IT Innovation Fit, Misfit, and Organizational Performance:
A Theoretical Model and Empirical Test in U.S. Manufacturing Plants*

Robert G. Fichman

Nigel P. Melville¹

September 10, 2008

*Presented at the 2008 WISE Conference in Paris, France

Abstract
This paper extends the dominant paradigm of IS innovation research by analyzing the extent to which organizations adopt electronic networking technologies (ENT) sub-optimally and measuring the implications on organizational performance. We advance the concept of innovation misfit by drawing on extant economic and organizational theories to develop a set of propositions linking the degree of misfit to productivity. Principal propositions are tested using a proprietary dataset of over 25,000 U.S. manufacturing plants. Empirical analysis supports the basic idea of innovation misfit, while raising new and interesting questions for future research.

¹We thank Ron Ramirez, Vern Richardson, Vijay Gurbaxani, and Kenneth L. Kraemer for valuable comments and suggestions. The paper has also benefited from suggestions from seminar participants at the University of Hawaii and the University of California, Irvine. We also thank Maggie Levenstein, Stan Sedo, and Clint Carter at the Michigan Census Research Data Center for helpful guidance and critical administrative and data support. Financial support for this research at the Michigan Census Research Data Center is gratefully acknowledged. All remaining errors are ours. The research in this paper was conducted while the authors were Special Sworn Status researchers of the U.S. Census Bureau at the Michigan Census Research Data Center. Research results and conclusions expressed are those of the authors and do not necessarily reflect the views of the Census Bureau. [This paper has been screened to insure that no confidential data are revealed.]

¹ Presenting author.
1. Introduction

A large literature has examined antecedents of organizational adoption of information technology (IT) (Swanson 1994; Fichman 2000). The typical goal of these studies has been to identify a set of characteristics that distinguish leading adopters from laggards in some context (Fichman 2004), resulting in a profile of an innovator (and its mirror image, the profile of a laggard). The innovator profile studies assume that managers are rational in their decision making, (Swanson and Ramiller 2004) sizing up whether adoption makes sense given their particular situation. However, due to bounded rationality firms may not adopt at the optimal level. When this occurs, there is an *innovation misfit*: a gap between a firm’s normatively expected level of innovation and its actual innovation behavior.

Given that the IS adoption literature has used an instrumental logic that more is better, and the IT business value literature tends to focus on complementarities (Bresnahan et al. 2002), there is a lack of knowledge about what happens under conditions of innovation misfit. Intuitively, we would expect that firms that have the normative profile of laggards but adopt like leaders would be less prone to innovate successfully and should have fewer opportunities to exploit the fruits of innovation should they manage to assimilate the technology against the odds. Conversely, firms that have the profile of a leader but act like laggards should experience an opportunity cost for missing out on a (presumably) beneficial technology that similarly positioned competitors are exploiting with full abandon. This premise has not been analyzed to our knowledge in prior research and forms the basis of our study.

The structure of the paper is as follows. In Section 2 we develop the concept of innovation misfit, which we divide into two types—*over-adoption misfit*, which occurs when a firm’s actual adoption exceeds the level suggested by its normative innovation profile, and *under-adoption misfit*, which occurs when actual adoption lags the level indicated by its normative innovation profile. We theorize about the impact of each type of misfit on organizational performance and consider explanations for why pervasive misfits might be observed in practice. In Section 3 we develop a conceptual model for IT innovation misfit that synthesizes our arguments and guides empirical analysis. In Section 4 we conduct an empirical
examination of the effects of innovation misfit on labor productivity using a proprietary data set capturing the adoption of electronic networking technologies (ENT) in over 25,000 US manufacturing plants in 1999 gathered by the US Bureau of the Census. Results support our premise that over-adoption misfit negatively moderates the relationship between ENT adoption and labor productivity. Under-adoption misfit does as well, though to a lesser extent. Specifically, we find that over-adoption reduces the impact of ENT on labor productivity by about 28%, while under adoption reduces it by about 12%. Interestingly, we find that in addition to the negative moderating effect, under-adoption misfit has a small but positive direct effect on labor productivity. In Section 5 we propose several possible explanations for these findings, and wrap-up with a discussion of the implications of our work for theory and practice.

2. Prior Research

In this section, we outline the dominant instrumental logic of IS innovation research, provide an intuitive motivation for the existence of innovation misfits, and apply relevant theory. This lays the groundwork for our conceptual model, which we present in the next section.

2.1 Dominant Instrumental Logic of IS Innovation

Why do some firms take the lead in adopting IT innovations, while others lag? This is a central question in the IT innovation diffusion field, one that IT researchers have addressed by formulating models wherein a set of explanatory factors (e.g., organizational structure, resources, competitive environment) are used to predict the timing and extent of innovation adoption (Fichman 2004). These factors can be used to identify the profile of innovation leaders (and laggards) in some context.

Prior research has been guided by a great diversity of reference theories, including those based on economic rationality, competitive effects, organizational learning, institutional theory, the resource-based view of the firm, and complementarities. Nevertheless, this stream has been dominated by an economic-rationalistic-logic wherein the firms expected to take the lead in practice are those that normatively should lead (Fichman 2004). Sometimes the rationales for predictor variables are explicitly economic. But more often, rationales revolve around one or both of two implicitly economic themes, one concerning a firm’s
ability to adopt, and the other concerning a firm’s degree of innovation-related needs and/or opportunities (Table 1). A criticism of the dominant paradigm is that the normative assumption does not always hold, i.e., some firms that normatively speaking should take the lead do not, and vice versa. This is the central motivating issue of this paper.

Table 1: Prior Research - Antecedents of Organizational IT Innovation

<table>
<thead>
<tr>
<th>Construct</th>
<th>Link to Ability to Adopt &amp; Innovation Needs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organizational Size</td>
<td>Proxy for other variables that are positively related to innovation, including scale, slack resources, professionalism and specialization (Tornatzky and Fleischer 1990, pg. 162). A larger scale of activities promotes the ability to amortize innovation costs (Fichman and Kemerer 1997).</td>
</tr>
<tr>
<td>Top Management Support</td>
<td>Top managers are responsible for allocating appropriate resources to an innovation effort, aligning incentives, monitoring progress towards goals, and intervening to get innovation projects back on track (Sharma and Yetton 2003). Supportive managers are more likely to engage in these innovation management tasks. Also, top managers are more likely to support innovation projects they see as well-aligned with a firm’s strategic needs and abilities (Huigang et al. 2007).</td>
</tr>
<tr>
<td>Compatibility</td>
<td>Technologies that are more compatible require fewer organizational changes to implement and are less likely to provoke resistance, resulting in lower costs and less risk of underutilization or failure. More compatible technologies should also have greater complementarities, leading to magnified innovation returns (Zhu 2004).</td>
</tr>
<tr>
<td>Organizational Knowledge and Skill</td>
<td>Organizations with greater knowledge and skill have lesser costs associated with organizational learning and are more likely to make good decisions through out the innovation process (Fichman and Kemerer 1997).</td>
</tr>
</tbody>
</table>

2.2. Beyond the Dominant Logic: Intuitive Considerations of Under- and Over-Adoption Misfit

Models of innovation adoption often assume a normative stance by employing rationales that link innovation to abilities or needs. When a prediction model incorporates factors that capture (or at least correlate with) innovative abilities, then firms positioned favorably on this model (i.e., firms that fit the profile of an innovator) are better-positioned to profit from being a leading adopter, because they are more likely to have an abundance of resources and expertise to apply to the effort. This is important, because innovation is not easy. While the decision to adopt requires no special talent, the path from adoption to assimilation and then on to captured business value presents a series of challenges and potential obstacles (Avison et al. 2006). Firms may choose the wrong instance of the technology, or improperly configure it, or install it badly, or find the intended users resist it, or fail to make crucial complementary changes to the surrounding organization (Brynjolfsson and Hitt 2000; Melville et al. 2004). Firms with higher innovative abilities should be more likely to skillfully navigate these obstacles and achieve deeper assimilation at lower cost, resulting in higher innovation returns. However, firms that innovate despite having an innovation profile suggestive of a laggard—i.e., those with an over-adopter
type misfit—are less likely to be able to translate the adoption decision into thorough assimilation and positive organizational impacts. Such firms tend to have fewer resources, less expertise, and cultures and structures that are less conducive to innovation. The net result for them should be higher costs, lower benefits, and so diminished returns to innovation.

