT investment. We then examine the interaction between the roles of internal capabilities and of cities. The results in Columns (1) through (4) of Table 3 show the coefficients of the main results; those in Columns (5) and (6) show marginal effects.

We first show that urban leadership holds for WEI, supporting Hypothesis 1. Specifically, an increase in location size has a significantly positive effect on investment in WEI technology. The results in Columns (1) and (2) show that establishments located in sizeable MSAs are significantly more likely to invest, other things being equal. The marginal effects based on the full specification in columns (5) and (6) suggest that, on average, being in a higher-density area increases the likelihood of investing 3.26 or 4.55 percentage points, depending on the measure. This is a large amount, given that the percentage of firms investing is just 11.92%.

Increases in capability significantly increase the probability of investment, which supports Hypotheses 2a and 2b. Table 3 shows that increases in establishment capabilities have a significantly positive effect (at the 1% level) on investment in WEI in all specifications. The marginal effect estimates in columns (5) and (6) suggest that a one-standard-deviation increase in the log of the number of programmers at the establishment increases average probability by 7.99 percentage points.[35] A one-standard-deviation (equivalent to one unit by construction) increase in our composite establishment

---

because the confidence interval encompasses the possibility of a positive coefficient under a null that it is positive.

[35] These quantities are computed by multiplying the marginal effect by the change in the variable.

capability factor has a similar effect, increasing the use of WEI by 8.12 percentage points. In sum, increases in internal establishment capabilities significantly reduce costs.

Increases in organizational capabilities have a similar, but smaller, impact. A one-standard-deviation increase in the log of the number of programmers increases the likelihood of investing in WEI by 1.21 percentage points, which is statistically significant at the 1% level. Increasing the composite organizational capability factor by one standard deviation increases the likelihood of investing in WEI by 1.54 percentage points. Although these effects are smaller than those for establishment capabilities, they are still large when compared to the sample rates for WEI.

We next examine the extent to which internal capabilities and cities are substitutes. The results in Columns (3), (4), (5) and (6) of Table 3 present the main results of our paper. There is considerable evidence that internal capabilities substitute for the benefits of city location, supporting Hypotheses 3a and 3b. The key effects are all in the expected direction. The interaction coefficients are significant at the 5% level and the marginal effects are significant for over half of the observations. Later, in Tables 5 through 8, we show the results of a number of robustness checks. The key effects of high-density location and internal capabilities are similar to those in Columns (5) and (6). The results in column (5) show that establishments outside sizeable MSAs benefit 1.72 percentage points more than establishments in sizable MSAs from a one-standard-deviation increase in the log of organizational programmers. Similarly, establishments outside sizeable MSAs benefit 1.51 percentage points more from a one-standard-deviation increase in the log of establishment programmers. Interestingly, while establishment programmers have a much stronger effect than organizational programmers (marginal effects of 7.99 percentage points versus 1.21 percentage points), the extent of substitution between cities and establishment capabilities is almost equal to that of cities and organizational capabilities.[36] This suggests that while only a fraction of

[36] We also examined the robustness of our marginal effects to changes in where the marginal effects are evaluated. We examined the distribution of marginal effects for nonzero capabilities, since capabilities are expected to increase the likelihood of investment only when they are nonzero. For programmers (column 5), of the 8,739 establishments

organizational capabilities are mobile, they perform similar activities to establishment capabilities and are able to substitute similarly for cities.[37] The marginal contribution of internal resources to innovation appears to be lower in cities than in other areas.

Figure 1 presents another view of the main results. It presents the predicted probabilities of investing in WEI using the results in Column (3) of Table 3 under different combinations of location size and internal capabilities. In this figure, we use the log of programmers as our measure of internal capabilities; results using our composite measure of capabilities are qualitatively similar. We now discuss Figure 1a, which presents the results for organizational capabilities, and we compare the results briefly to the results for establishment capabilities in Figure 1b, which are qualitatively similar: Figure 1a shows that establishments located in sizeable MSAs have a greater likelihood of WEI investment (Hypothesis 1). When the firm has no internal organizational capabilities, location in a sizeable MSA increases the probability of investing considerably, from 11.6% to 16.1%.[38] This 4.5 percentage-point increase is the difference at the intercept. Moreover, Figure 1a provides support for Hypothesis 2a: the upward sloping lines show that the probability of action increases as organizational capabilities increase, whether or not the establishment is in a sizeable MSA.

Figure 1a also demonstrates how the prediction of Hypothesis 3a shapes behavior: The curve depicting establishments in sizeable MSAs is flatter than that for other establishments. The marginal impact of increasing organizational capabilities is lower for organizations in sizeable MSAs. Changing from low (OC_i = 0) to average (OC_i = 1.738) capabilities increases the probability of investing in WEI by 1.5 percentage points for establishments in low-density areas, and increases the probability by 0.5

---

with positive organization and establishment programmers, over 99% of organizational capabilities interactions are negative and significant at the five-percent level and over 90% of establishment capabilities interactions are negative and significant at the five-percent level. For capabilities defined by factors (column 6), of the 8,739 establishments with positive organization and establishment capabilities (as defined by programmers), 74% of the organizational capabilities interactions are negative and significant at the five-percent level and over 95% of the establishment capabilities interactions are negative and significant at the five-percent level.

[37] These results are also robust to allowing organizational capabilities in the same location (MSA or state) or in the

percentage points for establishments in sizeable MSAs. Organizational capabilities, however, are unable to completely substitute for the benefits of an urban location. Even for very capable organizations that are one full standard deviation above the mean (OC_i = 4.028), rates are 2.3 percentage points higher in sizeable MSAs than in rural areas.[39] In summary, we find no evidence that organizational capabilities and cities are complements.

These results illustrate the importance of establishment capabilities, organizational capabilities, and location size in reducing costs. Although it is difficult to compare them because of differing metrics, the slopes of the organizational capabilities lines in Figure 1a are not nearly as steep as the slopes of the establishment capabilities lines in Figure 1b. For establishments located outside of a sizeable MSA, an increase in establishment capabilities from zero to their mean level (EC_i = 0.510) increases the likelihood of investing in WEI by 4.1 percentage points, while an increase in establishment capabilities from zero to their mean level (OC_i = 4.028) increases the probability by 1.6 percentage points. These results reflect the coefficient estimates on establishment and organizational capabilities in the first two rows of Table 3.

