Balancing Incentives in Pharmaceutical Research

Iain M. Cockburn, Boston University and NBER

Rebecca M. Henderson, MIT Sloan School and NBER

Scott Stern, Northwestern University and NBER

This revision January 2004

The cooperation and assistance of the firms that contributed the data used in this study is greatly appreciated. Jeff Furman, Pierre Azoulay, and Mercedes Delgado provided outstanding research assistance. We thank Robert Gibbons, Bengt Holmstrom, Tom Hubbard, Jonathan Levin, Dan Levinthal, Fiona Scott Morton, David Mowery, seminar participants at numerous universities and conferences and two anonymous referees for helpful comments; however the conclusions and opinions expressed in this paper are our own, and any errors or omissions remain entirely our responsibility. This research was funded by POPI, the Program for the Study of the Pharmaceutical Industry at MIT and by the MIT Center for Innovation in Product Development under NSF Cooperative Agreement Number EEC-9529140. Their support is gratefully acknowledged. Stern completed this research while a Health and Aging Fellow at the National Bureau of Economic Research, whose hospitality is gratefully acknowledged. All correspondence to: Professor Iain Cockburn, School of Management, Boston University, 595 Commonwealth Avenue, Boston, MA 02215, cockburn@bu.edu.

Balancing Incentives in Pharmaceutical Research

ABSTRACT

When effort is multi-dimensional, firms will optimally "balance" the provision of incentives. Setting high-powered incentives along one dimension raises the returns to providing high-powered incentives along other dimensions which compete for the worker's effort and/or attention (Holmstrom and Milgrom, 1991). We test for this effect in the context of for-profit pharmaceutical laboratories using detailed data on individual research programs. Consistent with this complementarity hypothesis, there is both cross-sectional and time-series evidence that firms providing strong promotion-based incentives for scientists to invest in basic research are more likely to provide strong incentives to supply effort towards applied research.

JEL Classification Numbers: L2, L65, O32.

Keywords: innovation, incentives, multitasking, complementarity.

I. INTRODUCTION

Recent research has described firms both as systems of complementary activities (Milgrom and Roberts, 1990; Ichniowski, Shaw and Prennushi, 1997) and as systems of complementary incentives (Holmstrom and Milgrom, 1991, 1994; Baker, Gibbons, Murphy, 1994, 2001). This research suggests that in many contexts, complementarities between organizational practices may arise from the contracting problems inherent in a multi-task agency setting. For example, when output is generated by workers (or work groups) exerting effort across two or more different tasks, the firm will optimally "balance" incentives across these tasks. If it does not, workers will inefficiently allocate too much effort towards those tasks with the highest marginal return to them. Indeed in some circumstances it may be optimal to pay nothing but a fixed wage to avoid this problem (Holmstrom and Milgrom, 1991, henceforth "H&M"). This idea provides valuable insight into the observation that adopting a specific organizational practice in isolation often fails to pay off.

Despite the importance of these ideas for our understanding of the theory of the firm, empirical characterizations of multi-dimensional incentive systems are surprisingly scarce. In part, this may be because in internal organizational settings where multiple dimensions of effort matter, measuring comparable incentive instruments across firms and over time requires detailed firm-level data that is difficult to obtain and interpret.

Anderson and Schmittlein (1984) provide an early study of how the incentives of sales agents relates to factors such as the degree of monitoring and whether the salesperson is a long-term employee. More recent cross-sectional studies explicitly test for complementarity in incentives *across* firm boundaries (Baker and Hubbard, 2002; Slade, 1996; Brickley, 1999). In a study of gasoline retailers, Slade (1996) provides evidence that differences in non-gasoline service offerings, such as a convenience store, influence the incentives provided by gasoline wholesalers. In the study most closely related to ours, Baker and Hubbard (2004) also exploit a shock to monitoring technology (for truck drivers) to test for the presence of a multi-task incentive system; however, whereas Baker and Hubbard exploit a literal shift in the monitoring technology (resulting from the diffusion of novel information technology), our approach is based on observing a shift in the nature of the activities undertaken by pharmaceutical research firms which resulted in a more precise monitoring technology.

In particular, this paper models and tests the complementarity hypothesis *within* firm boundaries in the context of pharmaceutical research. Using a detailed dataset compiled from extensive fieldwork and the internal records of a sample of nine representative firms over 15 years we move beyond cross-sectional approaches common to prior studies to evaluate the complementarity hypothesis by exploiting variation both within and across firms in our sample.

Pharmaceutical research provides a particularly interesting setting in which to explore multi-task agency problems. In the first place, although the question of exactly how incentives are provided for basic and applied research has important implications for the rate of technological innovation (Romer, 1990; Lazear, 1996), there is little systematic empirical evidence about how such incentives are provided in industrial laboratories (Hauser, 1998). Cohen and Levinthal, (1989), for example, suggest that a firm's ability to invest in absorptive capacity may have important implications for research productivity, but we know little about how firms might motivate their researchers to make these kinds of investments. For IO and organizational economists, understanding how firms provide incentives to *internal* researchers (who relinquish intellectual property claims on discoveries made during their employment) is important for understanding the conditions under which R&D will take place in the confines of an integrated firm (Holmstrom, 1989; Aghion and Tirole; 1994; Lerner and Merges, 1998; Gans and Stern, 2000).

The long-run level of research productivity in pharmaceutical drug discovery depends on the level of effort devoted towards two distinct activities: short term, "applied" effort devoted to the discovery and synthesis of chemical compounds, and longer term, more "basic" effort devoted to improving the scientific competencies of the firm and the odds that any particular chemical compound might become a profitable drug.

Effort devoted to short terms, or applied, research has always been relatively easy to measure since it can be approximated by the number of patented compounds generated by a research group. Historically, however, it has been difficult to measure effort devoted to longer term, more "basic" research, and prior to the late 1970s, firms in consequence avoided the use of high powered incentives, rewarding their researchers largely through the use of fixed salaries. Starting in the late 1970s, however, the pharmaceutical industry experienced a significant exogenous shock to the technology of drug discovery which made monitoring of effort towards long-term activities more effective (and may have also changed the absolute level of returns to the long term research

component). In particular, during this period, it became feasible to use a monitoring system based on the provision of incentives to publish in the scientific research literature as a signal for effort devoted to the development of long-term research competencies.

Pharmaceutical research thus transitioned from a world in which firms needed to reward two competing activities (short term applied and longer term basic), only one of which could be measured with any precision, to a world in which two mutually complementary activities continued to be important but in which not only the first but also the second could be approximately measured. Firms responded heterogeneously to this "shock": while some firms quickly adopted a research organization which provided high-powered promotion-based incentives designed to encourage efforts in basic research, other firms were much slower to do this, eschewing internal incentives based upon basic research outputs (such as scientific publications) well into the 1990s. We use this heterogeneity among firms and over time to evaluate whether firms offering high-powered incentives for basic research were more likely to provide higher-powered incentives for applied research. As such, we are evaluating an important implication of the multi-task model: in response to an exogenous shock which shifts the ability to measure effort exerted along one dimension, do firms increase the incentive intensity for other tasks competing for workers' time?

Using data from nine firms over fifteen years we establish three results. First, we demonstrate substantial variation among firms and across time in the intensity with which they provide incentives for basic or long term research. The primary mechanism used to do this appears to have been the internal labor market of the firm. By actively taking advantage of the monitoring capabilities of the public research community, and rewarding research workers' participation in "open science" through practices such as using publication in the refereed literature as a criteria in promotion decisions, some firms provided powerful incentives to supply effort along this dimension. Other firms did not use these practices, or applied them less intensively. Second, we find evidence for significant variation in the provision of high-powered incentives to do applied research: some firms rewarded research teams with substantially higher budgets following better-than-expected (important) patent output, while in others this effect was much more muted. Third, we find evidence in a variety of "cuts" of the data for a quantitatively and statistically significant positive association between the use of these two instruments.

The correlation between the incentives for short and long term research subsequent to the

transition to "rational" drug design, may, of course, be due to factors unrelated to the firm's response to the multitask agency problem. Rather than rely on a single argument for identification (for example, by simply assuming the exogeneity of certain instruments), our approach is to identify the most likely sources of bias and to provide direct controls for these effects. For example, after showing the presence of a positive correlation in the context of a pooled data analysis, we demonstrate an even stronger positive correlation in a more demanding "differences-in-differences" estimator, including fixed effects for each individual research program along with time trends for each therapeutic area. While the limited number of firms in our sample makes us cautious about over interpreting these results, we view them as supporting the H&M hypothesis about the role of "balance" in the provision of research incentives.

The paper begins with a discussion of the nature of drug discovery research, the change that occurred in the 1970s and its implications for the management and organization of pharmaceutical research. Section III reviews the H&M multitask agency model and derives empirical implications. Section IV discusses our data and some of the reduced form properties of the data set. Section V reviews our empirical findings, and Section VI concludes. A supporting appendix discuss the construction of each of the incentive measures in greater detail.

II. EFFORT & INCENTIVES IN PHARMACEUTICAL RESEARCH

The process of drug discovery and development is complex and extends over several years. In the "research" phase, also referred to the drug discovery process, researchers attempt to find compounds that may plausibly be developed into drugs by demonstrating their therapeutic effects in animals. In the second, or "development" phase, these compounds are tested in humans and undergo rigorous review by the Food and Drug Administration. The two phases require distinct skills and knowledge and are nearly always carried out by quite different people. In this paper, we focus only on the research phase.