Similar patterns should be observed to the extent the innovation prediction model includes variables that capture innovation needs/opportunities. In this case, firms that rate favorably on this model are positioned to derive greater benefits from being an adoption leader, virtually by definition. Such firms are more likely to have a high degree of compatibility between the innovation and their strategies, processes, values and skills, and to operate in industries where the innovation is particularly well-suited. Also, they should be more persistent in their assimilation efforts because they quite rightly foresee a greater potential reward at the end of the road. On the other hand, firms that are poorly positioned on variables that capture innovation needs and opportunities may be more prone to having their commitment drain away as obstacles are encountered, and as realization dawns that perhaps the firm is not the best candidate for innovation after all. Should assimilation be achieved despite obstacles, the innovation returns will tend to be diminished owing to the relative lack of needs for the innovation or opportunities to exploit it.

The above arguments suggest that when a theoretical model of innovation is based on an instrumental logic—and most are—firms exhibiting a high fit between expected and actual innovation should have higher returns to innovation than those with an over-adoption misfit.

The implications of innovation misfits in the form of under-adoption are more complex. At the most basic level it could be argued such firms pay an opportunity cost in that they have a lesser level of adoption—and so a lesser level of adoption-driven benefits—than peer firms that “right-adopted.” However, a more interesting issue is whether and how being an under-adopter affects the returns to any given level of adoption. On the one hand, such firms might be viewed as being well qualified on a structural level—if not over qualified—for the level of adoption they have chosen. Unlike over-adopters, which as argued above would tend to lack the resources and expertise needed to translate adoption into business value,
under-adopters would occupy a position on the innovation profile relative to their own behavior suggesting a greater abundance of resources and expertise than even right-adopting peers. On the other hand, this sort of misfit is evidence of a lack of *mindfulness* among decision makers about innovation timing. (We discuss the concept of innovation mindfulness in more detail in the section that follows.) If managers are not mindful in deciding when to adopt, perhaps they will also be unmindful in other decisions related to IT implementation and use. For example, although they may possess latent resources and expertise that could smooth the implementation process, they may fail to deploy those resources appropriately.

As a result, while we envision a straightforward, negative relationship between over-adoption type misfit and returns to innovation, we see a more complex relationship between under-adoption misfit and innovation returns. We will address the issue of how over- and under-adoption affect innovation returns in more detail when we present our conceptual model and propositions in the next section. First, we provide additional conceptual foundations related to the origin of innovation misfits.

**2.3 Innovation Misfit and Organizational Performance: Theoretical Considerations**

We briefly review conceptual perspectives that are consistent with deviations from economic rationality and that steer managers away from making profit-maximizing decisions. Our point of departure for this discussion is Swanson and Ramiller’s (2004) concept of *innovation mindfulness*. They classify an organization as innovating mindfully when it “attends to innovation with reasoning grounded in its own facts and specifics,” and further explain that these situational specifics “… can be quite complex, including, among other issues, the innovation's ramifications for operational efficiencies and strategic advantage; the organization's preparedness for the change involved; the quality and availability of complementary resources needed; implications for various common and conflicting interests, both internally and in interfirm relationships; and the effects of adoption on the firm's legitimacy with outside constituencies” (pg. 4). It is noteworthy that most of these situational specifics have an obvious normative interpretation related to innovation abilities or needs/opportunities, and that they pertain to managerial
decision making throughout the adoption and implementation process. A mindless organization, by contrast, makes adoption decisions without a reasoning grounding in these situational specifics.

These definitions suggest a strong linkage between innovation mindfulness and innovation fit. In particular, we should expect that mindful organizations will more likely to consciously seek a state where there is a good fit between their specific characteristics and actual innovation behavior, and will hence be less prone to experience inadvertent innovation misfits. By contrast, the mindless organization should be highly prone to such misfits.

Despite these clear overlaps, it would be a mistake to treat mindfulness/mindlessness and innovation fit/misfit as equivalent concepts, because mindfulness is a cognitive element predominantly associated with a process, while innovation fit is a structural element associated with a firm’s innovation profile. Also, managers can mindfully attend to certain specifics (e.g., implications of adoption for their own prestige) that do not necessarily relate to instrumental impacts of adoption. Nevertheless, where mindlessness is pervasive, we expect that innovation misfits will follow. This means that the conditions that lead to more pervasive mindless adoption should also lead to more pervasive innovation misfits.

Mindless innovation can arise from a number of causal mechanisms, including: institutional isomorphism (the tendency of organizations to come to resemble each other more than they should from an instrumental standpoint (DiMaggio and Powell 1983)), information cascades (the tendency of decision makers to base choices on observations of others’ choices more so than private information (Bikhchandani et al. 1998)) and managerial fashions (the tendency of organizations to adopt innovations that are the most prominent topics of discourse (Abrahamson 1996). Though the mechanisms differ, these perspectives all provide explanations for why some firms might tend to go along with the pack without giving much attention to situational specifics as a factor in adoption decisions.
2.3.1 Managerial Fashions

Swanson and Ramiller give a privileged position to managerial fashions as an explanation for widespread mindlessness, and in fact include among their propositions that “mindlessness in innovating with IT will be observed more widely the more fashionable the organizing vision” (2004, pg. 571). We agree this is a particularly salient perspective in the context of IT innovations, given the endemic level of hype seen in the IT domain, and thus we will focus our attention here.

Abrahamson defines management fashions as “relatively transitory collective beliefs, disseminated by the discourse of knowledge entrepreneurs, that a management technique is at the forefront of rational management progress” (Abrahamson and Fairchild 1999, p. 709) Management fashion researchers have argued that the forces of fad and fashion can have a powerful effect on innovation diffusion processes. They posit that knowledge entrepreneurs (such as academics, consultants, business gurus, and technology vendors) have an interest in focusing managerial attention on emerging innovations. They exploit strong norms of progress among managers to create interest in new ideas, which are portrayed as being on the leading edge of managerial practice.

Abrahamson (1991) argues that the forces of fad and fashion can provoke two kinds of divergences from the normative ideal: (1) community adoption of technically inefficient innovations (i.e., technologies that are not actually beneficial to most firms in a population), and (2) community rejection of technically efficient innovations (technologies that would be beneficial to most firms in a population). The first kind of divergence implies a population-level bias towards over-adoption, while the latter kind of divergence implies a bias towards under-adoption.

It is worth pausing here to reflect more deeply on what a bias toward over-adoption or under-adoption really means at the level of the population and the level of the firm. Abrahamson’s dichotomy of technically efficient and technically inefficient innovation is a useful rhetorical device, but over simplifies the situation in the case of IT. The potential benefits of emerging IT are not static, and neither are they uniformly distributed across a population of potential adopters. Rather, technologies improve over time as
products embodying the technology broaden and mature, as complementary products and services emerge, and as the community learns how to best deploy and use the technology. Some improvements come with the passage of time and greater vendor investment, while others arise from network effects surrounding actual use of the innovation in a growing adopter population. In the early portion of the diffusion cycle, the adoption tends to be economical for just a few firms, and just for a narrow range of uses within those firms. As time goes on and the adopter community grows, the innovation becomes economical for increasing numbers of firms and uses within firms. An innovation whose inherent superiority is high has the potential to eventually become economically attractive to most firms, while an innovation whose inherent superiority is low will only ever become attractive to a few. In the economic ideal, all firms would queue up in the order of their innovation-related abilities and needs, and then jump across the adoption threshold only when the technology has improved enough to warrant adoption—if it ever does.