Moreover, these figures also are suggestive about the relative importance of internal versus local external capabilities in lowering innovation costs. For example, an establishment with high organizational capabilities (one standard deviation above the mean) that is located in a low-density location has a lower rate of investment (15.0%) than a similar establishment with zero organizational capabilities that is located in a high-density location (16.1%). Yet, an establishment with high establishment capabilities in a low-density location has a higher rate (24.6%) than a similar establishment with mean values for establishment capability and a high-density location (19.5%).

Overall, these results suggest that internal capabilities are substitutes for cities when investing in

---

same industry (three- or six-digit NAICS) to have a separate effect on costs.

[38] Simulations assume establishment capabilities are equal to zero.

[39] To ensure the results are not driven by functional form, we estimated a quadratic specification of capabilities. The coefficients on establishment capabilities squared, organizational capabilities squared, and their interactions with the city dummies are both statistically and economically insignificant. Results are available upon request.

complex technologies. Yet, they also suggest that internal establishment capabilities are more effective at lowering costs than are organizational capabilities.

5.2. Substitutability of Organizational Capability and Establishment Capability

In Table 4, we present evidence that establishment and organizational capabilities are substitutes (Hypothesis 4). No matter how we measure capabilities, the interaction between organizational and establishment capabilities is negative and significant at the one-percent level.[40] The results in Column (5) show that an increase in the log of organization programmers by one standard deviation will decrease the marginal effect of establishment capabilities by 0.89 percentage points when establishment capabilities are at mean values. This magnitude is of moderate size when compared to the effect of establishment capabilities on investment. By comparison, an increase in the log of establishment programmers from 0 to its mean value will increase the probability of investment by 4.15 percentage points when organizational capabilities are equal to zero, that is, when the establishments have no organizational capabilities.

The results for our composite measure are similar. Assuming a value for our composite establishment factor of one, an increase of one standard deviation in organizational capabilities will decrease the marginal effect of establishment capabilities by 5.16 percentage points. By comparison, an increase in the establishment capability composite variable from 0 to 1 will increase the probability by 8.68 percentage points. Overall, these results suggest that there exists significant substitution between establishment and organizational capabilities in establishment investment decisions in WEI.

5.3. Robustness Checks

In Tables 5 through 8, we show the results of a number of robustness checks. In Table 5, we explore the exogeneity assumptions relating to location and capability. In Table 6, we explore different

[40] Capability is potentially endogenous if establishments in weak organizations hire more programmers to implement a *planned* investment in process innovations. Nevertheless, the negative correlation on the interaction term still suggests substitutability between establishment and organizational capabilities irrespective of the direction of causality. Including a three-way interaction between z_i, OC_i, and EC_i yields nearly identical results. The

definitions of location size and investment decision. In Table 7, we examine whether our results are driven by industry competition. In Table 8, we examine whether our results apply to both services and manufacturing.

Columns (1) and (2) of Table 5 only use data from establishments in multi-establishment firms that reported that their IT investment decisions are made locally at the establishment. Our baseline specification implicitly assumes that the error terms of establishment decisions are independently distributed. This assumption is particularly likely to hold for this subset of establishments. Although this reduces our sample size considerably, thereby leading to a loss of significance in some cases, the signs of all results remain the same.

For the remainder of Table 5, we use instrumental variables techniques to examine the assumption of exogenous capabilities. In particular, instrumental variables probit regressions were used. Following Maddala (1983, p. 247–52), we used Amemiya Generalized Least Squares.[41] We define five instruments. First, we instrument for a firm's establishment capabilities with the establishment capabilities of other establishments in other firms in the same two-digit NAICS industry in the other locations that the firm has an establishment. Second, similarly, we instrument for a firm's organizational capabilities with the organizational capabilities of other establishments in other firms in the same two-digit NAICS industry in other locations where the firm has an establishment. These instruments should be correlated with the capabilities of an establishment but not with the propensity of the establishment to invest in WEI, conditional on its industry. Third and fourth, we use two instruments for the interaction of establishment capability and sizeable MSA. First, we interact the previous instrument for establishment capabilities (i.e., instrument 1) with a dummy for sizeable MSA. Second, we use establishment capabilities at other establishments in other industries in the same location. These capabilities will be

coefficient on the interaction term is economically small and statistically insignificant.

[41] In the first stage, the endogenous variables are treated as a linear function of the instruments and the exogenous variables. The second stage probit uses the predicted values for the endogenous variables from the first stage.

affected by the same local supply conditions but will not be directly correlated with the decision to invest. We construct our fifth instrument, the interaction of organizational capabilities and sizeable MSA, by interacting the above instrument for organizational capabilities (i.e., instrument 2) with a sizeable MSA dummy.[42] We therefore have five main instruments for four potentially endogenous variables.

In Columns (3) and (4), we use instruments for the establishment and organizational capabilities variables. We do not instrument for the interactions of these variables with location. While significance on the interaction term for establishment capability is lost and the significance on the interaction term for organization capability is only 89.9 percent when capability is defined by programmers, all other significance remains and the signs do not change. The main results do not appear to be driven by the endogeneity of the capabilities variables. Nevertheless, it is also possible that establishments or organizations in particular locations are more capable. Therefore, to examine robustness, in Columns (5) and (6), we use all five instruments for the four potentially endogenous variables, namely, establishment capability, organizational capability, and their interactions with being in a high-density area. Establishment and organizational capability are still positively associated with investment, and while some significance is lost, the marginal effect of the interaction of being in a high-density area with either capability measure is negative. In summary, the results are robust to instrumental variables techniques.

In Table 6, we explore the robustness of the variable definitions. Columns (1) through (4) show that the results are robust to different definitions of what constitutes a city. In our base specification, we define a city as an MSA with a population of over 500,000. Columns (1) and (2) define a city as any MSA. Columns (3) and (4) define a city as a location with a population density greater than the sample median (610 people per square mile). We also ran regressions that use three kinds of MSA: small (<500,000), medium (500,000-1 million), and large (over 1 million). The establishment capabilities interaction is largely monotonic in location size. The main divide for organizational capabilities is

---

[42] We do not use the organizational capabilities equivalent of the second establishment capabilities instrument,

between MSAs and non-MSAs.[43] In all cases, the qualitative results remain the same.[44] Columns (5) through (8) of Table 6 check the interaction of location and capabilities on different technologies. The results for Internet development languages (columns (5) and (6)) and PC servers (columns (7) and (8)) are similar to the results for WEI.