For much of this century, the technology of drug research was dominated by a technique commonly described as "random" drug discovery. Under this regime, large numbers of candidate compounds would be tested for pharmaceutical activity in an "animal model" or a relatively crude cell culture or assay. For example, a search for hypertensive therapies might involve injecting large numbers of candidate compounds into hypertensive dogs to explore the degree to which they

reduced blood pressure, while a search for therapies effective against anxiety might involve administering compounds to rats and then observing their behavior in stressful situations.[1] Molecules showing pharmaceutical activity would then be subjected to further testing, and modified to improve their pharmacological properties. In most cases, the "mechanism of action" — the specific biochemical and molecular pathways responsible for a compound's therapeutic effect — was not well understood. While random drug discovery was not entirely divorced from more fundamental scientific research conducted within the public sector, in general it was not critical that pharmaceutical researchers be active players in the community of public science.

While given the name "random" (largely by those seeking to pioneer the next generation of drug discovery technology), this method of drug discovery was neither unproductive nor unscientific. Many powerful drugs were discovered using this set of techniques, and it continues to be employed, particularly in those cases in which the biochemical or genetic mechanisms underlying a disease are not well understood. Effective "random" drug discovery is neither automatic nor unskilled. It requires, first, the development of the "applied" or "short term" ability to generate plausible candidates and to elaborate their chemical structure. (An effective group, for example, on finding a compound that appears to show some activity in an animal screen, will devote a great deal of effort to synthesizing related compounds that might show the same or better activity but that might also have superior pharmacological profiles.) This kind of activity can often be localized to particular research groups, and approximately measured by the number of patentable compounds that they generate. It also requires, however, a deep knowledge of medicinal chemistry (the ways in which particular chemical structures were metabolized and reacted to by the human body), and of both animal and clinical models and their interpretation. For example, a recent story in the WSJ explains how a Japanese researcher was able to identify a particularly promising compound by observing that it changed the way in which a rat's whiskers twitched (WSJ ref, 2003). This skill, while not "basic" or "fundamental" in the classic sense of university research, nevertheless shares many of its characteristics. Its impact tends to be long term in nature, and to be difficult to localize. Knowledge in medicinal chemistry, in clinical pharmacology, and in the nature of animal models

[1] For example, one test involved throwing rats into buckets of water and observing how long they continued to struggle.

may be useful across the firm in unpredictable ways. It is also extraordinarily hard to measure, being largely embedded in the tacit knowledge shared by experienced chemists and drug developers.

The need to encourage researchers to invest in both kinds of activity presented the pharmaceutical firms with a classic incentive design problem. In the ideal case one would of course wish to reward researchers on the basic of their real output – or in this case on the degree to which they discover highly profitable drugs. Unfortunately using drug profitability as a metric is problematic. It typically takes between 10-15 years to develop a new drug, and successful drug introductions are a classic problem in team production, drawing deeply on a wide range of "downstream" skills, including clinical trial design, statistics, manufacturing regulatory affairs and sales and marketing, so that drug profitability is inevitably a very noisy (and very delayed) measure of research effort.

The degree to which a research group generates patentable compounds is a reasonable measure of short term effort, but the vast majority of patented compounds prove to be worthless – current industry estimates are that only 2% of promising leads become marketed drugs, let along profitable drugs – and thus patent output alone, while a reasonable measure of the effort devoted to short term, applied research, is an unsatisfactory measure of what the firm would really like to measure – "excellent research" – or the appropriate balance of long and short term research.

In principle, pharmaceutical firms might have been able to address this dilemma by allocating the tasks of "basic" and "applied" work to different groups within the firm, with incentives tailored to each task, as H&M suggest. Hoffman-La Roche, for example, created the "Roche Institute" to pursue fundamental research in biological systems. However our fieldwork suggests that this approach had significant drawbacks, and it was not widely pursued. Such groups tended to degenerate into "ivory towers" – producing a large number of scientific papers but contributing little to the process of drug discovery. Effective adoption of "rational" drug discovery seems to depend on a tight integration between basic and applied research (Gambardella, 1995; Henderson and Cockburn, 1996). The dominant means used to accomplish this integration was to organize researchers into small teams (4-7 PhDs), responsible both for staying at the leading edge of their particular disciplines and for working together to translate this fundamental knowledge into

promising compounds.[2]

We believe that this organizational design creates exactly the kinds of tension that are captured in the H&M model. Managers of these research groups have to encourage workers to supply effort in both short and long term research activities. They cannot easily condition incentives on a useful measure of output, and their employees must optimally engage in two, complementary activities that compete for researchers' time. Under the "random" drug discovery regime it was particularly difficult to measure the degree to which researchers were engaging in "longer term" or "more basic" activity, and to our knowledge the vast majority of firms used low powered incentives, (largely salary and subjective promotion benefits), as H&M would predict.

This changed in the late 1970s. In 1978, Squibb announced the discovery of the anti-hypertensive drug Captopril. This marked a watershed in the technology of drug discovery, since it was the first drug to be discovered through the use of an *in-vitro* (literally "in glass" as opposed to *in-vivo*, or "in life") screen that duplicated a particular mechanism of action, rather than through the use of an animal model.[4]

This technology, commonly called "rational" or "mechanism based" drug discovery, offered a powerful new research tool, and was gradually adopted across the industry over the course of the next fifteen years. Research-oriented pharmaceutical companies began to make substantial investments in basic research in disciplines such as biochemistry or cell biology, and to invest much more heavily in understanding and accessing publicly funded science.

This change is critically important for our purposes here because it made it much easier to measure the degree to which drug researchers were engaging in "longer term" or "more basic" research activities. As research switched away from a model in which the important long term knowledge was largely idiosyncratic, tacit knowledge about animal models and the complexities of human physiology to a model in which the important long term knowledge was of much more of a

[2] There are, of course, exceptions to this generalization. For example many pharmaceutical firms currently maintain small groups of researchers charged with the development of expertise in genomics, a new area of science that will probably have a very significant impact on the drug discovery process.

4 For a fuller discussion of the discovery of Captopril, see Henderson (1994). Note that researchers had used speculation about drugs' mechanism of action as a research tool long before the discovery of Captopril. Sir James Black, for example, discovered the first of the beta-blockers in the early 1960s by exploiting his hypothesis that blocking the heart's beta receptors would lower blood pressure. But he did not make this discovery by screening

classically "scientific" nature – knowledge of fundamental genetics and biochemistry – so firms were able to take advantage of the community of public science to monitor their researchers much more effectively. Evaluating the quality of basic research is prohibitively difficult for managers who are not themselves at the cutting edge of the relevant science. Pharmaceutical firms attempting to provide incentives to perform this kind of basic research rely instead on the set of institutions that have evolved to evaluate publicly funded biomedical researchers.

The reward system of "open science" is based on publication, peer review and priority, with a clearly established public rank hierarchy in most disciplines. (Merton, 1973; Dasgupta and David, 1994; Stephan, 1996; Stern, 1999). Firms who encourage their research workers to participate in "open science" can benefit from this system in two ways. First, they can use worker's success in publishing in peer-reviewed journals and in garnering respect from their scientific peers as informative signals of the level of effort devoted towards basic research. Second, the rank order tournament aspects of this reward system translate straightforwardly into the tournament internal to the firm. By promoting researchers on the basis of their publication record and on their standing in the public rank hierarchy of their field, or on the criteria used by the publicly funded scientific community, a firm could provides high-powered incentives to supply effort directed towards this particular kind of long term or basic research.

For an individual researcher, effort devoted towards understanding fundamental biological principles is a substitute for short term or applied effort devoted towards translating scientific knowledge into the discovery of potential drugs. Staying at the leading edge of a scientific discipline requires devoting substantial effort to publication, basic laboratory work and to remaining connected to the wider research community. Translating this knowledge into the discovery of potentially commercializable new drugs, however, requires devoting effort to working in an interdisciplinary applied research team. Rewarding researchers solely on the basis of their ability to work as part of this team and to generate immediate output increases the risk that the researchers will fail to make the time-consuming effort intensive actions required to be at the leading edge of fundamental science, or that they will attempt to free ride on the scientific work of others. Similarly, only rewarding effort devoted to basic research might lead researchers to focus solely on advancing their own careers at the expense of effort that might be productively invested in the search for new

---

compounds against isolated beta receptors.

drugs. We argue, therefore, that the adoption of rational drug discovery provides an intriguing setting natural experiment that allows us to test H&M's hypothesis.

III. A MODEL OF INCENTIVES IN PHARMACEUTICAL RESEARCH

Theoretical work on incentive contracting has generated a number of important propositions about the structure of contracts between principals and agents in situations where the agent is required to perform multiple tasks (Holmstrom and Milgrom, 1991, 1994; Baker, 1992).[5] One of the most salient propositions is that in these multitask agency settings the optimal incentive regime is "balanced" — the degree to which high-powered incentives can be offered on any single dimension will be a function of the precision with which effort on every other important dimension can be measured, so that the degree to which high-powered incentives will be offered along any one dimension will be correlated with whether high-powered incentives are offered along other relevant dimensions. To see this more clearly, we briefly review the H&M model, adapting their general framework to the specific setting of long and short term research in pharmaceutical drug discovery.