Over-adoption at the level of the population is then a relatively straightforward concept: it occurs when there is a bias among firms towards jumping across the innovation threshold too soon and/or with too much enthusiasm, meaning that the average returns for the population as a whole are lower than they would have been. It does not necessarily mean that a technology is poor or that no firms should ever have adopted; rather it means there was a collective bias among firms towards adopting earlier and or more broadly than would have optimized their collective returns. Conversely, population-level under-adoption occurs when there is a bias towards late adoption and a too-narrow commitment to use. For example, it could be that the “bust” following dot-com boom went too far, with—at least for a time—firms being more reluctant that they should have been to invest in IT innovations.

Over-adoption at the level of the firm is more complex, and can be viewed in either of two ways: our relative notion of over-adoption, where a firm takes an earlier place in the adoption queue than its innovation profile would suggest it should, and absolute over-adoption, where a firm crosses the innovation threshold before it should to optimize returns (regardless of what place in the queue it holds).
In the absence of a large population-level bias, firms that exhibit relative over-adoption will also tend to have over-adopted on an absolute level.²

3. Theoretical Model of the Impact of Innovation Misfit on the Returns to IT Adoption

We now formalize arguments pertaining to innovation fit and misfit in a theoretical model. We develop two propositions under the maintained assumption that there is no strong population-level bias. This occurs when the systematic forces that propel irrational exuberance (e.g., hype) are roughly balanced by forces causing excessive caution (e.g., organizational inertia). We then relax this assumption and theorize about how population effects might affect principal propositions.

3.1 Innovation, Innovation Misfit, and Organizational Efficiency

Our model posits a main path of causation in which favorable positions on normative innovation profile characteristics lead firms to engage in greater extent of innovation (earlier and more extensive adoption) with respect to a focal technology, which, in turn, produces increased efficiency (Figure 1, Path #1 and Path #2; Table 2). We explicate each in turn. The positive association between the normative innovation profile and extent of innovation (Path #1) will hold to the extent that managers are aware (explicitly or implicitly) of what the innovation profile characteristics are in some context and prefer normatively rational outcomes. The factors that actually constitute this profile must be tailored to the theoretical context at hand, although there are certain characteristics (e.g., organizational size and scale, skills and expertise, top management support, compatibility) that are favorable to innovation leadership in a wide variety of contexts. We discuss these in the next section on construct operationalization.

² We note that while the absolute notion of over-adoption might hold some surface appeal, it is not feasible to measure in practice, and is not very interesting theoretically, as its relationship to performance is tautological.
The rationale for a positive link between greater IT innovation and organizational efficiency (Figure 1, Path #2) is slightly more complicated. At the most basic level, the adoption of innovative IT can be viewed as creating a particular kind of IT capital. According to micro-economic theory, a rational firm should invest in any particular form of capital just to the point where the last unit of investment produces no more value than it costs; since costs of IT capital are positive, the gross marginal productivity contribution of IT should likewise be positive (Hitt and Brynjolfsson 1996). A large number of studies have verified a positive link between IT capital and productivity, both when IT is defined as aggregated IT investment (Brynjolfsson and Hitt 1996) and when it is defined as the use of some particular technology (Mukhopadhyay et al. 1997; Hitt et al. 2002; Barua et al. 2004). Although additional arguments can be developed for why IT innovation adoption should promote productivity (for example, IT innovations might enable a more efficient organization of production) the interpretation of IT innovation as producing a specialized form of capital suffices for our current purpose.
**P1**: The greater the adoption of IT innovation, the greater is firm productivity, *ceteris paribus*.

Now consider innovation misfit—the degree of misalignment between the extent to which a firm is innovating versus the extent to which it is theoretically expected to be innovating—and its implications on firm performance (Figure 1, Path #4). In the general case, a mindful firm that seeks normatively rational outcomes will attempt to choose a relative position in the population-level adoption queue that reflects its relative position on the innovation profile. A firm whose innovation profile is indicative of the greatest abilities and needs with respect to the innovation will usually be best served to take the lead, while a firm that has the profile of a laggard will usually be best served to defer adoption.

To make things more concrete, imagine an emerging IT for which the normative profile of an innovator is dominated by just one factor: firm size. Perhaps this hypothetical technology is most suitable for large scale production and requires the kind of professionalism and expertise that is most often found in large firms. Now imagine two firms have adopted relatively early, one large and one small. The large firm would be a right-adopter with a high degree of innovation fit (Figure 2a); the small firm would be an over-adopter (Figure 2b). Of the two firms, we would expect the right-adopting large firm to have greater returns to innovation than the over-adopting small firm. The large firm is well suited to innovating in the early part of the overall diffusion cycle, but the small firm is poorly suited to it. In particular, the small firm is at a structural disadvantage (due to a lower production scale) and also has a greater risk of an excessively costly and poor quality implementation (due to the lack of expertise). The small firm might still get some productivity enhancement from adoption, but less than the large firm with the more appropriate production scale—and less than the small firm would get were it to wait until a later time, when, as would normally occur, the technology will be simpler, cheaper, and broader in its potential applicability.
In addition to the structural disadvantage just described, over-adoption implies a lack of mindfulness among managers in timing adoption. In this hypothetical situation, the key organizational fact for managers to attend to in making the adoption decision is firm size, and yet the behavior of the small firm suggests they did not attend to it. Swanson and Ramiller define innovation mindfulness as a general property of an organization in its engagement with a technology, and if so, we would expect mindlessness in adoption timing to be associated with mindlessness throughout the implementation process. Thus, a firm-level over-adoption misfit is not only evidence of a structural disadvantage, but also an indicator of a potential implementation execution disadvantage arising from mindlessness, which should also tend to lower innovation returns. These arguments lead to our first proposition indicated in Figure 1, Path #4:

**P2:** Firms with an over-adoption misfit will experience a performance penalty relative to right-adopting firms in the form of lower returns to any given level of adoption.

While the theoretical implications of over-adoption misfit are relatively straightforward as explained above, the effects of under-adoption misfit are not. Returning to our hypothetical emerging technology, imagine a pair of late-adopting firms, one large and one small. In this case, the small firm will be classified as a right-adopter with a high innovation fit (innovation profile of a laggard profile paired with late adoption), while the large firm will have an under-adoption type misfit (innovation profile of a leader paired with late adoption).
In this situation the impact of under-adoption will be ambiguous. On the one hand, the large firm should be in a better structural position to adopt, as evidenced by a profile (large size) consistent with innovation leadership. Even though we are now looking at adoptions occurring at a later point in the diffusion cycle—where the innovation has come within the economic reach of smaller firms—there should still be some structural size advantage. On the other hand, there is evidence of a lack of mindfulness in the large firm in timing the adoption decision, which implies lesser mindfulness towards the technology in general. This lack of mindfulness may increase the chance of implementation execution problems. For example, resources may exist in abundance but not be properly allocated to the implementation project. Or, managers may fail to recognize the need to make complementary changes to other organizational practices (e.g., decentralization of decision making).

To summarize, we propose that under-adoption type misfit will lead to two opposing effects. Such firms will tend to be structurally better positioned to get improved innovation returns, but are less likely to be mindful in terms of project execution, which tends to lower returns. This leads to our second proposition:

**P3:** Firms with an under-adoption misfit will experience a performance benefit or a performance penalty relative to right-adopting firms.

- **P3a:** The performance implications of under-adoption misfit will be positive if the magnitude of the positive structural effect of having the profile of an innovation leader is larger than the negative implementation execution effect of lower innovation mindfulness.
- **P3b:** The performance implications of under-adoption misfit will be negative if the magnitude of the positive structural effects of having the profile of an innovation leader is smaller than the negative implementation execution effect of lower innovation mindfulness.