Sizeable MSAs not only have greater external resources, but may also have stronger competition in industries that sell non-tradeable goods and services. Thus, our city dummy may also capture the effects of competition. To examine this hypothesis, the results in Table 7 show whether our conclusions are robust to the inclusion of variables measuring competition. In Columns (1) and (2), we include the total number of other establishments in the establishment's same six-digit NAICS and county, and in Columns (3) and (4), we include the total employment in the same six-digit NAICS and county. We also interact these variables with establishment and organizational capabilities.

If our city dummy proxies for external resources beyond the effects of competition, our core results should remain qualitatively unchanged. They do. In particular, Columns (1) and (3) show that our results using establishment and organizational programmers remain unchanged, regardless of whether we include establishments or employment in the same industry-county. Columns (2) and (4) show that our results using establishment factors and organizational factors are qualitatively the same: signs and significance continue to hold on the interaction with establishment capabilities, while signs continue to hold on the interaction with organizational capabilities.[45] As a further robustness check, we also examined whether these results were robust to estimating these regressions using only establishments in service industries. Service sector establishments will more likely be influenced by competition with local

---

because it is not clear how organizational capabilities of establishments in a city will be correlated.

[43] The coefficients on the interactions of establishment capabilities and city sizes are -0.0357, -0.0796, and -0.0989 for small, medium, and large MSAs, respectively (non-MSA is the base). For organizational capabilities, the coefficients are -0.0408, 0-.0408, and -0.0456, respectively. Full results are available from the authors on request.

[44] The inference also does not change with other city definitions, including MSAs with a population of over one million, a continuous population measure, and a continuous density measure. Results are available upon request.

[45] The interaction with organizational capabilities in column (4) is significant at the 87.46 percent level.

establishments. Again, these results are qualitatively the same.[46] Overall, the proxies for competition have little impact on our core results.

We also examined whether our results were specific to any particular sector of the economy. The results in Table 8 show that the substitutability between internal and external factors holds in both manufacturing and services.

6. Discussion and Conclusions

In this study, we find extensive statistical evidence of localization of substitution between internal and external inputs into innovation. We show that establishments located in large urban areas innovate as if they face fewer constraints and have lower costs. We also find a symmetric role for internal capabilities: establishments that are in firms with a greater number of IT personnel invested in WEI technology more frequently, as did those with prior experience with related non-Internet applications. Overall, we conclude that the marginal contribution of internal capabilities to investment and co-invention in a process innovation is lower for establishments in cities than for establishments elsewhere.

More generally, we find that establishments engaged in co-inventive activity draw upon a variety of resources: internal establishment capabilities, internal organizational capabilities, and external purchased services. We provide a framework for measuring the contribution of each of these channels to new process innovation. In contrast to prior work, we find that all of these channels are substitutes for one another as inputs into innovative investment.

These results have implications for understanding the sources of co-inventive activity required for process innovation, as well as managerial implications for the optimal location of innovative activity. In particular, these findings suggest that the advantages of agglomeration will be most important for single-establishment firms that have been unable to develop internal capabilities for innovative activity. The findings are consistent with those of researchers who have argued that agglomeration of firms with

[46] For brevity, these results are not included in any tables. They are available from the authors upon request.

similar input demands can provide benefits through the provision of complementary third party services. These benefits will be most valuable among small firms and for firms in young or infant industries, where internal capabilities and business processes are still being developed.[47] More generally, the findings are consistent with Saxenian's (1996) observation that managers at firms that anticipate innovating will be better off locating near other firms that are innovating.

Our findings are also consistent, albeit more speculatively, with those of researchers who have argued that as industries mature and average firm size increases, there is less need for the complementary resources and knowledge transfer found in cities. As a result, firms may relocate to shape their innovative activities (Furman, Kyle, Cockburn, and Henderson, 2005), or economize on transportation costs or save on wages (Duranton and Puga, 2001). Nevertheless, caution is warranted in following this line of reasoning. We have examined only one reason why firms would desire urban locations. Firms may agglomerate in the same location for a variety of related reasons: knowledge transfer, labor market pooling, knowledge spillovers, and transportation costs among them.

Finally, our results have implications for ongoing research about outsourcing. It is a comparatively unexplored theme in outsourcing research whether the location of an establishment shapes the propensity of establishments to use market-mediated external channels. Our evidence about investment in WEI suggests location is a determinant of the outsourcing decision. Furthermore, our results suggest that location will matter more when the firm has fewer internal capabilities. Using direct measures of outsourcing to better understand the roles of location and internal capabilities in this context is an interesting subject for future research.

[47] For example, our results are consistent with the global distribution of firms engaged in software development. While small independent firms engaged in software development in countries such as India and Ireland cluster in a relatively small number of areas, the location of large U.S. firms that produce software products or services is distributed throughout the U.S. and worldwide.

References

Ai, Chunrong, and Edward C. Norton. 2003. Interaction terms in logit and probit models. *Economics Letters* 80: 123-129.

Arora, Ashish. 1996. Testing for complementarities in reduced-form regressions: A note. *Economics Letters* 50(1): pages 51-55.

Arora, Ashish and Alfonso Gambardella. 1990. Complementarity and External Linkages: The Strategies of the Large Firms in Biotechnology. *Journal of Industrial Economics* 38(4): 361-79.

——— and ———. 1994. Evaluating technological information and utilizing it: Scientific knowledge, technological capability and external linkages in biotechnology. *Journal of Economic Behavior and Organization* 24(1): 91-114.

Arora, Ashish, Alfonso Gambardella, and Salvatore Torrisi. 2004. In the Footsteps of Silicon Valley? Indian and Irish Software in the International Division of Labor. In Timothy Bresnahan and Alfonso Gambardella (eds). *Building High-Tech Clusters: Silicon Valley and Beyond.* Cambridge: Cambridge University Press.

Astebro, Thomas. 2002. Noncapital investment costs and the adoption of CAD and CNC in U.S. metalworking industries. *RAND Journal of Economics* 3(4), 672-688.

Attewell, Paul. 1992. Technology Diffusion and Organizational Learning: The Case of Business Computing. *Organizational Science* 3 (February): 1–19.