We begin with a simple model of the provision of incentives for research workers in an employment relationship (i.e., one in which the firm hires the workers and owns the output of their research). Consider an environment where the firm's profits are dependent on two distinct research activities, "short term" and "long term" research. For each dimension of effort i (S=short term, L=long term), the researcher chooses an effort level, e_i, yielding output $Y(e_S,e_L)$ with Y increasing in e_S, and e_L. Assume that, in each period, the firm observes two contractible signals:[6]

$$x_S = e_{S,} + \eta_S$$

$$x_L = e_L + \eta_L, \text{ where } \boldsymbol{\eta} \sim N(O,\sigma)$$

Both short-term and long-term research can only be measured with error, and so the firm implements an incentive contract to maximize (joint) surplus taking into account the level of precision of the vector of signals. For simplicity, we assume that the signals are independent of each other.

Simply put, the firm's problem is to offer incentives according to the vector of observed signals to elicit the optimal (feasible) level of effort. By placing structure on the agent's preference

5 We owe an enormous debt to Bengt Holmstrom for the discussion that follows. Needless to say, he is not responsible for any errors that it may contain.

6 Importantly, we assume that the signal vector is observable, contractible and unbiased. In a model which allows for subjective signals or incorporates the role of reputation, the relationship between signals and optimal

function (specifically on the cost function for supplying effort), it is possible to solve for the firm's optimal incentive scheme. Following H&M, assume that the (risk-averse) agent trades off expected income against the cost of effort, that effort is costly ($c_i > 0$), and that the cost function is supermodular for effort along each dimension ($c_{ij} > 0, \forall\ i \neq j$) .[1] If the compensation contract specifies a wage of w(x), then the agent's expected utility is assumed to take the form:

$$u(CE) = E\{u[w(x) - C(e_S, e_L)]\}$$

Where

$u(w) = -e^{-r(w)}$ and CE denotes the agent's "certainty equivalent" money payoff. The coefficient r measures the agent's risk aversion. The principal is assumed to be risk neutral. We further assume that the incentive scheme imposed by the firm takes the form of a linear reward structure relating the agent's wage to the observable signals[7]:

$$w = \alpha_0 + \alpha_S x_S + \alpha_L x_L$$

where α_S and α_L are the incentive intensities implemented by the firm for short term/applied and long term/basic research, respectively and α_O is the fixed component of salary. The agent's certainty equivalent, CE, is then:

$$CE = \alpha_0 + \alpha_S x_S + \alpha_L x_L - C(e_S, e_L) - \tfrac{1}{2} r \alpha_L^2 \sigma_L^2 - \tfrac{1}{2} r \alpha_H^2 \sigma_H^2$$

Expected gross benefits for the firm, Y (e_S, e_L) are:

$$Y(e_S, e_L) = p_S e_S + p_L e_L$$

Then, following Holmstrom and Milgrom, total surplus, TS, is given by:

$$TS(\alpha_S, \alpha_L) = Y(e_S, e_L) + CE$$

Or by:

$$TS(\alpha_S, \alpha_L) = p_S e_S + p_L e_L - \alpha_0 + \alpha_S x_S + \alpha_L x_L - C(e_S, e_L) - \tfrac{1}{2} r \alpha_L^2 \sigma_L^2 - \tfrac{1}{2} r \alpha_H^2 \sigma_H^2 \quad (1)$$

TS ($\alpha_S, \alpha_L; \theta$) is supermodular if:

$$\frac{\delta^2 TS}{\partial \alpha_S, \delta \alpha_L} \geq 0, \quad \frac{\delta^2 TS}{\partial \alpha_S, \delta \theta} \geq 0, \quad \frac{\delta^2 TS}{\partial \alpha_L, \delta \theta} \geq 0, \qquad (2)$$

---

incentives will be more subtle, and empirical predictions more difficult to come by (Baker, et al, 2001).

7 Rather than following the detailed (and familiar) derivation under which linearity is in fact optimal (Holmstrom and Milgrom, 1987) , we assume linearity to focus on the relationship among incentive instruments.

Note that

$$\frac{\delta^2 TS}{\partial\alpha_S, \delta\alpha_L} = (p_S - \alpha_S)\frac{\partial^2 e_S}{\partial\alpha_S\alpha_L} + (p_L - \alpha_L)\frac{\partial^2 e_L}{\partial\alpha_S\alpha_L} - \frac{\partial e_S}{\partial\alpha_L} - \frac{\partial e_L}{\partial\alpha_S}$$

If the cost function $C(e_S,e_L)$ is quadratic, then

$$\frac{\partial^2 e_S}{\partial\alpha_S\alpha_L} = \frac{\partial^2 e_L}{\partial\alpha_S\alpha_L} = 0$$

and if effort supply function is supermodular in α_S and α_L, so that the marginal cost of effort along one dimension is increasing in the level of effort along the other dimension,

$$\frac{\partial e_S}{\partial\alpha_L} = \frac{\partial e_L}{\partial\alpha_S} > 0 \qquad \text{so } \frac{\delta^2 TS}{\partial\alpha_S, \delta\alpha_L} > 0,$$

For our purposes, however, the more interesting question is whether a shock to the firm's ability to measure effort will lead to correlated movements in incentive intensities, i.e. whether

$$\frac{\delta^2 TS}{\partial\alpha_S, \delta\theta} \geq 0, \quad \frac{\delta^2 TS}{\partial\alpha_L, \delta\theta} \geq 0$$

for the case when $\theta = -\sigma_L$. (Increasing measurement precision of the long-term research effort activity means that the error with which the signal is measured, σ_L , falls.)

Notice, first, that an exogenous shock that raises the "importance" of one of the activities (if, for example, $\theta = p_L$, p_L increases), incentive intensities will not, in general move together. To see this, note that:

$$\frac{\delta^2 TS}{\partial\alpha_S, \delta p_L} = \frac{\partial e_L}{\partial\alpha_S} < 0, \text{ while } \frac{\delta^2 TS}{\partial\alpha_L, \delta p_L} = \frac{\partial e_L}{\partial\alpha_L} > 0$$

Thus if the transition to rational drug discovery meant only that basic research became an increasingly important factor in the "production function" of research, we would not, in general, expect the intensity of incentive instruments to move together. Indeed, if effort devoted to basic research cannot be measured, then it may become even more important to rely on fixed salaries to compensate researchers, and the weight given to effort devoted to short term research might well fall.

However an exogenous shock to measurement precision, σ_L will increase the returns to the firm of increasing incentive intensity along both dimensions. To see this formally, note that:

$$\frac{\delta^2 TS}{\partial \alpha_S, \delta \sigma_L} = 0$$

while

$$\frac{\delta^2 TS}{\partial \alpha_L, \delta \sigma_L} = (p_S - \alpha_S)\frac{\partial^2 e_S}{\partial \alpha_S \sigma_L} + (p_L - \alpha_L)\frac{\partial^2 e_L}{\partial \alpha_S \sigma_L} - 2r\alpha_L \sigma_L \geq 0 \text{ if } \sigma \text{ falls, since}$$

$$\frac{\partial^2 e_S}{\partial \alpha_S \sigma_L} = \frac{\partial^2 e_L}{\partial \alpha_S \sigma_L} = 0 \text{ by first order conditions.}$$

Or, for the case that $\theta = -\sigma_L$, TS ($\alpha_S, \alpha_L, \theta$) is supermodular. These predictions have important empirical implications. Note that in a broad class of models the conditions for complete supermodularity do not hold, and we would not expect incentive instruments to be complements to each other. However in the case on which we focus here, in which the "shock" is an improvement in measurement productivity, then the system is supermodular, and we can expect α_S and α_L to be positively correlated with each other (Holmstrom and Milgrom, 1991; Athey and Stern, 1998)[8] Notice too that if returns to long term effort are increasing over the period, (p_L increasing) our test for supermodularity will be conservative.[10]

Of course, in order to argue that positive covariation between incentive elements implies complementarity requires that we address potential alternative statistical sources of positive covariation -- namely positive correlation among the factors driving the adoption of each incentive element (Arora, 1996; Athey and Stern, 1998). To do so, we note that, under the complementary hypothesis, this covariation test should be robust to conditioning on other observable factors which may be associated with the adoption process for each incentive element. As such, rather than imposing exclusion restrictions (Arora, 1996) or estimating a full structural model of adoption (Athey and Stern, 1998), our empirical approach is to evaluate the covariation test using several different "cuts" of the data, each chosen to control for the most likely alternative sources of positive correlation between the two incentive instruments.

8 More generally, the incentive intensities will be positively correlated if the stochastic shocks are statistically *associated* (a strong form of positive correlation).

10 This prediction depends only on the supermodularity of the effort supply function (i.e., different tasks are substitutes in effort), and on linearity in the incentive scheme (i.e., there is no interaction between signals in the contract).

IV. DATA SOURCES AND CONSTRUCTION

Our empirical task is thus to measure α_S and α_L and to explore the correlation between them, making plausible the hypothesis that any correlation we observe is not a function of other plausible sources of correlation.