Were it found that under-adoption misfit increases returns, this could be taken as evidence that the positive structural effect of having the profile of a leader (large size) dominates the negative effect of mindlessness in project execution. Conversely, were it found that under-adoption decreases returns, this could be taken as evidence the negative mindlessness effect outweighs the positive structural effect.

### 3.2 Effect of Population Bias

So far, we have been operating under the assumption that there is no bias towards innovating too early or late in the population of firms as a whole. Now we consider how predictions might change were there a strong population-level bias in one way or the other. When there is significant population bias towards
adopting too soon on an *absolute* level—meaning that firms are prone to adopt earlier than would be normatively ideal—then *relative* over-adoption at the firm level will still be an indicator of both a structural disadvantage and an implementation execution disadvantage. If anything, we would expect to see an even stronger penalty for firm-level over-adoption in this case. The firm would have jumped ahead of its expected place in an adoption queue that is itself ahead of where it should be.

By contrast, some degree of under-adoption might be viewed as evidence of a mindful process of swimming against a tide that contains many mindless adopters (Fiol and O’Connor 2003). Thus we would expect the productivity penalty for firm-level under-adoption (if any) to be lessened when there is a population bias towards over-adoption. For example, during the height of the dot-com boom many firms engaged in misguided ecommerce initiatives, which suggests the possibility of better organizational performance for those that were more circumspect.

**P4a:** In the presence of an early population-level bias in adoption, the impact of firm-level over-adoption misfit is exacerbated.

**P4b:** In the presence of an early population-level bias in adoption, the impact of firm-level under-adoption misfit (if any) is mitigated.

When there is a population-level bias toward adopting too late, this should have the opposite of the effects described above. Some amount of firm-level over-adoption might be seen as a positive sign of mindfulness, and if so, the negative effects of an over-adoption misfit could be moderated, or possibly eliminated entirely. By contrast, the firms with an under-adoption misfit might be seen as especially mindless, increasing the chance of an overall performance penalty.

**P5a:** In the presence of a late population-level bias in adoption, the impact of firm level over-adoption misfit is mitigated.

**P5b:** In the presence of a late population-level bias in adoption, the impact of under-adoption misfit (if any) is exacerbated.

The net result is that we expect the moderating relationship posited in Proposition 2 to hold at least as strongly when there is a population-level over-adoption as when there is no bias, but to get increasingly weaker as a population bias moves towards adopting too late. Regarding Proposition 3, we expect a population-level biases will shift the balance between the positive and negative facets of firm-level under-
adoption. In particular, the negative facet will increase in the presence of a population under-adoption bias, and decrease in the presence of a population over-adoption bias.

4. Empirical Analysis

Our empirical analysis proceeds in three stages. First we use data on the adoption of electronic networking technologies (ENTs) by US manufacturing plants to estimate a normative adoption prediction model that includes as antecedents several variables relating to a plant’s abilities and needs with respect to ENTs. Next, we compute measures for over- and under-adoption misfit based on the residuals from this model. Third, we estimate a set productivity regressions that allow us to examine the effects of misfit on performance. In the sections below we describe our data sources and measures, and then present our statistical models.

4.1 Data Sources

The dataset used for this study was constructed by combining data from three separate proprietary data sources collected and maintained by the Center for Economic Studies, U.S. Census Bureau. Together, these datasets represent the most comprehensive and accurate measurement of the U.S. manufacturing sector and are widely used to study technology and productivity (Power 1998; Bartelsman and Doms 2000; Dunne et al. 2000; Atrostic and Nguyen 2005). The first data source is the Census of Manufactures (CM), which includes numerous economic variables on the universe of U.S. manufacturing establishments – more than 300,000 in 1997 – for years ending in numerals 2 and 7. Variables include revenue, payroll, employees, cost of materials, energy consumed, value of shipments, capital expenditure, and capital stock.³ The second data source is the Annual Survey of Manufacturers (ASM), which spans years between Census years and comprises a subset of CM variables for a stratified sample of manufacturing establishments (roughly 50,000 in 1997). The third source of data for this study is the 1999 ASM Computer Network Use Supplement (CNUS). CNUS contains data on a range of electronic networking technologies (ENT) in use at the plant, including various technologies intended to improve

³ See Miranda and Jarmin {CES-WP-02-17, The Longitudinal Business Database, Javier Miranda, Ron Jarmin} for details.
firm efficiency, such as the use of the Internet, electronic data interchange (EDI), and intranets. Matching these three data sources provides a detailed sample of economic activity and ENT use across more than 25,000 U.S. manufacturing plants.

4.2 Measures

**IT Innovation Extent.** Our modeling strategy demands a measure of IT innovation extent that has two characteristics. First, it should pertain to technologies with sufficient reach and range to have a measurable impact on overall plant performance. For example, spreadsheet software used by a small group of engineers would not suffice. Use of electronic data interchange (EDI), by contrast, is likely to impact plant productivity (Hitt et al. 2002). Second, it should capture actual use, rather than investment: “merely examining the dollars invested in IT may not be an accurate reflection of the effectiveness of IT because the extent of its usage may vary across industries, firms, or process.” (Devaraj and Kohli 2003, p. 274). The CNUS dataset contains variables consistent with both characteristics: wide reach and range and actual use.

We utilize six items on the survey that capture the actual use of wide-ranging infrastructural electronic networking technologies: (1) use of local area networks (LAN), (2) EDI, (3) the Internet, (4) extranets, (5) intranets, and (6) other electronic networking technologies (Table 3). Each of the six ENT items is a dichotomous metric (use / don’t use), which we treat as an interval-ratio variable with two scale points. Thus, our measure of IT innovation extent is the sum of the number of ENTs reported to be in use at the plant. It is worth noting that the number of innovations adopted as of a particular date tends to be positively correlated with both average earliness of adoption for those innovations, and the average extent of implementation (Zmud and Apple 1992; Fichman 2001). The logic is that if a firm tends to adopt earlier, then there is more chance for those adoptions to have occurred prior to the survey date, which will lead to a larger number of innovations reported as in use. Also, a firm that has adopted an ENT earlier will have had time to implement those ENTs more extensively by the time of the survey date.
**Organizational Efficiency.** When theorizing about the business value impacts of information technology, it matters greatly which measure of performance is employed (Hitt and Brynjolfsson 1996). Measures focused on productivity suggest use of the theory of production and notions of efficiency as the underlying perspective. Profitability measures, by contrast, suggest theories related to strategy and competitive advantage. We adopt the productivity perspective for two reasons. First, our theoretical interest lies in how innovation misfits affect the returns to innovation, and the productivity perspective supports a more parsimonious theoretical model, largely free of potential complications relating to the actions of competing firms and consumers. Second, our empirical data pertains to IT use in a manufacturing setting. Here, productivity is relatively easy to measure (e.g., the ratio of revenue to labor) and the direct link between IT use and productivity tends to be strong (Lichtenberg 1995; Brynjolfsson and Hitt 1996; Hitt and Brynjolfsson 1996).