Athey, Susan, and Scott Stern. 2003. An Empirical Framework for Testing Theories about Complementarities in Organizational Design. Working Paper, Northwestern University, Evanston, IL.

Austin, Robert D., Mark J. Cotteleer, Richard L. Nolan, and George Westerman. 1999. Tektronix, Inc.: Global ERP Implementation. Harvard Business School Case 9-699-043.

Bendolay, Elliott and F. Robert Jacobs. 2005. *Strategic ERP Extension and Use.* Stanford University Press.

Bresnahan, Timothy, Erik Brynjolfsson, and Lorin Hitt. 2002. Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence. *Quarterly Journal of Economics* 177: 339-376.

Bresnahan, Timothy and Alfonso Gambardella. 1998. Division of Labor and the Extent of the Market. In *General Purpose Technologies and Economic Growth*, ed. Elhanan Helpman, Cambridge: MIT Press.

——— and ———. 2004. *Building High-Tech Clusters: Silicon Valley and Beyond.* Cambridge: Cambridge University Press.

Bresnahan, Timothy, and Shane Greenstein. 1996. Technical Progress in Computing and in the Uses of Computers. *Brookings Papers on Economic Activity, Microeconomics*: 1-78.

Bresnahan, Timothy, and Manuel Trajtenberg. 1995. General Purpose Technologies: ‘Engines of Growth’? *Journal of Econometrics* 65 (1): 83-108.

Brynjolfsson, Erik, and Lorin M. Hitt. 2000. Beyond Computation: Information Technology, Organizational Transformation and Business Performance. *Journal of Economic Perspectives* 14(4): 23-48.

Cassiman, Bruno, and Reinhilde Veugelers. 2006. In Search of Complementarity in Innovation Strategy: Internal R&D and External Knowledge Acquisition. *Management Science* 52(1): 68-82.

Cohen, Wesley M. and Daniel A. Levinthal. 1990. Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly* 35: 128–52.

Columbo, Massimo and Rocco Masconi. 1995. Complementarity and Cumulative Learning Effects in the Early Diffusion of Multiple Technologies. *Journal of Industrial Economics* 43(1): 13-48.

David, Paul A. 1969. A Contribution to the Theory of Diffusion, Memorandum No. 71, Stanford Center for Research in Economic Growth, Stanford University.

Davis, James C. and J. Vernon Henderson. 2002. Headquarters’ Location Decisions, Working Paper, Department of Economics, Brown University.

Duranton, G. and D. Puga. 2001. Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products. *American Economic Review* 91(5): 1454–77.

——— and ———.2002. From Sectoral to Functional Specialization, NBER Working Paper #9112, Cambridge, MA.

Forman, Chris. 2005. The Corporate Digital Divide: Determinants of Internet Adoption. *Management Science* 51: 641-654.

Forman, Chris, Avi Goldfarb, and Shane Greenstein. 2002. Digital Dispersion: An Industrial and Geographic Census of Commercial Internet Use, NBER Working Paper #9287, Cambridge, MA.

———, ———, and ———. 2005a. How Did Location Affect Adoption of the Commercial Internet? Global Village vs. Urban Leadership. *Journal of Urban Economics* 58(3): 389-420.

———, ———, and ———. 2005b. Technology Adoption In and Out of Major Urban Areas: When Do Internal Firm Resources Matter Most, NBER Working Paper #11642, Cambridge, MA.

Furman, Jeffrey L., Margaret K. Kyle, Iain Cockburn, and Rebecca Henderson. 2005. Public & Private Spillovers, Location, and the Productivity of Pharmaceutical Research. Working Paper, School of Management, Boston University.

Greenstein, Shane. 2001 Commercialization of the Internet: The Interaction of Public Policy and Private Action. in (eds) Adam Jaffe, Josh Lerner, and Scott Stern, *Innovation Policy and the Economy.*

Griliches, Zvi. 1957. Hybrid Corn: An Exploration in the Economics of Technical Change. *Econometrica* 25 (October): 501-522.

Henderson, J. Vernon. 2003. Marshall's Scale Economies. *Journal of Urban Economics* 53(1): 1–28.

Hubbard, Thomas N. 2000. The Demand for Monitoring Technologies: The Case of Trucking. *Quarterly Journal of Economics* 116: 533-560.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. 1993. Geographic localization of knowledge spillovers as evidenced in patent citations. *Quarterly Journal of Economics* 108(3): 577-598.

Karshenas, Massoud and Paul Stoneman. 1993. Rank, Stock, Order, and Epidemic Effects in the Diffusion of New Process Technologies: An Empirical Model. *Rand Journal of Economics* 24(4): 503–28.

Kelley, Maryellen and Susan Helper. 1999. Firm Size and Capabilities, Regional Agglomeration, and the Adoption of New Technology. *Economics of Innovation and New Technology* 8: 79–103.

Kenney, Martin. 2003. The Growth and Development of the Internet in the United States. in (ed) Bruce Kogut, *The Global Internet Economy*. Cambridge: MIT Press.

Kolko, Jed. 1999. Can I Get Some Service Here? Information Technology, Service Industries, and the Future of Cities, Working Paper, Department of Economics, Harvard University.

Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge England: Cambridge University Press.

Marschak, T. 2004. Information Technology and the Organization of Firms. *Journal of Economics and Management Strategy* 13: 473-515.

Mowery, David C. and Timothy S. Simcoe. 2002. The Origins and Evolution of the Internet. in *Technological Innovation and Economic Performance,* (eds.) R. Nelson, B. Steil, and D. Victor. Princeton, NJ: Princeton University Press.

Nolan, Richard L. 2001. Cisco Systems: Building Leading Internet Capabilities. Harvard Business School Case 9-301-133.

Porter, Michael E. 1998. Clusters and competition: New agendas for companies, governments, and institutions. in *On Competition,* (ed.) Michael E. Porter. Boston: Harvard Business School Press.

Saxenian, Annalee. 1996. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press.

Singh, Jasjit. 2004. Multinational Firms and Knowledge Diffusion: Evidence Using Patent Citation Data. Working Paper, INSEAD, Singapore.

Stigler, George. 1951. The division of labor is limited by the extent of the market. *Journal of Political Economy* 59: 185-193.

Vernon, Raymond. 1963. *Metropolis 1985.* New York: Doubleday.