Recall that pharmaceutical researchers are organized into small groups. A key empirical issue is thus whether to measure α_S and α_L at the group or individual level. We argue here that pharmaceutical firms implement α_S at the group level, and that they implement α_L at the individual level, largely because x_S is measured at the group level, whereas x_L can be measured at the level of the individual. Long term or basic research is usually conducted primarily by individuals, and the structures of public science allow each individual's effort level to be monitored.[11] In contrast, as we noted above, successful new drugs are typically the result of the joint effort of a research team composed of 4-7 PhD scientists.[12] Since in general the firm cannot observe the separate contribution of each member of these teams, it may optimally choose to provide a "group-level" incentive, or a "bonus" to the group's overall budget. Nonetheless, this may still provide powerful incentives for individuals: team members can then allocate this bonus among themselves, within the constraints established by the internal procedures of each firm choosing to increase wages, to hire new researchers or to purchase expensive capital equipment. Since the teams are so small, the firm can ameliorate the problem of rewarding team production by providing rewards for successful applied research at the *group* level, giving each research group discretion in how to allocate this "bonus" (Holmstrom, 1982) while at the same time remaining confident that the small size of the group will prevent any significant free riding by individual researchers who might otherwise seek to maximize the effort that they devote to basic research at the expense of the group.

Our results are obtained from a unique data set built from the detailed internal records of a sample of nine research-oriented pharmaceutical companies who, taken together, spend about 25% of the total amount of privately funded pharmaceutical research conducted worldwide[13]. Data on

---

11 Modern biology is increasingly conducted in large labs, and in these kinds of projects it may be difficult to identify the contribution of the individual. However in the period covered by our data this was not yet a major concern.

12 The mean level of funding for a single therapeutic area is $1.6m (1985 $). The detailed headcount data that we obtained from a few of the firms in our sample suggest that this is roughly sufficient to employ 4-7 PhD level researchers, when overhead and support costs are factored in.

13 The data are provided under guarantees of strict confidentiality and anonymity so we can discuss the makeup of the sample only in broad terms. The sample is relatively representative of the industry as whole, in terms

individual research programs expenditures are supplemented by patent data and a measure of the degree to which the firm provides incentives for basic research in its promotion policies. (Cockburn and Henderson (1994) and Henderson and Cockburn (1994; 1996) discuss the construction of this data set in greater detail.

Measuring the incentive to do basic or long term research To measure the intensity with which firms provide incentives to engage in "fundamental" research (i.e., α_L), we use a variable, PROPUB, that is derived from over a hundred interviews with senior managers and scientists at our sample of pharmaceutical firms. In order to minimize the problem of retrospective bias, the interviews designed to construct a comprehensive history of the development of cardiovascular drugs at each firm.[14] Each respondent was questioned in detail about the ways in which research was organized over time, but the questions were linked to specific events in the history of the firm (e.g., who worked on the development of this beta-blocker? what happened? were they rewarded? why or why not?). PROPUB was then constructed by assigning each firm in each year a value on a 5-point Likert scale based on the degree to which the firm's promotion policies were based on a researcher's standing in the external scientific community, where a value of 1 indicates that the firm placed no value at all on a researcher's reputation in the external community in rewarding his or her efforts and a value of 5 indicates that it was a central criteria in such decisions.

PROPUB has been found to discriminate effectively among firms in terms of their R&D productivity and is also correlated with several alternative measures of a firm's commitment to the world of public science and of its rate and extent of scientific publication activity (Henderson and Cockburn, 1996; Cockburn, Henderson, and Stern, 2000). However since the use of a subjectively constructed Likert scale will always raise questions, we also employ an alternative measure ("HIGH" PROPUB DUMMY) which is equal to 1 after a firm has increased its PROPUB level and is zero otherwise. While this measure exploits less of our qualitative information than PROPUB, it provides a more unambiguous index of the changing incentives for basic research within each firm

---

of size, technical or commercial performance, and geographic distribution (with firms headquartered in both the United States and Europe).

14 The data are provided under guarantees of strict confidentiality and anonymity so we can discuss the makeup of the sample only in broad terms. The sample is relatively representative of the industry as whole, in terms of size, technical or commercial performance, and geographic distribution (with firms headquartered in both the United States and Europe).

in the sample.[15] Our results are robust to the use of either measure.

Across firms, differences in PROPUB reflect significant differences in the promotion policies of the firm (ranging from strong restrictions on scientific publishing and the active discouragement of basic research initiatives to the use of a promotion system not dissimilar to that of a university biology department – promotion based on publication record and external recommendation letters from leading scientific researchers in the public sector). Within a firm, "switches" in the PROPUB regime reflect a significant change in the firm's use of promotion incentives to encourage basic research. Over the sixteen year and nine firm sample, there were five "switches" from a lower to a higher regime are observed, so that there are fourteen distinct "firm / basic research incentive level" regimes.[16]

Applied Research Incentives

To assess the internal incentives provided to supply effort towards applied research, we look to the firms' internal capital market, and to research funding decisions. Internal capital markets can play an important role as a reward mechanism for workers, ameliorating agency problems within the firm (Hart, 1995; Stein, 1997). In the context of pharmaceutical firms, we observe drug discovery teams in different therapeutic areas competing with one another for resources, with variation in project funding decisions interpretable as a highly visible reward for success. (A "therapeutic area" is a sub-market within the pharmaceutical industry. For example depression, anxiety and hypertension are all separate therapeutic areas.) By varying a research team's budget in response to observed output, a firm provides incentives for the team's workers to supply effort to generate positive signals.

We measure the intensity of incentives to supply effort in applied research by estimating the sensitivity of drug discovery team research budgets to observed success in producing "applied"

15 In other work, we explore several alternatives, such as PUBFRAC (the percentage of patent authors who also publish in the referred literature). Though less subjective, these quantitative measures suffer from two limitations. First, they measure *outcomes* rather than incentive policies, and, second, they cannot be constructed for the full period covered by our detailed R&D investment data.

16 Since adopting a higher level of PROPUB may take time, for firms in which we observe a switch from a lower to a higher level of PROPUB, we allow for a "transition" period during the first year of implementation by excluding these "switching" periods from our sample. All of the results presented in Section V are robust to various different treatments of the adjustment process, such as creating a one-year "band" around the switching dates (including the year before and after) and ignoring the adjustment process altogether.

output in the form of potentially marketable compounds, where we measure this output in terms of the number of "important" patents applied for in a given year. To ensure comparability across firms, we restrict ourselves to a measure of "important" patent counts, that is inventions for which patent applications were filed in at least two of three major jurisdictions (the U.S., Europe, and Japan.) This controls for variation across firms in their propensity to patent "marginal" discoveries or in their national environment (patent counts based on single country grants will tend to be biased towards domestic firms).[1] We define a patent as important if it was subsequently granted in two of the three major patent jurisdictions (the USA, Europe and Japan). Important patents provide a particularly useful measure of applied output in this setting since the pharmaceutical industries is one of the few industries in which patents both correspond to particular products (a drug is a single patentable molecule) and in which they are central to competitive advantage (Levin, et al, 1987; Cohen, et al, 2000).[17] We assume that the timing of the firm's patent filings is a good measure of the time at which decision-makers acquire objective information about a research group's recent production of potentially commercializable compounds. Finally, we match these patents to underlying research expenditures using a classification scheme based on standard therapeutic class codes (such as the IMS Worldwide Therapeutic Classification Scheme) modified to reflect the organizational structure of the firms in the sample.[18] All patents are counted by earliest world-wide priority date of the invention.

We estimate this sensitivity by constructing a simple model of R&D investment at the research program level, which allows the team's research budget (and thus observed expenditures) to be driven both by the need to provide incentives and by technological opportunity. The key assumption of the model is that changes in the research budget for a given drug discovery team from

---

[17] Derwent's *World Patent Index* compiles comprehensive data on international patent filings, allowing us to identify those granted in multiple jurisdictions. Application costs rise roughly proportionately with the number of jurisdictions, and firms rarely file in all possible jurisdictions, let alone all major markets (e.g. all OECD countries.) By excluding inventions where the firm does not file in at least two out of three major jurisdictions, we are therefore left with a count of "important" patents. Derwent's database goes back to 1962, though much less comprehensive data is available before 1970.

Advances in molecular biology have spawned a number of developments that are "basic" in the sense of being fundamental to advances in science, but that have nonetheless proved to be patentable. However during the time covered by our sample, only a small share of total research expenditures were devoted to such areas. For example, although the average firm in our sample produced several hundred US patents per year, Kaplan, Murray and Henderson (2001) estimate the average pharmaceutical firm produced less than five biotechnology patents through 1990.

[18] Where we were not confident about this matching, research programs and patents are assigned to a

year *t-1* to year *t* reflect *both* a "bonus" payment reflecting the team's "excess" productivity over and above the expected level of applied research output in year *t* $\tilde{I}_{i,j,t} = \alpha_S x_{s,i,j,t-1}$ where $x_{s,i,j,t-1}$ is the "shock" to applied research productivity by a research team in year t-1, *and* changes that reflect "efficient" investment insofar as the firm adjusts its research expenditures according to "news" from period *t-1* about underlying technological and market opportunities (Pakes, 1981; Abel, 1984),

$$I^*_{i,j,t} = I^*_{i,j,t-1} + \beta^x X_{i,j,t-1} + \beta^z Z_{i,j,t-1} \qquad (3)$$

where $X_{i,j,t-1}$ is the shock to technological opportunity realized by program *i* in firm *j* in period *t-1*, $Z_{i,j,t-1}$ are opportunity shocks external to this program but observed by the firm, and $I^*_{i,j,t-1}$ is the optimal level of expenditure in the prior period.