**Table 3: Construct Operationalization**

<table>
<thead>
<tr>
<th>Construct Name &amp; Definition</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT Innovation Extent:</td>
<td>ENT Adoption: Sum of six dichotomous ENT items: “Do you use x technology?” Where x is (LAN, EDI, internet, extranet, intranet, other). (Hitt et al. 2002; Banker et al. 2006)</td>
</tr>
<tr>
<td>Organizational Efficiency:</td>
<td>Labor Productivity: Ratio of total value of shipments to total employment. Prior period Labor Productivity is used in the adoption model. (Atrostic and Nguyen 2005)</td>
</tr>
<tr>
<td>Organizational Size:</td>
<td>Plant Size: For the Adoption Model: Sum of standardized values of total sales and total employment. For Productivity Model: Six plant size groupings represented by dummies. (Astebro 2002; Colombo and Delmastro 2002; Atrostic and Nguyen 2005)</td>
</tr>
<tr>
<td>Multi-Plant:</td>
<td>Multi-Plant: Dummy set to 1 for plants that are part of multi-plant firms (Colombo and Delmastro 2002; Atrostic and Nguyen 2005)</td>
</tr>
<tr>
<td>Plant Age:</td>
<td>Plant Age: Time since plant opened, in years. (Dunne 1994; Colombo and Delmastro 2002; Atrostic and Nguyen 2005)</td>
</tr>
<tr>
<td>Employee Skill:</td>
<td>Non-Production Labor Ratio (NPLR): Ratio of non-production labor to total labor. (Colombo and Mosconi 1995; Doms et al. 1997; Atrostic and Nguyen 2005)</td>
</tr>
<tr>
<td>Capital Intensity:</td>
<td>Capital Intensity: Ratio of capital to total labor. (Atrostic and Nguyen 2005)</td>
</tr>
<tr>
<td>Materials Intensity:</td>
<td>Materials Intensity: Ratio of materials expense to total labor. (Atrostic and Nguyen 2005)</td>
</tr>
<tr>
<td>Over-Adoption Misfit/ Under-Adoption Misfit:</td>
<td>Misfit Hi/Misfit Lo: Difference between fitted values in the Stage 1 adoption regression and actual values. Misfit Lo is 1 if in lower quartile, zero otherwise. Misfit Hi is 1 if in lower quartile, zero otherwise. (See Altman 2006) for use of residuals in two-step regression.</td>
</tr>
<tr>
<td>Industry:</td>
<td>Industry dummies: 3-digit NAICS (Atrostic and Nguyen 2005)</td>
</tr>
</tbody>
</table>

**Adoption Antecedents.** Our goal here is not to propose or test any particular theory of IT adoption. Rather, the goal is to use the available data to build a normative innovation profile using factors that can
be interpreted as capturing a firm’s innovation-related abilities or needs. We selected variables for the adoption prediction model based on prior studies of manufacturing technology that took an economic perspective (e.g., (Doms et al. 1997; Astebro 2002; Colombo and Delmastro 2002) and on previous studies making use of the CNUS dataset (Atrostic and Nguyen 2005). See Table 3 for rationales linking these variables to adoption.

**Labor Productivity Antecedents.** We selected variables for the labor productivity prediction model based on prior research (Atrostic and Nguyen 2005). See Table 3 for rationales linking these variables to labor productivity.

**Adoption Misfit.** Adoption misfit is defined as the degree of misalignment between a firm’s actual innovation and the level indicated by a normatively-derived innovation profile. While measures of innovation extent are directly available, the latter is not. One possible measurement strategy would be to conduct a primary survey asking managers the extent to which they perceive that their firm tends to over- or under adopt IT innovations. However, we expect such a measure would be extremely noisy, and would suffer from severe respondent biases. Another possibility would be to assume that a greater degree of agency conflicts, susceptibility to fads and fashion, mindlessness, and other non-normative firm-level characteristics are associated with a greater degree of misfit, and to use such variables to form a proxy measure of misfit. Data limitations preclude this approach.

A third strategy—and the one chosen here—is to build on an empirical economics literature that employs regression residuals (Kakwani 1993; Gouyette and Pestieau 1999; Jayasuriya and Wodon 2003). For example, Sen (1981) uses deviations from regressions of life expectancy on per capita income to suggest that certain countries are more effective than others in leveraging capital for human development. In our context, we assume that most firms are rational adopters and have a low innovation misfit, and further, that there is no population bias. Under these conditions, a regression of ENT adoption on innovation antecedents yields a fitted line with residuals. The residuals indicate the degree of misfit between actual adoption and theoretically expected adoption. If this measure of misfit reflects more than just
measurement noise—and if our theory about innovation misfit is correct—we would expect to see a connection between this residual-based measure of misfit and organizational performance. Specifically, we operationalize misfit as the difference between a firm’s actual extent of adoption of IT innovations and its expected extent of adoption based on our theoretical model. Over (under)-adoption misfit is a dichotomous variable that is 1 if the plant is in the top (bottom) 25 percentile of residuals, and zero otherwise.4

4.3 Empirical Modeling

Our empirical model is designed to capture key dimensions of our conceptual model (Figure 1) with sufficient accuracy and precision to enable robust testing of developed propositions. Model development proceeds in three stages: (1) normative adoption prediction model, (2) computation of misfit based on residuals, and (3) productivity regression.

In Stage 1 we estimate a linear adoption prediction model for IT innovation extent similar to that used in prior research (Cooper and Zmud 1990; Fang et al. 2005; Ramamurthy et al. 2008) with ENT adoption as the dependent variable and prior-period labor productivity, size, skill, capital to labor ratio, multi-plant status, plant age, and industry controls as independent variables. A firm with a favorable score on these antecedents has the normative profile of an innovator, which means that it is better-positioned in terms of needs and abilities to exploit the benefits of earlier/more extensive ENT adoption. For example, firms with more skillful employees are likely to be better at innovating and hence have a more favorable score, ceteris paribus. By contrast, a firm with unfavorable scores has the profile of an innovation laggard, one that is relatively poorly positioned to benefit from earlier/more extensive ENT adoption. Appending an error term, we obtain the following empirical model:

\[ ITInnov = \beta_0 + \beta_1 \text{Prod} + \beta_2 \text{Skill} + \beta_3 \text{MP} + \beta_4 \text{Age} + \beta_5 \text{CapLab} + \beta_6 \text{Size} + \text{Control} + \varepsilon_{it} \]  

(1)

The residuals from Stage 1 are key to our operationalization of misfit, which we compute in Stage 2 of our analysis. A large residual above the line indicates that the firm is innovating beyond the level

---

4 Use of continuous scales based on residuals tend to be highly correlated with ENT adoption, leading to multi-collinearity problems. Details are provided in the next section.
suggested by its normative innovation profile, i.e., over-adoption. However, it is an open empirical question as to whether the residuals exhibit small or large variance, and hence, what might reasonably be considered “large.” Likewise, it is an open conceptual question as to how we might deem what an appropriate cutoff is to distinguish right adopters from over adopters. Given our priors, we might assume that the residuals will be Gaussian, and that the tails of the distribution might well represent over or under adoption. Another approach is to set a baseline cutoff of 25% and conduct a sensitivity analysis at different cutoffs. As there is no prior framework of misfit on which to build, we choose the latter approach due to its generality. Finally, there is the issue of whether to use the raw residuals as a continuous measure of misfit, or, whether to perform a transformation. Based on the potential for multicollinearity, and given the paucity of prior research, we chose to trichotomize the residuals. We classified over-adopters as those plants with residuals falling in the top quartile for the sample as a whole, and under-adopters as those in the bottom quartile.\(^5\) The top 25% and bottom 25% cuts-offs correspond to substantial divergences between expected and actual behaviors. We emphasize that the objective of this paper is to demonstrate that such cutoffs exist with implications for performance, rather than propose or test a specific magnitude of cutoff.

It may be the case that our measure of misfit is picking up nothing but measurement noise in the regression. If this is the case, we would not expect to see any association whatsoever between IS innovation misfit and organizational performance. In contrast, if we were to observe an association, one that is consistent with our propositions implications of misfit, this would lend support to our theory—though it would not prove causality, a point to which we shall return in the next section.\(^6\)

Finally, in Stage 3 we construct a labor-productivity model examining the effects of innovation misfit on the returns to IT innovation adoption. The baseline model is based on prior studies of IT labor productivity (including those based on the CNUS dataset) and includes productivity as the dependent variable and IT innovation, skill, capital to labor ratio, multi-plant status, materials to labor ratio, and

\(^5\) Results are robust to alternative cutoff levels, including 10, 20, and 30 percentile.