Table 1: Variable definitions and glossary

Variable	Definition
Y	A latent variable for net benefit of undertaking the investment project
$\overline{B}$	Gross benefit of undertaking the investment project
$\overline{C}$	Total cost of undertaking the investment project
B()	Gross benefit of the investment without the establishment-specific benefit random variable
C()	Total cost of the investment that relates to internal capabilities
c()	Monetary costs of purchased services
u^B	Establishment-specific random benefit variable
u^C	Establishment-specific random cost variable
u	Establishment-specific random variable equal to u^B-u^C
ec	Establishment capabilities invested in the project
oc	Organizational (firm) capabilities invested in the project
ps	Purchase of local services to facilitate the investment project
me	Indicator variable for multi-establishment status
x	Establishment characteristics such as number of employees and industry
t	Time
z	Location size (defined by population size or population density)
EC	Total establishment capabilities
OC	Total organizational capabilities
$g(t)$	Relationship between total cost of undertaking the investment project and time
$h(t)$	Relationship between gross benefit of undertaking the investment project and time
α_0	The regression constant
α	The estimated coefficient on population size, z_i
β	The estimated coefficient on organizational capability, OC_i
γ	The estimated coefficient on establishment capability, EC_i
δ	The estimated coefficient on organizational capability times population size, OC_iz_i
ϕ	The estimated coefficient on establishment capability times population size, EC_iz_i
λ	The estimated coefficient on organizational capability times establishment capability, OC_iEC_i
θ_1	The estimated coefficient on being a multi-establishment firm, me_i
θ_2	The estimated coefficients on the vector of control variables, x_i
Φ	The Normal cumulative distribution function (c.d.f.)
WEI	"Within-Establishment Internet"—process innovations that reduce the costs of communications internal to the establishment based on Internet (TCP/IP) technologies.

Table 2: Descriptive Statistics

	Mean	Standard Deviation	Minimum	Maximum	Number of Observations
Full Data Set					
Log(Programmers in Organization+1)	1.7382	2.2898	0	8.5535	86,879
Log(Programmers in Establishment +1)	0.5100	1.0189	0	6.2166	86,879
Organizational Capability Factor	0	1	-0.4433	5.9924	86,879
Establishment Capability Factor	0	1	-0.3768	44.6452	86,879
MSA Population over 500,000 Dummy	0.7371	0.4402	0	1	86,879
Multi-Establishment Firm Dummy	0.4479	0.4973	0	1	86,879
Log(Establishment Employment)	5.3376	0.7248	4.605	10.933	86,879
WEI	0.1192	0.3240	0	1	86,879
Use an Internet Language	0.0658	0.2480	0	1	86,879
Have a PC Server	0.5513	0.4974	0	1	86,879
MSA Population over 500,000					
Log(Programmers in Organization+1)	1.7865	2.3393	0	8.5535	64,038
Log(Programmers in Establishment +1)	0.5606	1.0899	0	6.2166	64,038
Organizational Capability Factor	-0.0100	0.9578	-0.4433	5.9924	64,038
Establishment Capability Factor	0.0385	1.0877	-0.3768	44.6452	64,038
Other areas					
Log(Programmers in Organization+1)	1.6029	2.1393	0	8.5535	22,841
Log(Programmers in Establishment +1)	0.3682	0.7691	0	6.2166	22,841
Organizational Capability Factor	0.0281	1.1094	-0.4433	5.9924	22,841
Establishment Capability Factor	-0.1080	0.6861	-0.3768	36.7297	22,841
Single Establishment Firms					
Log(Programmers in Establishment +1)	0.5330	0.9668	0	6.2166	47,966
Establishment Capability Factor	-0.0347	0.8381	-0.3768	36.7297	47,966
Multi-Establishment Firms					
Log(Programmers in Establishment +1)	0.4817	1.0791	0	6.2166	38,913
Establishment Capability Factor	0.0428	1.1676	-0.3768	44.6452	38,913

Table 3: Main Results

	Coefficients				Marginal Effects at Mean Values		
	Direct Effect Only		Direct Effect and Interaction Effect			Direct Effect and Interaction Effect (based on columns 3 and 4)	
	(1)	(2)	(3)	(4)		(5)	(6)
Coefficient	Capability Defined by Programmers	Capability Defined by Factors	Capability Defined by Programmers	Capability Defined by Factors	Marginal Effect	Capability Defined by Programmers	Capability Defined by Factors
OC	0.0152	0.0427	0.0399	0.0822	$\frac{\partial Y_i}{\partial OC_i}$	0.0053	0.0154
	(0.0057)**	(0.0102)**	(0.0082)**	(0.0192)**		(0.0016)**	(0.0034)**
EC	0.2670	0.2091	0.3395	0.3823	$\frac{\partial Y_i}{\partial EC_i}$	0.0784	0.0812
	(0.0085)**	(0.0197)**	(0.0193)**	(0.0252)**		(0.0024)**	(0.0054)**
OC*z			-0.0285	-0.0480	$\frac{\partial^2 Y_i}{\partial z_i \partial OC_i}$	**-0.0075**	**-0.0145**
			(0.0078)**	(0.0203)*		**(0.0022)****	**(0.0067)***
EC*z			-0.0804	-0.1847	$\frac{\partial^2 Y_i}{\partial z_i \partial EC_i}$	**-0.0148**	**-0.0541**
			(0.0206)**	(0.0314)**		**(0.0057)****	**(0.0104)****
z (MSA Population > 500,000 Dummy)	0.0995	0.1308	0.2070	0.1377	$\frac{\partial Y_i}{\partial z_i}$	0.0326	0.0455
	(0.0180)**	(0.0182)**	(0.0257)**	(0.0186)**		(0.0051)**	(0.0062)**
Multi-Establishment Firm Dummy	0.1126	0.1481	0.1100	0.1464	$\frac{\partial Y_i}{\partial me_i}$	0.0308	0.0483
	(0.0274)**	(0.0189)**	(0.0273)**	(0.0188)**		(0.0077)**	(0.0062)**
Log(Establishment Employment)	0.2318		0.2299		$\frac{\partial Y_i}{\partial \ln(empl)_i}$	0.0644	
	(0.0132)**		(0.0132)**			(0.0037)**	
Observations	86871	86871	86871	86871		86871	86871
LL	-24550.40	-25914.41	-24528.56	-25861.03		-24528.56	-25861.03

For columns (1) through (4) standard errors are in parentheses; for columns (5) and (6) mean standard errors are in parentheses. All regressions are weighted to reflect the actual geographic distribution of establishments from County Business Patterns and include dummy variables for three-digit NAICS and month of survey. Significance levels do not change if standard errors are clustered by firm. Key results in **bold.**

+significant at 90% confidence level.