In addition, we assume that the firm's internal measure of technological opportunity cannot be distinguished from the signal it receives about the team's applied research output shock (i.e., $X_{i,j,t-1} = x_{S,i,j,t-1}$ and that $Z_{i,j,t-1}$ can be partitioned into "news" observable to both the firm and econometrician ($z_{i,j,t-1}$) and a shock observable to the firm but not to the econometrician ($\zeta_{i,j,t-1}$). Subtracting $I^*_{i,j,t-1}$ from both sides and modeling total investment $I_{i,j,t} = I^*_{i,j,t} + \tilde{I}_{i,j,t}$ yields an expression for the overall change in expenditure after accounting for both the group-based incentive payment and the firm's response to technological opportunity:

$$\Delta I_{i,j,t} = (\alpha_s + \beta^x) x_{s,i,j,t-1} + \beta^z z_{i,j,t-1} + \zeta_{i,j,t-1} \qquad (4)$$

Under this model, data on research program investment and applied research output can be used to estimate $\gamma^S = \alpha_S + \beta^x$, the sensitivity of research program budgets to the prior period's unanticipated applied research output.[19] Notice that since any increases stemming from individual promotions go into the level of the group's budget and we are here focusing on changes, we can separately identify the two instruments.

To estimate (4), we must derive a measure for $x_{s,i,j,t-1}$, the observed shock to applied

---

"Misc/NEC" class and not used in the analysis.

[19] Overall, the sign and magnitude of γ^S are ambiguous theoretically since, while we expect α_S, the firm's optimal investment response to applied research output shocks depends on whether there are increasing or diminishing returns to effort in a particular therapeutic area. In applying this model to data , we must ensure that the estimate of γ^S controls for unobserved factors correlated with applied research output "shocks" and increases in R&D funding problem, which we largely address through the use of a differences-in-differences estimator with firm-program fixed effects. We describe the empirical strategy in detail in Section V.

research output for a given therapeutic program in a given year. The details of our derivation are provided in Appendix A, but essentially, we first calculate the expected level of patents (our measure of applied research output) for each team for each year by regressing the level of patents as a function of the historical patent production rate of the team and $I_{i,j,t-1}$ .We then use the fitted values from this regression as our measure of the "predicted" level of patents for that team for that year. Finally, we define PATENT SHOCK as the difference between the observed and predicted level of patenting for that research program and SHOCKi,j,t as:

$$x_{S,i,j,t-1} = SHOCK_{i,j,t} = \frac{PATS_{i,j,t-1} - E[PATS_{i,j,t-1}]}{E[PATS_{i,j,t-1}]} * I_{i,j,t-2} \qquad (5)$$

which is PATENT SHOCK adjusted for the scale of the research program. Since this measure is observed by the firm when choosing the research team's budget for year t, the firm is able to implement the investment equation (4).

Funding Variables. Our data on research investment are taken from a database on research expenditures for several hundred individual research programs conducted by firms in this sample between 1975-1990. These data were assembled from confidential internal records, and great care was taken to treat data consistently across firms and over time. Pharmaceutical research takes place in two distinct phases: pre-clinical (or "discovery" research) and clinical (i.e., development); here we focus exclusively on the former.[20] RESEARCH is thus the level of expenditures on pre-clinical discovery research in a given firm-program-year, deflated to 1986 dollars by the NIH biomedical research deflator. We measure the "bonus" to the research budget, ΔRESEARCH, as the first difference of RESEARCH. Similarly, FIRM RESEARCH is just the sum of RESEARCH over all observed programs of a firm in a given year.

Sample selection. With a complete, balanced data set (all firms participating in all programs in all years from 1975-1990), the data set would consist of 7040 firm-program-year observations. The data set is unbalanced, however, affecting the size of the sample. First, and most importantly, firms initiate and discontinue research programs throughout the sample. We only include

20 By focusing exclusively on the discovery phase of pharmaceutical research, we avoid the complexities of modeling the multi-year multi-stage development phase whereby individual drugs are moved through clinical development and testing for regulatory approval. Also note that external research grants and licensing or joint-venture payments are sometimes included in the data (as appropriate); however, these types of funding arrangement represent a very small share of the total during the period of our sample

observations for which a research program is "active" in the sense that the firm actively engaged in at least some research in a particular therapeutic area (resulting in the loss of 2319 potential observations). As well, some firms are involved in mergers and some firms' discovery spending is not observed continuously between 1975-1990 (resulting in a net loss of 978 observations). Further, 1164 observations are removed because both ΔRESEARCH and PATENT SHOCK*$RESEARCH_{t-1}$ are equal to 0. Finally, since we are interested in whether firms who have a given level of PROPUB tend to be more responsive to applied research outputs shocks in their capital budgeting, we allow a one-year "adjustment" for those firms who switch PROPUB during the sample period, resulting in the loss of 139 observations. Taken together, these sampling rules result in a final data set of 2417 observations which we use throughout our empirical analysis.

Table 1 provides variable definitions, and Table 2 reports the summary statistics). On average, each firm in the sample has just above 10 distinct drug discovery teams spending $1.58 million per year (in constant 1986 dollars) and obtaining 3.30 important patents per year (Table 2). On average, program receive a modest "boost" over time; however, ΔRESEARCH varies widely across programs and over time. Although some programs produce more than 15 patents per year, no patents are produced in 30 percent of program-years, and, for 76% of the annual observations, fewer than five patents are produced.

While the promotion policy variable PROPUB is centered around the mean of the 5-point Likert scale, there exists substantial variation along these dimensions both across firms and across time (ANOVA reveals that the variance is evenly divided across the within-firm and between-firm dimensions). The measure capturing the presence of a "switch" in PROPUB ("HIGH" PROPUB DUMMY) captures more than one quarter of the sample.

V. CORRELATION OF BASIC AND APPLIED RESEARCH INCENTIVES

Recall that our key equation is equation (4)

$$\Delta I_{i,j,t} = (\alpha_s + \beta^x) x_{s,i,j,t-1} + \beta^z z_{i,j,t-1} + \zeta_{i,j,t-1} \qquad (4)$$

where $\gamma^S = \alpha_S + \beta^x$ is the firm's budgetary response to unanticipated changes in research output, and our key empirical question is that of the degree to which α_S is correlated α_L.

Table 3 presents an estimate of γ^S obtained by estimating (4) using data on annual research

expenditures and patents at the level of the individual research program. We regress the first difference of research expenditures $\Delta I_{i,j,t}$ or $\Delta RESEARCH_{i,j,t}$ against (a) our measure of the applied research output "shock" (i.e., $x_{s,i,j,t-1}$ represented as SHOCK), (b) our measures of external technological opportunity shocks, (i.e., $z_{i,j,t-1}$) and (c) controls for an overall time trend, the scale of the program ($I_{i,j,t-1}$, or $RESEARCH_{i,j,t-1}$) and "momentum" in the research funding process ($\Delta I_{i,jt-1,}$ or $\Delta RESEARCH_{i,j,t-1}$). Since $z_{i,j,t-1}$, is difficult to measure directly, we use "news" in the patent applications of related research programs both inside the firm and at a sample of competitor firms to proxy for changes in technological opportunity.[21] I_{t-1} is included as a control for size and to capture any higher-order time series properties.

The very high variance of the dependent variable (and the starkness of our investment model) is reflected by the low R^2 for the regression,[22] but our main variable of interest, the "shock" to observed applied research output has a positive coefficient, as expected, and is strongly significant. The magnitude of this coefficient is sensible: it implies that a one-standard-positive-deviation "surprise" in SHOCK has about a $140,000 (or approximately 9%) impact on the budget of the average program. Finally, there is a great deal of variation in γ^S both across firms and across basic research incentive "regimes" within a firm (recall that our measure of basic research incentives is a categorical variable with specific "switch" dates for individual firms): we can conclusively reject homogeneity of γ^S along each of these dimensions (these results are available from the authors upon request). This result holds with or without the other covariates in the model, and whether or not their coefficients are allowed to be regime-specific, or are constrained to be equal across sub samples.

These results suggest that firms do indeed respond heterogeneously to unexpected shocks. Recall, however, that $\gamma^S = \alpha_S + \beta^x$. The firm's budgetary response to unexpected shocks reflects both its rewards to effort and its responsiveness to technological opportunity. The remainder of the paper is devoted to evaluating whether this variation in γ^S can be tied to the provision of basic research incentives (i.e., to the level of PROPUB). In other words, is α_S correlated with α_L, or with

21 The control measures for competitors' patents are drawn from a broader cross-section of 29 leading worldwide pharmaceutical firms.

22 Another way to think of this is to recognize that research program budgets are highly auto correlated. While firms do adjust these expenditures, either through marginal changes to program budgets or by opening or closing programs, year-on-year the average changes are quite small. (See Cockburn and Henderson, 1994).

PROPUB?

We begin with two very simple reduced-form summary analyses. The goal of these preliminary exercises is to explore whether a correlation between basic and applied research incentives can be found even using relatively crude measures and quite aggregate data, and so to motivate the more nuanced panel data analysis that follows.

In Table 4, we compute the average change in research funding for individual firm-program-years depending on whether the firm-program receives a positive or negatively signed applied research output shock (SHOCK > or < 0) and on whether the firm is associated with a "HIGH" or "LOW" PROPUB regime. The differences are dramatic: in low PROPUB regimes, a positive SHOCK is associated with a budget "boost" of $180,000 relative to a negative SHOCK. In contrast, in high PROPUB regimes, the budget boost almost doubles, to over $350,000 (the conditional means in all four boxes are statistically significant from each other at the 1% level). In other words, drug discovery programs operating in a high PROPUB regime are associated with a much higher sensitivity to patent output shocks.