\(^6\) We thank a seminar participant at UC Irvine for underscoring this useful counterfactual.
industry controls as independent variables. Then we examine main effects (Proposition 1) and moderation effects (Propositions 2 and 3) by adding dummy variables for over- and under-adoption and interaction terms between these dummies and IT innovation as measured by ENT adoption. (Note that when dummies are simultaneously included for both over- and under-adoption, then these dummies capture divergence from the base case, which is represented by firms residing in inter-quartile range on the adoption prediction residuals, for which there is no dummy included. These cases are treated here as right-adopters.) The empirical model is as follows:

$$\text{Prod} = \beta_0 + \beta_1 \text{MisfitLo} + \beta_2 \text{MisfitHi} + \beta_3 \text{ITInnov} + \beta_4 \text{Skill} + \beta_5 \text{CapLab} + \beta_6 \text{MP} + \beta_7 \text{MatLab} + \text{Control} + \epsilon_i$$  (2)

To recap, a negative and statistically significant coefficient on the Misfit X ITInnov terms in (2) combined with a statistically significant main effect of ITInnov will provide support for Proposition 2, Proposition 3, and Proposition 3b.

4.4. Results

Summary Statistics

Summary statistics for principal variables are presented in Table 4 below. We provide the means and standard deviations for the whole dataset (N=31,436). To give a sense of the difference between the profiles of innovation leaders and laggards, we also provide the means and standard deviations for plants that adopted 0-2 ENTs (low innovation, N=17,643) versus plants that adopted 3-6 ENTs (high innovation, N=13793).

Due in part to the large size of the dataset, all but a few correlations are significant. As expected, ENT adoption has positive and significant correlations with size, multi-plant, skill, and capital intensity. It is also positively and significantly correlated with labor productivity, plant age and materials intensity. (Note that while the positive direct correlation with plant age is counter to expectations, the relationship turns negative after controlling for other variables in the adoption regression.)

As expected, labor productivity is positively and significantly correlated with multi plant, skill, capital intensity, and materials intensity. It is also positively and significantly correlated with plant size and age, though the correlations are quite small.
Table 4: Descriptive Statistics and Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>MEAN (SD)</th>
<th>CORRELATION COEFFICIENT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All (N=31436)</td>
<td>Low Innov. (N=17,643)</td>
</tr>
<tr>
<td>1. ITINNOV</td>
<td>2.229 (1.307)</td>
<td>1.260 (0.748)</td>
</tr>
<tr>
<td>2. PROD (LP)</td>
<td>275.037 (375.149)</td>
<td>230.699 (307.930)</td>
</tr>
<tr>
<td>3. SIZE</td>
<td>232.774 (496.229)</td>
<td>136.535 (340.829)</td>
</tr>
<tr>
<td>4. MP</td>
<td>0.668 (0.471)</td>
<td>0.530 (0.499)</td>
</tr>
<tr>
<td>5. AGE</td>
<td>20.166 (7.092)</td>
<td>19.641 (7.212)</td>
</tr>
<tr>
<td>6. SKILL</td>
<td>0.278 (0.184)</td>
<td>0.274 (0.182)</td>
</tr>
<tr>
<td>7. CAPLAB</td>
<td>102.973 (212.074)</td>
<td>79.534 (190.024)</td>
</tr>
<tr>
<td>8. MATLAB</td>
<td>148.440 (252.779)</td>
<td>123.847 (208.704)</td>
</tr>
</tbody>
</table>

*p < .1; ** p < .05; *** p < .01
Note: High innovation denotes top half of innovators; low innovation defined similarly.

Regarding our principal variable, \( ITInnov \), we computed tetrachoric correlations between all six of its dichotomous components. All these correlations were positive and significant with the exception of two involving the “other” item, which had very low magnitude. Principal components analysis on the six dichotomous variables revealed that almost half of the variation is explained by the first component. All six items loaded positively on this first component, consistent with a unidimensional construct.

**Stage 1 Regression**

The results of the ENT adoption prediction model (Equation 1) are presented in Table 5. As expected, the coefficients are positive and significant (\( p < .01 \)) for size, multi-plant, skill, and capital intensity, and negative and significant for plant age. Also the coefficient for prior-period labor productivity is positive and significant. We also estimated an alternative model based on valued-added rather than output. A similar pattern of results obtained.

Although variance explained by \( R^2 \) is fairly high (31%) there is still some variance that is not explained.

As discussed earlier, unexplained variance arises from innovation misfit (possibly arising from mindless adoption), but could also be due to several other factors, including other unobserved normative and non-adoption predictors (that are not strongly correlated with included variables), measurement error, and unexplainable random variation. Plotting residuals versus fitted values and use of standard diagnostics
such as White’s test reveal no dependencies between fitted values and residuals, validating our use of residuals as a measure of innovation misfit.

### Table 5: Adoption Prediction Model

<table>
<thead>
<tr>
<th></th>
<th>Coefficient/Std. Err.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DV = Ent Adoption</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.698*** (0.092)</td>
<td></td>
</tr>
<tr>
<td>Labor Productivity</td>
<td>0.121*** (0.010)</td>
<td></td>
</tr>
<tr>
<td>Skill</td>
<td>0.121*** (0.010)</td>
<td></td>
</tr>
<tr>
<td>Multi-Plant</td>
<td>0.385*** (0.016)</td>
<td></td>
</tr>
<tr>
<td>Plant Age</td>
<td>-0.004*** (0.001)</td>
<td></td>
</tr>
<tr>
<td>Capital-Labor</td>
<td>0.067*** (0.006)</td>
<td></td>
</tr>
<tr>
<td>Plant Size</td>
<td>0.309*** (0.004)</td>
<td></td>
</tr>
<tr>
<td>Industry</td>
<td>-</td>
<td>R²</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.311</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td>31,436</td>
</tr>
</tbody>
</table>

*p < .1; ** p < .05; ***p < .01
Note that labor productivity, skill, capital-labor, and size are log-transformed to mitigate nonlinearities. Also, labor productivity is lagged by two years because one rationale for adoption of technology is that it is a response to historically lower productivity. Capital-labor is lagged by two years due to data limitations.

### Stage II Regression

The results of the labor productivity models (Equation 2) are presented in Table 6. Model I is the baseline model, Model II adds ENT Adoption, and Model III adds ENT Adoption and the Over- and Under Adoption Misfit direct and interaction effects. As expected, skill, capital intensity, materials intensity, and multi-plant are positive and significant across all three models. ENT adoption is positive and significant (p < .01) in Models II and III, in support of Proposition 1. The coefficient for ENT adoption x Misfit Hi is negative and significant (p < .01), suggesting that over-adoption misfit negatively moderates the effect of ENT adoption on organizational performance, in support of Proposition 2. The coefficient for ENT adoption x Misfit Lo is also negative and significant (p < .05), which in support of Proposition 3b that under-adoption misfit negatively moderates performance. The coefficient for Misfit Hi is negative and weakly significant (p < .1), while that of Misfit Lo is significant (p < .01). We interpret these results in Section 6. All models explain about 80% of variance in labor productivity.
Table 6: Labor Productivity Model

<table>
<thead>
<tr>
<th>DV: Labor Productivity</th>
<th>Model I</th>
<th>Model II</th>
<th>Model III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.342*** (0.034)</td>
<td>2.283* (0.034)</td>
<td>2.118*** (0.036)</td>
</tr>
<tr>
<td>Skill</td>
<td>0.015*** (0.003)</td>
<td>0.012*** (0.003)</td>
<td>0.005*** (0.003)</td>
</tr>
<tr>
<td>Capital-Labor</td>
<td>0.084*** (0.003)</td>
<td>0.081*** (0.003)</td>
<td>0.075*** (0.003)</td>
</tr>
<tr>
<td>Multi-Plant</td>
<td>0.109*** (0.005)</td>
<td>0.096*** (0.005)</td>
<td>0.065*** (0.006)</td>
</tr>
<tr>
<td>Materials-Labor</td>
<td>0.558*** (0.004)</td>
<td>0.554*** (0.004)</td>
<td>0.546*** (0.004)</td>
</tr>
<tr>
<td>Innovation Extent</td>
<td>0.109*** (0.005)</td>
<td>0.096*** (0.005)</td>
<td>0.065*** (0.006)</td>
</tr>
<tr>
<td>MisfitLo</td>
<td></td>
<td></td>
<td>0.148*** (0.012)</td>
</tr>
<tr>
<td>MisfitHi</td>
<td>-0.033* (0.020)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Innovation Extent *</td>
<td>-0.013** (0.007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MisfitLo</td>
<td>-0.030*** (0.006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industry</td>
<td>0.799</td>
<td>0.801</td>
<td>0.803</td>
</tr>
<tr>
<td>N</td>
<td>31423</td>
<td>31424</td>
<td>31423</td>
</tr>
</tbody>
</table>