*significant at 95% confidence level.

**significant at 99% confidence level.

Table 4: Are Establishment Capabilities and Organizational Capabilities Substitutes?

	Coefficients				Marginal Effects at Mean Values (based on columns 2 and 4)		
	Capability Defined by Programmers		Capability Defined by Factors			Capability Defined by Programmers	Capability Defined by Factors
Variable	(1)	(2)	(3)	(4)	Marginal Effect	(5)	(6)
OC	0.0501	0.0624	0.0775	0.0972	$\frac{\partial Y_i}{\partial OC_i}$	0.0109	0.0256
	(0.0046)**	(0.0068)**	(0.0112)**	(0.0188)**		(0.0012)**	(0.0036)**
EC	0.3234	0.3764	0.2328	0.3913	$\frac{\partial Y_i}{\partial EC_i}$	0.0814	0.0868
	(0.0073)**	(0.0141)**	(0.0157)**	(0.0248)**		(0.0017)**	(0.0045)**
OC*EC	-0.0241	-0.0231	-0.0457	-0.0438	$\frac{\partial^2 Y_i}{\partial OC_i \partial EC_i}$	**-0.0039**	**-0.0516**
	(0.0018)**	(0.0018)**	(0.0088)**	(0.0086)**		**(0.0005)****	**(0.0095)****
OC*z		-0.0162		-0.0255	$\frac{\partial^2 Y_i}{\partial z_i \partial OC_i}$	-0.0038	-0.0061
		(0.0064)*		(0.0198)		(0.0018)*	(0.0065)
EC*z		-0.0639		-0.1708	$\frac{\partial^2 Y_i}{\partial z_i \partial EC_i}$	-0.0123	-0.0484
		(0.0143)**		(0.0289)**		(0.0039)**	(0.0095)**
z (MSA Population > 500,000 Dummy)	0.0672	0.1451	0.1271	0.1364	$\frac{\partial Y_i}{\partial z_i}$	0.0238	0.0446
	(0.0155)**	(0.0224)**	(0.0180)**	(0.0185)**		(0.0045)**	(0.0060)**
Multi-Establishment Firm Dummy	0.0692	0.0686	0.1351	0.1340	$\frac{\partial Y_i}{\partial me_i}$	0.0193	0.0438
	(0.0199)**	(0.0199)**	(0.0189)**	(0.0189)**		(0.0056)**	(0.0062)**
Log(Establishment Employment)	0.2540	0.2518			$\frac{\partial Y_i}{\partial \ln(empl)_i}$	0.0709	
	(0.0089)**	(0.0089)**				(0.0025)**	
Observations	86872	86872	86871	86871		86872	86871
LL	-25199.06	-25185.82	-25823.42	-25779.55		-25185.82	-25779.55

Standard errors are in parentheses. All regressions are weighted to reflect the actual geographic distribution of establishments from County Business Patterns and include dummy variables for three-digit NAICS and month of survey. Key results in **bold.**
+significant at 90% confidence level.
*significant at 95% confidence level.
**significant at 99% confidence level.

Table 5: Exploring the Exogeneity Assumptions: Robustness to Establishment-Level Decisions and Instrumental Variables

	Subset of Firms		Instrumental Variables			
	Establishment-Level Choices Only		Instrument for Establishment Capability and Organizational Capability		Instrument for Establishment Capability, Organizational Capability, Establishment Capability*City, and Organizational Capability*City	
	(1)	(2)	(3)	(4)	(5)	(6)
	Capability Defined by Programmers	Capability Defined by Factors	Capability Defined by Programmers	Capability Defined by Factors	Capability Defined by Programmers	Capability Defined by Factors
$\frac{\partial Y_i}{\partial OC_i}$	0.0040	0.0111	0.0096	0.0261	0.0045	0.0309
	(0.0043)	(0.0062)*	(0.0045)*	(0.0068)**	(0.0025)+	(0.0127)*
$\frac{\partial Y_i}{\partial EC_i}$	0.0647	0.0509	0.0743	0.0787	0.0972	0.1485
	(0.0074)**	(0.0074)**	(0.0115)**	(0.0173)**	(0.0301)**	(0.0440)**
$\frac{\partial^2 Y_i}{\partial z_i \partial OC_i}$	-0.0183	-0.0118	-0.0118	-0.0381	-0.0046	-0.0006
	(0.0089)*	(0.0154)	(0.0072)	(0.0128)**	(0.0016)**	(0.0350)
$\frac{\partial^2 Y_i}{\partial z_i \partial EC_i}$	-0.0252	-0.0506	-0.0084	-0.1110	-0.0117	-0.0874
	(0.0197)	(0.0211)*	(0.0355)	(0.0508)*	(0.0158)	(0.0481)+
$\frac{\partial Y_i}{\partial z_i}$	0.0141	0.0092	0.0246	0.0283	0.0072	0.0089
	(0.0205)	(0.0228)	(0.0061)**	(0.0067)**	(0.0049)	(0.0074)
$\frac{\partial Y_i}{\partial me_i}$			0.0135	0.0246	0.0145	0.0053
			(0.0153)	(0.0064)**	(0.0083)+	(0.0097)
$\frac{\partial Y_i}{\partial \ln(empl)_i}$	0.0637		0.0708		0.0116	
	(0.0122)**		(0.0055)		(0.0156)	
Observations	6708	6708	86792	86792	86792	86792
LL	-3640.56	-3748.45	-25551.23	-26808.13	-26355.95	-26354.48

Values represent marginal effects at means. Standard errors are in parentheses. Columns (1) and (2) are weighted to reflect the actual geographic distribution of establishments from County Business Patterns. All regressions include dummy variables for three-digit NAICS and month of survey. Interaction coefficients are significant except $\frac{\partial^2 Y_i}{\partial z_i \partial OC_i}$ in columns (2) and (6) and $\frac{\partial^2 Y_i}{\partial z_i \partial EC_i}$ in columns (1), (3), and (5).