A second method for evaluating the overall presence of a correlation between γ^S and PROPUB involves a simple two-step procedure. In the first stage, we estimate the budget's sensitivity to SHOCK for each of our firm-basic research "regime" combinations (recall that there are a total of 14 "regimes" across the sample); or in other words, following (9), we estimate 14 individual γ^S estimates, one for each firm across the span of time over which that firm maintains a constant level of basic research incentives. We then evaluate the correlation between this regime-specific estimate of the sensitivity to research outputs shocks and PROPUB.[23] Though there are only 14 distinct regimes, the results are encouraging: the Pearson correlation coefficient between PROPUB and $\hat{\gamma}^S$ is 0.499 (significant at 5%). In Table 5, we present two simple regressions of $\hat{\gamma}^S$ on PROPUB; even after controlling for a time trend, PROPUB has a positive coefficient (significant at the 10% level).[1]

These results are highly suggestive, and are certainly consistent with our core hypothesis:

---

23 Note that if β^x were a constant, so that variation in $\hat{\gamma}^S$ reflected variation in α^S, this would provide a clean test of the H&M hypothesis.

[24] Of course, because both of these variables are measured with substantial error (a problem we address below), the estimated coefficient is likely downward-biased.

under high PROPUB regimes, firms offer more aggressive incentives for the generation of applied output. However, our analysis so far has not accounted for the potential impact of unobserved heterogeneity on the correlation between PROPUB and γ^S. If the levels of these variables are jointly determined by an unobserved factor (or if the factors determining each variable are correlated with each other), then the correlation among incentive intensities may be due to unobserved heterogeneity rather than to complementarity. At the same time, if these unobserved factors are independent of each other, this will introduce "noise" into the observed correlation of incentive intensities, and so weaken the power of a correlation test for inferring complementarity among incentive instruments. For example, suppose that the intensity of incentives for applied research is determined by factors unrelated to the intensity of incentives for basic research (e.g., because of corporate culture or liquidity constraints), then the observed correlation between them will provide a downward-biased estimate of the importance of complementarities in the provision of incentives.

To address these concerns, we exploit the panel structure of our data to estimate the conditional correlation between basic and applied research incentives under several alternative assumptions about the nature of unobserved heterogeneity within our sample. Specifically, the remainder of the analysis is conducted at a more disaggregated level – taking the "firm-program-year" as the unit of observation. This allows us to take advantage of the full richness of our research program data and to introduce controls for both potential changes in β^X and for possible alternative drivers of correlation between basic and applied research incentives.

To understand this empirical strategy more precisely, recall that we defined γ^S to be the total response of the research budget (of research program i in firm j in year t) to the "surprise" in applied research output: $\gamma^S = \alpha_S + \beta^x$. We test for correlation between $\alpha_{S,i,j,t}$ and $\alpha_{L,i,j,t}$ by letting $\alpha_{S,i,j,t}$ be a function of $\alpha_{L,i,j,t}$ (i.e. $\alpha_{S,i,j,t} = \rho^0 + \rho^{S,L}\,\alpha_{L,i,j,t}$) yielding:

$$\gamma^S{}_{i,j,t} = \beta^x{}_{i,j,t} + \rho^o + \rho^{S,L}\alpha_{S,i,j,t}$$

Substituting back into (4) results in an empirical model to test for the presence of correlation using firm-program-year data:

$$\Delta I_{i,j,t} = (\beta^x{}_{i,j,t} + \rho^o + \rho^{S,L}\alpha_{L,j,t})x_{s,i,j,t-1} + \beta^z z_{i,j,t-1} + \zeta_{i,j,t-1} \quad (11)$$

where the test for complementarity is simply $\rho^{S,L}$. As discussed above, the key challenge in estimating this parameter (and therefore performing a consistent test for complementarity) is

accounting for the impact of variation in β^X .

To begin, we assume that β^X is unobservable and is uncorrelated with $\alpha_{L,i,j,t}$. Under this assumption, we can implement (11) by using PROPUB as a measure of $\alpha_{L,i,j,t}$ and regressing ΔRESEARCH on SHOCK and interactions of SHOCK with PROPUB. As in Table 3, we also include a time trend and controls for technological opportunity and other drivers of ΔRESEARCH. Table 6 reports these results. In model (6-1) we reconfirm our results from Table 3, with a regression showing a significant relationship between ΔRESEARCH and SHOCK. Model (6-2) provides our first detailed evidence that the overall sensitivity to SHOCK is positively associated with the level of PROPUB. Not only does the inclusion of a PROPUB interaction decrease the quantitative and statistical importance of SHOCK, but the coefficient suggests that the impact of PROPUB is quite large. Whereas a one-standard deviation in SHOCK is associated with less than a 5% increase in investment when PROPUB is at its lowest level, this same shock is associated with over a 19% increase when PROPUB is at its highest level. In (6-3), we include several controls associated with $z_{i,j,t-1}$ – two measures of information about technological opportunity ("NEWS" in COMPETITOR PATENTS and "NEWS" in RELATED PATENTS) as well as measures to account for the scale of the research program ($RESEARCH_{t-1}$) and potential serial correlation in the dependent variable ($\Delta RESEARCH_{t-1}$). Though these additional regressors enter significantly, their inclusion does not change our key result: the coefficient on SHOCK*PROPUB remains positive, of a similar magnitude, and with a similar standard error. In (6-4), we replace the time trend with year fixed effects, each interacted with SHOCK. These are jointly significant and result in a modest increase in the estimated parameter on SHOCK*PROPUB.[25] These results are consistent with the findings from Tables 4 and 5, and provide evidence consistent with the presence of complementarity between basic and applied research incentives.

However this interpretation is conditional on our assumption that variation in β^X is

25 We also have explored several robustness checks on these relatively "sparse" specifications (available from the authors). These include: using therapeutic class-specific fixed effects, incorporating several controls for changes in the firm's management structure (such as changes in the CEO, R&D Vice President, or changes in the process used in the capital budgeting process, and introducing additional lags of the dependent variable into the specification. As well, as discussed in Appendix B, we have explored specifications calculating SHOCK with alternative models of the firm's expectations process and using the "levels" version of SHOCK rather the percentage version used in Table 6. While several of these additional results contribute modestly to the regression's explanatory power, none is associated with a substantial change in either the magnitude or statistical significance of the SHOCK*PROPUB

uncorrelated with PROPUB. We therefore turn to more detailed analyses which control for the likely sources of correlation between β^X and PROPUB across programs, firms, and time. Three alternatives stand out. First, as discussed above, spurious correlation would be introduced if the use of both incentive instruments simply increased over time. Between the late 1970s and early 1990s, the use of promotion-based basic research incentives diffused widely throughout the pharmaceutical industry. The results in Table 6 indicate that overall changes over time are statistically significant, whether captured by a time trend or by year fixed effects. Although these variables have little or no impact on the coefficient of interest, we continue to include them in subsequent regressions in order to control for any omitted trends over time industry-wide variables. Second, there may be heterogeneity across therapeutic classes. It is possible, for example, that firms with higher levels of PROPUB are concentrated in therapeutic areas which tended to increase their sensitivity to SHOCK at a faster rate than the average. For example, the benefits from providing incentives for basic research seems to have increased especially rapidly in hypertension (Henderson, 1994; Cockburn, Henderson, and Stern, 2000). To the extent that patents (or a "surprise" in patenting) in these therapeutic areas also became more informative about applied research effort and technological opportunity, the correlation between PROPUB and the level of applied research incentives will reflect heterogeneity among firms in terms of their participation in different therapeutic areas. Third, it is possible that high PROPUB regimes are associated with firms and research programs which have "intrinsically" higher sensitivity to patents in the research budgeting process. For example, perhaps firms with higher levels of PROPUB have more "active" R&D managers who also tend to be more sensitive to applied research output in capital budgeting, or who simply have a taste for high powered incentives. In such an environment, exploiting the cross-sectional variation in the data will confound evidence of complementarity with evidence of a "taste" for incentives.