*p < .1; **p < .05; ***p < .01; Note that labor productivity, skill, capital-labor, and materials-labor are log-transformed to mitigate a nonlinear specification error observed in early regressions. Capital-labor is also lagged due to data limitations (high serial correlation mitigates potential error). Results are robust to alternate cutoffs for misfit, including 20% and 30%.

Robustness Checks

We used standard diagnostics to determine the extent to which our data satisfy assumptions of ordinary least squares regression for the ENT adoption and labor productivity regressions. We used scatter plot matrices, stem and leaf plots, and Cook’s D to examine outliers, influence (removing the observation substantially changes the coefficient estimates), and leverage (observation with extreme value on predictor). Based on these visual and statistical tests, no data points were removed. We plotted residuals and used q-q plots to examine normality. Using the Shapiro-Wilkes test, we could not reject the null of normality, providing support for robustness of inferences. Using White’s test, we found a small degree of heteroscedasticity in the labor productivity model (Stage 2), so we reran the specification with variance-corrected errors: no difference in the pattern of results was observed. Finally, the VIF test was used to rule out potential errors due to multicollinearity. The test was < 2 for all main variables, suggesting that multicollinearity is not a problem.

To complement our analysis based on interactive terms, we ran a split samples analysis. Specifically, we estimated Model II from Table 6 separately for the right-adoption (N=15,710), over-adoption (N=7,855),
and under-adoption (N=7,858) groups. With the exception of ENT adoption, the estimated coefficients for all independent variables were similar across the three groups and with the full sample results (although skill turned insignificant, probably due to the lowered sample sizes). In contrast, the coefficients for ENT adoption were positive and significant across all splits, but considerably higher for the right-adopter split (.144) than for the over-adopter (.083) and under-adopter (.060) splits. Thus, the split sample results reinforce the results from the interactive model.

We varied our assumptions to determine their impact on the sensitivity of results (Leamer 1983). First, we tried alternative values of misfit cutoff (20%, 25%, and 30%). We also used alternative measures of plant performance, including productivity based on total value of shipments as well as value added. The overall pattern of results did not change.

5. Discussion
5.1 Summary of Results
First, consistent with prior research we found greater ENT adoption among firms that were larger, had a higher ratio of non-production to production workers, were part of multi-plant firms, were more capital intensive, and were from manufacturing industries that typically lead with IT innovation (e.g., computers, and printing and publishing). Such firms are more likely to have a greater ability to innovate successfully due to greater resources and scale, greater managerial sophistication and professionalism, and greater skills and expertise. Second, we found greater productivity (as expected) among firms that were more capital and material intensive, were part of multi-plant firms, had higher had a higher ratio of non-production to production workers, and that resided in certain industries. The positive influence of capital and material intensity are standard results, and our coefficients are in line with prior results in general and those from CNUS dataset. (Black and Lynch 2001; Atrostic and Nguyen 2005). Managers in multi-plant firms tend to be more sophisticated and professional, which suggests that other things equal, they will tend to use technology in more effective ways. The positive effect for our skill proxy makes sense because IT requires skill to implement and requires more skill on an operational basis once implemented.
Third, as expected we found a strong positive association between ENT adoption and labor productivity, in support of Proposition 1. This is consistent with a long line of research showing a positive relationship between IT investment and productivity in general (Lichtenberg 1995; Brynjolfsson and Hitt 1996; Banker et al. 2006) and with prior analysis of CNUS data in particular (Atrostic and Nguyen 2005). The ENT coefficient of .111 (see Model III in Table 6) means that for each incremental ENT adopted, logged labor productivity goes up by .111. This corresponds to an increase, for each incremental ENT adoption, of 2.0% in logged labor productivity (starting from the sample mean of 5.62) and 11.7% in unlogged labor productivity (starting from sample mean of 275).  

Fourth, consistent with Proposition 2, we found that over-adopter misfit negatively moderates the relationship between ENT and performance. The coefficient for the interaction of over-adopter misfit with ENT adoption is -.030, suggesting a net ENT coefficient for over-adopters of .111-.030=.081. This means firms in the over-adopter group receive about an 8.3% boost for each incremental adoption (starting from the sample mean of 275), which is about 28% lower than what right-adopting firms receive (other things equal). As argued earlier, the performance penalty of over-adoptive misfit arises from two potential sources. First, firms with an over-adoptive misfit have lower scores on the innovation profile than right-adopting firms with the same level of adoption. Such firms are less likely to have the benefits of greater size, scale, sophistication, expertise and slack resources demanded by the level of ENT adoption they have chosen. They are also likely to reside in industries that have the greatest potential to exploit the benefits of more aggressive IT adoption. Second, over-adoptive misfit implies a lack of mindfulness in making adoption decisions. Managers in these firms chose to take the lead in adopting ENT even though they did not have especially favorable scores on variables comprising the innovation profile. Granted, some of these firms may have had compensating characteristics that were not captured in our model, but to the extent that this is the case, it implies that our results actually understate the negative impact of innovation misfit and mindlessness on innovation returns.

---

7 This large main effect may be attributed to 1) the direct effect of increased use of electronic networks; 2) if a firm has adopted more innovations as of a given date, they are likely to have adopted each of them earlier (on average), and to be using them more extensively (on average) than those that have adopted fewer innovations as of given date; and 3) potential complementarities among the ENTs.
Fifth, we also found a smaller but significant negative moderation effect for under-adoption misfit, suggesting support for Proposition 2b. The coefficient for the interaction of under-adoption misfit with ENT adoption is -.013, suggesting under-adopters have returns to ENT adoption that are 12% lower right adopters (starting from the sample mean of 275). As argued earlier, an under-adoption misfit has ambiguous causality: on the one hand, under-adopters have scores on the innovation profile that are more favorable than they “should” have, and this would suggest improved returns compared to right-adopting laggards. But on the other hand, under-adoption is a sign of a lack of mindfulness, which would suggest worse returns. Our results indicate the latter effect may be stronger in this study.

Sixth, of all our empirical results, there is just one that represents a puzzle: the significant and positive coefficient for the under-adoption dummy variable of .148. We had expected this coefficient to be insignificant (as was the coefficient for the over-adoption dummy). A positive coefficient for the under-adoption dummy suggests a fixed positive effect from membership in this group that is independent of the number of ENTs adopted. We have previously argued that there could be some potential positive productivity effects from under-adoption (i.e., higher scores than expected on the innovation profile). Also, it makes sense that the countervailing negative mindlessness effect might diminish in the presence of the kinds of population-level bias towards over-adoption we expect is more likely in this study due to the dot-com bubble. However, it is difficult to see how this would net out to a fixed productivity benefit from under-adoption unrelated to the level of ENT adoption, or how this effect squares with the negative moderation effect.