Table 6: Robustness to City Definitions and Different Technologies

	Different City Definitions				Different Technologies (City is defined as MSA population > 500,000)			
	Any MSA		County Density above the Sample Median (610 people per square mile)		Uses an Internet Language		Has a PC Server	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Capability Defined by Programmers	Capability Defined by Factors	Capability Defined by Programmers	Capability Defined by Factors	Capability Defined by Programmers	Capability Defined by Factors	Capability Defined by Programmers	Capability Defined by Factors
$\frac{\partial Y_i}{\partial OC_i}$	0.0055	0.0158	0.0063	0.0169	0.0017	-0.0003	0.0044	-0.0027
	(0.0016)**	(0.0034)**	(0.0016)**	(0.0037)**	(0.0012)	(0.0028)	(0.0015)**	(0.0024)
$\frac{\partial Y_i}{\partial EC_i}$	0.0786	0.0821	0.0782	0.0875	0.0465	0.0388	0.0727	0.0520
	(0.0025)**	(0.0058)**	(0.0028)**	(0.0051)**	(0.0019)**	(0.0028)**	(0.0033)**	(0.0049)**
$\frac{\partial^2 Y_i}{\partial z_i \partial OC_i}$	-0.0120	-0.0233	-0.0048	-0.0082	-0.0047	-0.0088	-0.0053	-0.0067
	(0.0027)**	(0.0083)**	(0.0018)**	(0.0053)	(0.0018)*	(0.0051)+	(0.0023)*	(0.0040)+
$\frac{\partial^2 Y_i}{\partial z_i \partial EC_i}$	-0.0179	-0.0775	-0.0042	-0.0300	-0.0033	-0.0214	-0.0241	-0.0485
	(0.0081)*	(0.0143)**	(0.0041)	(0.0090)**	(0.0038)	(0.0066)**	(0.0079)**	(0.0086)**
$\frac{\partial Y_i}{\partial z_i}$	0.0331	0.0455	0.0291	0.0429	0.0143	0.0295	-0.0138	-0.0089
	(0.0061)**	(0.0075)**	(0.0045)**	(0.0053)**	(0.0048)	(0.0065)**	(0.0055)**	(0.0049)+
$\frac{\partial Y_i}{\partial me_i}$	0.0308	0.0489	0.0308	0.0492	-0.0208	-0.0191	-0.0401	-0.0202
	(0.0077)**	(0.0063)**	(0.0075)**	(0.0061)**	(0.0056)**	(0.0056)**	(0.0068)**	(0.0046)**
$\frac{\partial Y_i}{\partial \ln(empl)_i}$	0.0649		0.0632		0.0138		0.0426	
	(0.0037)**		(0.0036)**		(0.0028)**		(0.0038)**	
Observations	86871	86871	86871	86871	86871	86871	86877	86877
LL	-24531.22	-25868.97	-24530.10	-25848.86	-18589.44	-19338.25	-54902.83	-55637.69

Values represent marginal effects at means. Standard errors are in parentheses. All regressions are weighted to reflect the actual geographic distribution of establishments from County Business Patterns and include dummy variables for three-digit NAICS and month of survey. Interaction coefficients are significant.

Table 7: Robustness to Competition

	Competition defined by number of establishments in same six-digit NAICS and the same county		Competition defined by total employment in same six-digit NAICS and the same county	
	(1)	(2)	(3)	(4)
	Capability Defined by Programmers	Capability Defined by Factors	Capability Defined by Programmers	Capability Defined by Factors
$\frac{\partial Y_i}{\partial OC_i}$	0.0090	0.0221	0.0083	0.0250
	(0.0022)**	(0.0054)**	(0.0028)**	(0.0077)**
$\frac{\partial Y_i}{\partial EC_i}$	0.0762	0.0674	0.0867	0.0683
	(0.0044)**	(0.0121)**	(0.0060)**	(0.0186)**
$\frac{\partial^2 Y_i}{\partial z_i \partial OC_i}$	-0.0046	-0.0089	-0.0062	-0.0102
	(0.0025)+	(0.0071)	(0.0024)**	(0.0067)
$\frac{\partial^2 Y_i}{\partial z_i \partial EC_i}$	-0.0176	-0.0632	-0.0124	-0.0582
	(0.0063)**	(0.0133)**	(0.0061)*	(0.0126)**
$\frac{\partial^2 Y_i}{\partial competition_i \partial OC_i}$	-0.0011	-0.0022	-0.0004	-0.0015
	(0.0005)*	(0.0017)	(0.0004)	(0.0011)
$\frac{\partial^2 Y_i}{\partial competition_i \partial EC_i}$	0.0018	0.0058	-0.0007	0.0027
	(0.0013)	(0.0033)+	(0.0009)	(0.0025)
$\frac{\partial Y_i}{\partial z_i}$	0.0263	0.0407	0.0281	0.0403
	(0.0056)**	(0.0066)**	(0.0054)**	(0.0064)**
$\frac{\partial Y_i}{\partial me_i}$	0.0301	0.0477	0.0304	0.0478
	(0.0076)**	(0.0061)**	(0.0077)**	(0.0062)**
$\frac{\partial Y_i}{\partial \ln(empl)_i}$	0.0638		0.0642	
	(0.0037)**		(0.0038)**	
$\frac{\partial Y_i}{\partial competition_i}$	0.0041	0.0030	0.0021	0.0024
	(0.0016)*	(0.0019)	(0.0010)*	(0.0012)*
Observations	86871	86871	86871	86871
LL	-24517.55	-25837.67	-24522.48	-25847.11

Values represent marginal effects at means. Standard errors are in parentheses. All regressions are weighted to reflect the actual geographic distribution of establishments from County Business Patterns and include dummy variables for three-digit NAICS and month of survey. Interaction coefficients are significant except $\frac{\partial^2 Y_i}{\partial z_i \partial OC_i}$ in column (2).

Table 8: Marginal Effects by Industry: Is the Result Industry-Specific or Economy-Wide?