We address each of these concerns by including controls that directly account for each factor.[26] Specifically, we interact $x_{S,i,j,t-1}$ with firm-program fixed effects, a time trend for each

---

coefficient

26 This approach can be contrasted with more "structural" solutions, such as imposing cross-equation restrictions regarding adoption drivers (Arora, 1996) or the estimation of a simultaneous equations model integrating the adoption and performance implications of complementarity (Athey and Stern, 1998)

research program, yielding a richer specification:

$$\Delta I_{i,j,t} = (\gamma^{0_{i,j}} + \gamma^{t}{}_{i}(t - t_0) + \rho^{S,L}\alpha_{S,j,t})x_{s,i,j,t-1} + \beta_{i,j} + \beta^{z} z_{i,j,t-1} + \zeta_{i,j,t-1} \qquad (12)$$

Interacting $x_{S,i,j,t-1}$ with a firm-program fixed effect controls for any cross-sectional variation in the "intrinsic" sensitivity of different research programs to applied research output. For example, if patents in a particular hypertension program are inherently more informative than patents in a particular depression program, (12) will control for these effects. As well, these fixed effects will control for the potential variation among managers in their "taste" for providing high-powered incentives. Controlling for changes over time, including year-specific and therapeutic class/year-specific dummies and interactions with $x_{S,i,j,t-1}$, nets out both an overall and class-specific trend in unobserved components of β^{X}. In other words, in (12), $\rho^{S,L}$ is the correlation between changes in the sensitivity to $x_{S,i,j,t-1}$ and changes in the level of PROPUB relative to the trend. This estimator is essentially a differences-in-differences estimator. However, in contrast to the classic differences-in-differences estimator, the hypothesis tested here concerns an interaction effect and so we require each of the individual effects to be interacted with $x_{S,i,j,t-1}$

Table 7 reports the results. In all of these regressions, we include a complete set of firm-program fixed effects and interactions of these with SHOCK. These interaction effects are jointly significant and substantially increase the explanatory power of the regression, indicating a high degree of heterogeneity among firm-programs in their average investment response to applied research output. To be consistent with a differences-in-differences estimator, these specifications rely exclusively on within-program variation in PROPUB and the presence of "switches" in the incentives provided for basic research. In this table, rather than using the Likert scale variable PROPUB, we use the "HIGH" PROPUB dummy, which is equal to one only for those years after the firm has "switched" from a lower level of PROPUB to a higher level of PROPUB, and is otherwise set equal to zero.[27] This measure is equal to one for a little more than one quarter of the sample, suggesting that it may be possible to identify our test exclusively on the "within" dimension.

27 All of the results in Table 7 are robust to using the five-point PROPUB variable instead of this differenced version.

This more stringent test provides further support for the "balance" hypothesis. After accounting for individual firm/program-specific interactions and program-specific time trend interactions with SHOCK, the magnitude of $\rho^{S,L}$ (our key parameter) increases substantially and remains at a similar level of statistical significance ($p < 0.01$ for all specifications). According to the "richest" specification, (7-4), for an average-sized program which realizes a one standard deviation SHOCK, there is a $390,000 incremental "boost" in the research budget after the firm switches to a higher level of PROPUB. This amount is more than 25% of the size of the average research program. In other words, after accounting for several sources of potential spurious correlation, the estimated relationship between basic and applied research incentives is stronger than in pooled data analysis conducted in Table 6.

Indeed, our evidence in favor of the complementarity hypothesis is somewhat strengthened when we consider the results from Tables 6 and 7 in concert. Recall that the key concern about the pooled analysis was the possible presence of a positive correlation between $\beta^{X}_{i,j,t}$ and PROPUB. However, in Table 7, after controlling for several sources of heterogeneity, we find that the magnitude on our key parameter increases and that a substantial share of the overall variation in ΔRESEARCH is associated with firm program-specific fixed effects. As such, the evidence from Table 7 is consistent with the hypothesis that the coefficient on PROPUB*SHOCK in Table 6 is, if anything, *underestimated*. If there was unobserved heterogeneity which was strongly and positively correlated with PROPUB, then either the "within" estimate of the coefficient in Table 7 would be much smaller in magnitude, or the fixed effects would have to account for a only a small fraction of the total variance.

For our result to be biased by any omitted independent variables driving incentive intensities, these would have to have a significant explanatory power above and beyond firm-program fixed effects and therapeutic class-specific trends. Meeting this challenge weakens the appeal of alternative interpretations of this result, since they must hold true both in the "pooled" or "between" dimensions of the data, and in the "within" dimension. Suppose, for example, that PROPUB and $\hat{\gamma}^{S}$ were driven by a common general organizational response to science-driven drug discovery, with changes in PROPUB reflecting the outcome of "doing science" in terms of actual tasks performed by workers, and the nature of human capital employed by the firm, and changes in the sensitivity of

research budgets to patent signals reflecting higher quality inventions, or a "science driven" capital budgeting process. This would certainly result in PROPUB and $\hat{\gamma}^S$ being correlated in the cross-section. But for the same to be true in the "within" dimension of the data there would have to be both (a) enough heterogeneity at the research program level in these effects that their "true" variation would not be accounted for by the fixed effects and therapeutic class-specific trends, and (b) sufficient co-movements over time in the "true" residual impact of adopting science-driven drug discovery on PROPUB and $\hat{\gamma}^S$ (as opposed to just noise) to generate a strongly positive association in the data. Absent effective instruments for the adoption of science driven drug discovery as distinct from pro-publication incentives we cannot definitively reject this hypothesis, but nonetheless we believe it to be unlikely. By and large, while firms adopted uniform incentive policies, the rate at which they adopted the particular techniques of science-driven drug discovery varied significantly across programs and thus we think it is very unlikely that the second condition holds in these data.

VI. CONCLUDING THOUGHTS

The principal finding of this paper is the presence of a positive correlation between measures of the use of promotion-based incentives for basic research and of team-based incentives for applied research. This correlation is both economically and statistically significant in a variety of different "cuts" of a panel dataset on R&D investment behavior of pharmaceutical companies. As in Ichniowski, Shaw and Prennushi (1997), our empirical strategy has been to exploit the full range of variation contained within a micro-level dataset to rule out a variety of potential sources of unobserved heterogeneity. The positive correlation between basic and applied research incentives exists whether we aggregate the data into a small number of distinct firm-regimes, exploit cross-sectional variation among individual research programs, or subject the hypothesis to a differences-in-differences test using only within-program variation over time.

This result is consistent with a key proposition of modern agency theory – that when a principal prefers agents to balance their effort across multiple tasks, if monitoring technology changes such that it is now possible to monitor effort devoted to two complementary tasks, incentive intensity will be "balanced" and increases in incentive intensity on one dimension will be associated with increases in incentive intensity on competing dimensions. Our interpretation of our results as

providing novel empirical support for this "complementarities" proposition is, however, tempered by our inability thus far to obtain data which would allow us to directly identify incentive intensity choices, and to rule out other potential explanations.

An interesting aspect of our investigation is the degree to which the types of incentives discussed in the abstract in contract theory are embedded in the design of the firm's internal organizational processes. We do not discount the efficacy of unidimensional monetary incentive schemes in environments where output is easily monitored and there is opportunity for specialization. But, to understand incentives in a complex environment such as an R&D laboratory, our results suggest that it is critical to account both for the possibility that incentives may be multidimensional, and for the firm's ability to provide these incentives through mechanisms such as the operation of its internal labor and capital markets. Aligning agency theory with the use of incentives in real organizations is likely to require quite careful tailoring of the empirical content of contract theory to concrete organizational and institutional settings.

APPENDIX A: ESTIMATING THE APPLIED RESEARCH OUTPUT SHOCK

"Surprises" in patenting play a key role in this paper as signals of effort supplied by research workers in applied research. This section discusses a variety of possible methods for constructing this "surprise" variable. Our measure of applied research output is "important patents" attributed to the research program. Pharmaceutical firms file patent applications on discoveries which show commercial promise promptly, and we believe that at least in this context they are a good measure of successful outcomes in applied research projects.[1] Using time series on each research group's patenting (PATENTS), one could construct a simple measure of the applied research output "shock" as the difference between the research group's observed and expected patenting rate,

$$PATENT\ SHOCK_{i,j,t} = PATENTS_{i,j,t} - \mu_{i,j,t}^{PATS} \qquad \text{(A1)}$$

where μ^{PATS} is an estimate of expected patent output from program *i* in firm *j* in year *t*. Two issues arise in adapting this formula to the investment sensitivity equation estimated in Section IV. First, this measure must be made comparable across programs, and so we need to take account of systematic technological differences across programs in the number of patents generated by a given amount of research spending. These differences may be large: a million dollars spent on screening for antibiotics may generate as much as five times as many patentable candidate compounds as a similar level of resources devoted towards cancer research. As well, programs vary widely in terms of their absolute size, and we need to adjust our shock measure so that the "budget sensitivity" parameter implies a proportional impact on the budget for programs of different size. Consequently, though the results do not depend on whether we control for these two "proportionality" problems, we address these issues in the empirical work by expressing the applied research output shock as:

$$SHOCK_{i,j,t-1} = \frac{PATENTS_{i,j,t-1} - \mu_{i,j,t-1}^{PATS}}{\mu_{i,j,t-1}^{PATS}} * I_{i,j,t-2} \qquad \text{(A2)}$$

To use (B2), we require a consistent estimate of μ^{PATS}, the expected level of patenting for a given firm-program-year. Obviously, the econometrician cannot construct an exact measure of this expectation; each firm has access to richer information about its own programs than outside

[28] We recognize that patents may be filed on discoveries which are quite far from commercial application: in this context, putting a candidate compound into clinical trials. There is also the possibility that strategic considerations may lead firms to delay filing applications, or to pursue large numbers of otherwise insignificant applications in an effort to construct a protective "thicket" around a core discovery. However prior work with these data, as well as interviews with firm personnel lead us to believe that these problems are unlikely to be a serious source of systematic bias. Note also that we count only "important" patents filed in two out of three major jurisdictions worldwide, and we date applications by their worldwide priority date.

observers. However, it is feasible to attempt to estimate the firm's expectation of patent output by making assumptions about what the firm pays attention to in this process. Here we discuss three alternatives, which attempt to "span the space" of reasonable models. Results presented in the body of the paper use only the third (most sophisticated) of these expectations models.