5.2 Summary of Contributions

We summarize the conceptual contributions of this research as follows. First, we conceptualize the notions of innovation fit and innovation misfit. We began by formalizing the concept of a normative innovation profile characteristics, which distinguish expected innovation leaders and laggards in some context, and then defined innovation fit and misfit based on the discrepancy between predicted innovation behavior (based on the normative innovation profile) and actual innovative behavior. Second, we
developed theoretical connections between innovation fit and innovation mindfulness. We posited that mindfulness—i.e., the extent to which firms take their individual “fact and specifics” into account in making innovation decisions—tends to promote higher degrees of innovation fit, while mindlessness leads to misfit. Third, we explicated various reasons for why misfits might be observed in practice and suggested a privileged role for the effects of managerial fashions in the context of IT adoption. Fourth, we developed theoretical rationales linking two types of innovation misfit—over-adoption misfit and under-adoption misfit—to firm performance. Under the assumption there is no population bias, we posited that over-adoption misfit will have an unambiguously negative effect on innovation returns, both because such firms are structurally poorly positioned to adopt, and also because low innovation fit is a sign of innovation mindlessness. In contrast, we argued that under-adoption misfit has an ambiguous effect on innovation returns, because even though under-adopters are structurally well positioned to adopt, they are more likely to lack mindfulness in adoption timing decisions, and by extension may be less mindful during implementation. Finally, we considered how the effects of a population bias towards too-early or too-late adoption might affect the link between innovation misfit and performance. We argued that a population-level bias towards too-early adoption will further magnify the ill effects of firm-level over-adoption, but will moderate the ill-effects of firm-level under-adoption. A population-level bias toward too-late adoption will have the opposite effect, moderating the ill-effects of over-adoption and magnifying the ill-effect of under-under adoption.

Our work also makes several empirical contributions. First, we develop an operational measure for innovation misfit using the residuals from an adoption prediction model, the idea being that firms with large residuals have a mismatch between predicted and actual innovative behavior. While not a perfect measure, it is consistent with a large stream of prior work that at least implicitly assumes that those firms that do take the lead in innovation tend to be those that, normatively speaking, should take the lead. It also has the advantage of being relatively easy to capture. (We consider the limitations of this measure and possible alternatives below). Second, we use a large dataset on the adoption of ENT in manufacturing to model the determinants of both innovation adoption and business value. We confirm prior results linking...
several variables to adoption, and more important, confirm our posited negative effect of over-adoption innovation misfit on returns to adoption. Also, our study is notable for being one of the first to examine IT innovation antecedents, IT adoption, and business value in a truly holistic way. Also it is one of only a few large-scale studies to demonstrate the extent of labor-productivity benefits of electronic networking technologies in manufacturing.

5.3 Implications for Research and Practice

For researchers, we break new ground in explaining the link between IT and performance. Prior studies have variously argued for a: (1) a direct link between IT and performance (Brynjolfsson and Hitt 1996), (2) an indirect link operating through mediating variables such as capabilities (Banker et al. 2006), and (3) a contingent link that depends on other variables, such as the presence of complementarities (Zhu 2004). Our study provides additional evidence for a contingent link between IT investment and performance. As expected, we show that innovation returns are substantially diminished among firms that exhibit an over-adoption misfit, and somewhat diminished among firms that exhibit an under-adoption misfit. This suggests that innovation fit/misfit holds promise as an important new variable in explaining the complex linkage between IT innovation and performance. Prior findings that IT adoption has a direct correlation with firm-level profit measures (e.g., (Hitt and Brynjolfsson 1997)) might be viewed as somewhat puzzling, in that it seems to imply a uniformly “more is better” mentality with regard to innovation. Our results indicate a more complex situation, where the extent to which “more is better” depends on the degree to which a firm fits the profile of a leading innovator.

Our results also shed some additional light on why the cumulative adoption so often follows an S-shaped pattern. Traditional innovation theory (Rogers 2003) holds that communication patterns produce this shape. In particular, news of the existence and benefits of an innovation are assumed to take different times to reach different actors (through mass market media and word-of-mouth), with the interplay between these two channels producing the distinctive S-shape. Attewell posited an alternative explanation, which is that firms sort themselves out based on ability to adopt, and wait until adoption
barriers have been sufficiently lowered. Assuming that the ability to adopt in a population of firms falls along a bell-shaped distribution, then if firms do sort themselves out according to ability to adopt, the would result in an S-shaped cumulative adoption curve. For our model to work, it is imperative for this self-sorting process to occur to some extent; otherwise the adoption prediction vector would not in any way reflect the ability to adopt. That fact that our model behaves as expected lends additional weight to the ability-based explanation for the ubiquitous adoption S-curve.

5.4 Limitations and Future Research

This study is subject to two potential limitations. First, our operational measure of innovation fit relies on an empirically derived vector of adoption predictors. Our approach intermingles genuine profile deviations with apparent deviations caused by measurement errors (Meilich 2006), and therefore assumes a fairly low incidence of errors in measuring profile characteristics. Since our data concern objectively observable quantities and were collected by the US Bureau of the Census we believe that respondents would be especially willing and able to give accurate answers, but this many not be true in other research. Also, our approach assumes that economically rational behavior is prevalent enough to drive the fitted results, but not so prevalent that few misfits remain to actually affect performance (Meilich 2006). We believe that in many situations these assumptions will hold; but researchers should be on guard for situations where it may not. Also, it important to note that these problems would tend to attenuate, rather than exaggerate the estimated impact of misfit (Dewar and Werbel).

A potential second limitation is that due to our use of secondary data, we had to rely on some comparatively generic innovation profile characteristics, rather than characteristics that were tailored to the specifics of the innovations at hand. For example, instead of a rich, context-specific measure of skills and knowledge, we used the ratio of non-production workers to production workers as a proxy measure. Likewise, in lieu of direct measures of managerial professionalism and sophistication, we employed the multi-plant designation as a proxy. In some respects it is useful to be able to show that even fairly generic—and thus more widely available—characteristics result in models that behave as expected. As
things stand now, some organizations no doubt were mistakenly classified as over-adopters because our proxies did not adequately capture their true positions on underlying theoretical constructs, or because they possessed favorable positions on characteristics that were excluded from our model specification (and that did not strongly correlate with those that were included). In a more complete model, there would fewer such firms. Therefore, we suggest that future work investigate whether the use of more complete and/or tailored profile characteristics does indeed lead to increased effects of misfit on performance.

Beyond addressing the potential limitations just mentioned, we suggest some additional avenues for future work. First, there is the question of what kinds of firm-level conditions lead to innovation fit and misfit? We have posited that innovation mindfulness and mindlessness as key antecedents. A number of directions are available to investigate this posited linkage and its implications. We view innovation mindfulness as one of the most important theoretical constructs to emerge from the IT innovation literature in the last decade, however as of yet no measures have been developed for the construct. We suggest that innovation fit/misfit could be positioned as an outcome variable in a nomological network used to evaluate the validity of measures for mindfulness/mindlessness. If the strong expected linkage were found, this would lend more credence to measures of both mindfulness and innovation fit as valid reflections of their underlying theoretical constructs. Also, a direct measure for mindfulness would allow us to sort out the ambiguous effects of under adoption misfit on performance by allowing us to decouple the positive structure effect of being “over qualified” from the negative project execution effect of being less mindful with regard to innovation. Finally, if a strong correlation were found between mindfulness/mindlessness and fit/misfit, then each might be viewed as suitable proxy for the other in future research. Each variable is quite challenging to capture in practice, but for different reasons: mindfulness/mindlessness requires primary data collection to capture directly, while fit/misfit requires data sufficient to estimate an adoption model. So we can easily foresee circumstances where the theoretical interest is in one construct, but the other is far more feasible to capture empirically.
Another important question concerns the macro-level conditions leading to pervasive over- or under-adoption misfits in a population of firms. Prior work related to institutional theory, managerial fashions, and information cascades has addressed the question of why many firms seem to ignore situational specifics in making managerial decisions. But little of this research has been directed specifically at the domain of IT innovation. We believe our work gives additional motivation to examine these implications of these mechanisms in an IT adoption context.

References