	Manufacturing Only (two-digit NAICS 31, 32, and 33)		Services Only (two-digit NAICS 42, 44, 45, 48, 49, 51, 52, 53, 54, 55, 56, 61, 62, 71, 72, and 81)	
	(1)	(2)	(3)	(4)
	Capability Defined by Programmers	Capability Defined by Factors	Capability Defined by Programmers	Capability Defined by Factors
$\frac{\partial Y_i}{\partial OC_i}$	0.0159	0.0205	0.0026	0.0126
	(0.0023)**	(0.0043)**	(0.0020)	(0.0046)**
$\frac{\partial Y_i}{\partial EC_i}$	0.0856	0.0902	0.0764	0.0762
	(0.0038)**	(0.0056)**	(0.0030)**	(0.0064)**
$\frac{\partial^2 Y_i}{\partial z_i \partial OC_i}$	-0.0083	-0.0171	-0.0038	-0.0097
	(0.0032)*	(0.0082)*	(0.0034)	(0.0106)
$\frac{\partial^2 Y_i}{\partial z_i \partial EC_i}$	-0.0140	-0.0463	-0.0133	-0.0492
	(0.0077)+	(0.0119)**	(0.0081)+	(0.0140)**
$\frac{\partial Y_i}{\partial z_i}$	0.0260	0.0187	0.0397	0.0611
	(0.0078)**	(0.0076)*	(0.0076)**	(0.0094)**
$\frac{\partial Y_i}{\partial me_i}$	0.0404	0.0965	0.0250	0.0298
	(0.0106)**	(0.0079)**	(0.0100)*	(0.0082)**
$\frac{\partial Y_i}{\partial \ln(empl)_i}$	0.0831		0.0626	
	(0.0052)**		(0.0047)**	
Observations	24240	24240	58767	58767
LL	-8945.55	-9342.02	-15377.88	-16265.29

Values represent marginal effects at means. Standard errors are in parentheses. All regressions are weighted to reflect the actual geographic distribution of establishments from County Business Patterns and include dummy variables for three-digit NAICS and month of survey. All interaction coefficients are significant. Significance levels do not change if standard errors are clustered by firm.

Figure 1a: Probability of Adoption by Organizational Capabilities

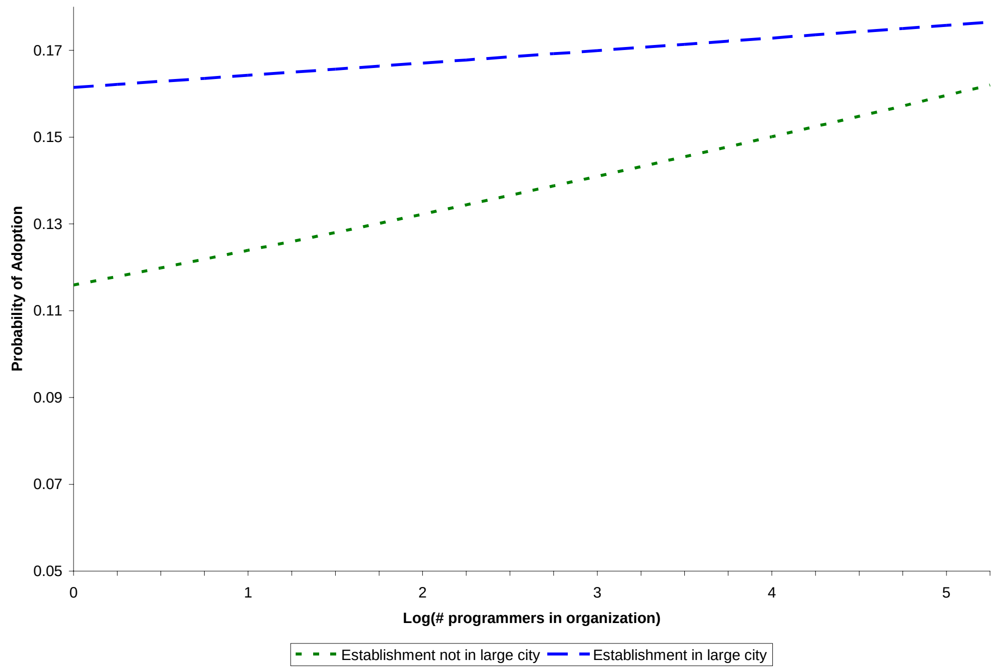
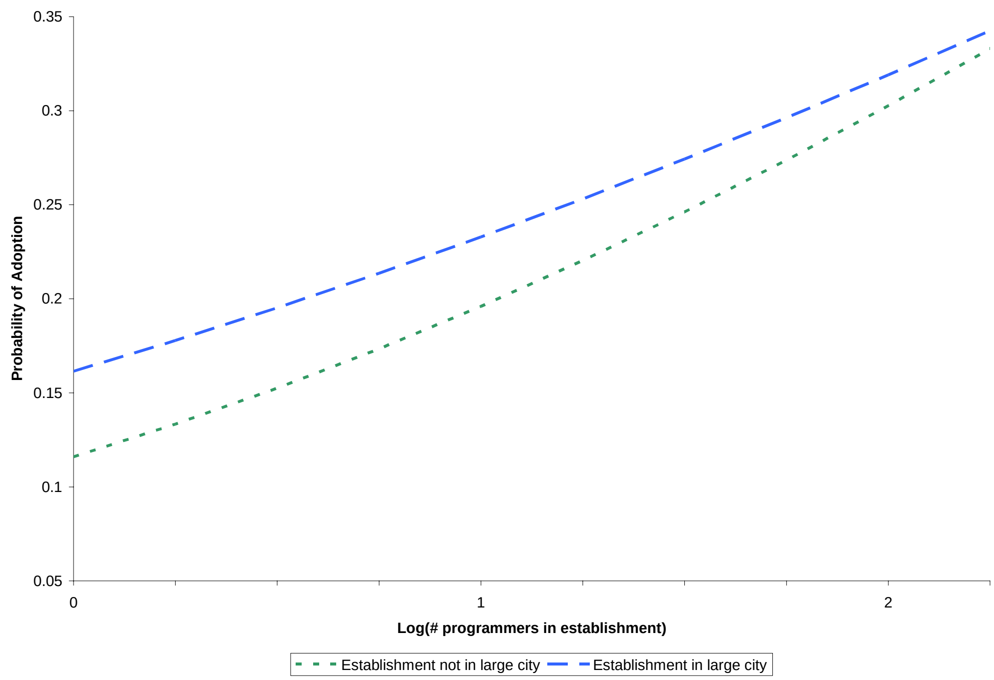


Figure 1b: Probability of Adoption by Establishment Capabilities



Predictions are based on a representative firm in the second half of 2000 with mean values of employment, industry effects, and multi-establishment status. Figure 1a assumes establishment capabilities are zero. Figure 1b assumes organizational capabilities are zero.