In the first, most naive, of these models, we assume that the firm's expectation is simply the level of patenting in the immediate prior year (i.e., annual patent counts follow a first-order Markov process):

$$\mu^{M}_{i,j,t} = PATENTS_{i,j,t-1} \qquad \text{(A3)}$$

This measure has obvious shortcomings, since it assumes that decision makers have only a very limited "memory" and are basing decisions on an extraordinarily limited information set. To construct our second measure, we model the firm's expectations about each research group's performance as being based on the assumption that patent counts follow a Poisson process whose rate parameter, $\mu^{POISSON}$ can be estimated from past realizations of patent output. Specifically, $\mu^{POISSON}$ is just the mean number of patents per period over all observed periods to date,

$$\mu^{P}_{i,j,t} = \frac{1}{t - T_0} \sum_{s=T_0}^{t-1} PATENTS_{i,j,s} \qquad \text{(A4)}$$

While data are only available for the period 1975-1990 for some of our other variables, we have much longer time series on the patent output of each research program, and $\mu^{POISSON}$ is constructed from as many as 30 years of data on patent counts. However while it incorporates the historical trend in patenting in the research program, $\mu^{POISSON}$ takes no account for the level of funding provided to each research program. It is reasonable to suppose that managers base their expectations about the level of patent output on both the history of patenting and the amount of resources currently available. To allow for this, we compute a third measure of expected patents, $\mu^{ADAPTIVE}$ based on the notion that manager's expectations are updated adaptively in response to both of these factors. $\mu^{ADAPTIVE}$ is constructed using a two-stage procedure based on a regression-based weighting of the $\mu^{POISSON}$ measure and $RESEARCH_{t-1}$, the level of funds provided to the research program in the previous period. To do this, we first compute $\mu^{POISSON}$, and then in the second stage run a Poisson regression of observed patenting on the level of research by each research group and $\mu^{POISSON}$ (this amounts to estimating the regression with a distributed lag on the dependent variable). $\mu^{ADAPTIVE}$ is thus the fitted value of the level of patenting resulting from this Poisson regression,

$$\mu^{A}_{t} = \exp(\hat{\lambda}_0 + \hat{\lambda}_P \mu^{P}_{t-1} + \hat{\lambda}_R RESEARCH_{t-1}) \quad \text{(A5)}.$$

Of course, it would be possible to extend the logic of $\mu^{ADAPTIVE}$ to take the fitted value from any model of the drivers of patenting productivity. Though $\mu^{ADAPTIVE}$ is a marked improvement over just using past realizations of the patenting process, the data we have available is only a small subset of the information available to the decision-maker in reality. However, we are reluctant to impose an overly sophisticated model (e.g., a vector autoregression model which minimizes ex-post forecasting

error) for two reasons. In the first place, it is simply counterfactual. Research managers can and do use sophisticated quantitative tools (Nichols, 1994), but a considerable body of research has demonstrated that practicing managers rely extensively on heuristics and rules of thumb. In the second place, an overly specified model is actually unhelpful in this context: a fully saturated statistical model will result in "shocks" which contain less and less "signal" about unanticipated performance and more and more true random noise! Consequently, $\mu^{ADAPTIVE}$ is our preferred measure of expectations, since we believe that it incorporates a realistic amount of information. While we doubt that research managers would update expectations without taking into account the amount of funds that had been invested in a program, we are skeptical that they account (in a consistent way) for factors such as the size or structure of the firm's overall research activities.

In Table A1 we present the results of estimating the Poisson regression which forms the basis for $\mu^{ADAPTIVE}$. The dependent variable is the count of patents applied for, and the explanatory variables are $\mu^{POISSON}$ and the log of RESEARCH. Estimated coefficients on both variables are highly significant and have the anticipated positive sign.

Table A2 summarizes the three alternative measures of expected patents. Note again that they are calculated from much longer time series on patenting and research expenditures than the sample used in the regressions in the body of the paper. For the MARKOV and POISSON measures, calculation is based only on years for which the program is "alive" i.e., the MARKOV measure is missing for those programs which were not at least minimally active in the immediately prior year, and the POISSON measure requires that the firm-program is minimally active in at least one year in the past. The final column of reports the sample statistics for the ADAPTIVE measure. In all cases, the reported descriptive statistics are for the sample of 2417 observations used in the main regressions. It is useful to note that the expectation for these three measures are similar though not identical (the average MARKOV expectation is 3.11, the average POISSON expectation is a lower value of 2.43, and the average ADAPTIVE expectation lies in between at 2.89).

BIBLIOGRAPHY

Abel, A. (1984). "R&D and the Market Value of the Firm: A Note," *R & D, Patents, and Productivity*, ed. Zvi Griliches. Chicago: University of Chicago Press. 261-270.

Aghion, P. and J. Tirole (1994). "The Management of Innovation," *Quarterly Journal of Economics*, 109 (4): 1185-1209.

Anderson, E. and D. Schmittlein (1984). "Integration of the Sales Force: An Empirical Examination," *RAND Journal of Economics*, 15: 385-95.

Arora, Ashish (1996). "Testing for Complementarities in Reduced Form Regressions: A Note," *Economics Letters*, 50(1): 51-55.

Athey, S. and S. Stern (1998). "An Empirical Framework for Testing Theories About Complementarity in Organizational Design," NBER Working Paper No. 6600.

Baker, G. (1992). "Incentive Contracts and Performance Measurement," *Journal of Political Economy*, 100: 598-614.

Baker, G., R. Gibbons, and K. Murphy (1994). "Subjective Performance Measures in Optimal Incentive Contracts," *Quarterly Journal of Economics*, 109:1125-56.

Baker, G., R. Gibbons, and K. Murphy (2001). "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics*, forthcoming.

Baker, G, and T. Hubbard (2003) "Incentives in trucking… (title?)" Mimeo, HBS

Brickley, J.A. (1999). "Incentive Conflicts and Contractual Restraints: Evidence from Franchising," *Journal of Law & Economics*, 42 (2): 745-74.

Cockburn, I., Henderson, R. and S. Stern (2000). "Untangling the Origins of Competitive Advantage," *Strategic Management Journal* 21: 1123-1145..

Cockburn, I. and R. Henderson (1994). "Racing to Invest?: The Dynamics of Competition in Ethical Drug Discovery," *Journal of Economics and Management Strategy*, 3(3): 481-519.

Cohen, W. and D. Levinthal (1989). "Innovation and Learning: The Two Faces of R&D," *Economic Journal*, 99(397): 569-596.

Cohen, W., Nelson, R. and J. Walsh (2000). "Protecting their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)," Working Paper No. 7552.

Dasgupta, P. and P. David (1994). "Towards a New Economics of Science," *Research Policy*, 23(5): 487-521.

Gambardella, A. (1995). *Science and Innovation in the US Pharmaceutical Industry*. Cambridge: Cambridge University Press.

Gans, J. and S. Stern (2000). "Incumbency and R&D Incentives: Licensing the Gale of Creative Destruction," *Journal of Economics Management and Strategy*, 9(4): 485-511.

Hart, O. (1995). *Firms, Contracts, and Financial Structure*. London: Oxford University Press.

Hauser, J. (1998). "Research, Development, and Engineering Metrics," *Management Science*, 44(12).

Henderson, R. (1994). "The Evolution of Integrative Capability: Innovation in Cardiovascular Drug Discovery," *Industrial and Corporate Change*, 3(3): 607-30.

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

Henderson, R. and I. Cockburn (1994). "Measuring Competence? Exploring Firm Effects in Drug Discovery," *Strategic Management Journal*, Winter Special Issue, 15: 63-84.

Holmstrom, B. (1982). "Moral Hazard in Teams," *Bell Journal of Economics*, 13: 324-40.

Holmstrom, B.. (1989). "Agency Costs and Innovation," *Journal of Economic Behavior and Organization*, 12(3): 305-27.

Holmstrom, B. and P. Milgrom (1987). "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 55 (2): 303-28.

Holmstrom, B. and P. Milgrom (1991), "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design," *Journal of Law, Economics and Organization*, 7: 24-52.

Holmstrom, B. and P. Milgrom (1994). "The Firm as an Incentive System," *American Economic Review*, 84: 972-991.

Ichniowski, C., K. Shaw, and G. Prennushi (1997). "The Effects of Human Resource Management Practices on Productivity," *American Economic Review*, 87(3): 291-313.

Kaplan, S., Murray, F. and R. Henderson (2003) "Discontinuities and senior management: assessing the role of recognition in pharmaceutical firm response to biotechnology" *Industrial & Corporate Change* Vol 12: pp 203-233.

Lazear, E. (1996). "Incentives In Basic Research" NBER Working Paper No. 5444, January 1996.

Lerner, J. and R. Merges (1998). "The Control of Technology Alliances: An Empirical Analysis of the Biotechnology Industry," *Journal of Industrial Economics*, 46(2): 125-156.

Levin, R., Klevorick, A., Nelson, R. and S. Winter (1987), "Appropriating the Returns from Industrial R&D," *Brookings Papers on Economic Activity*.

Merton, R. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*, Chicago: University of Chicago Press.

Milgrom, P., and J. Roberts (1990). "Bargaining Costs, Influence Costs, and the Organization of Economic Activity," *Perspectives on Positive Political Economy* eds. J. Alt and K. Shepsle Cambridge: Cambridge University Press.

Nichols, N. (1994). "Scientific Management at Merck: An Interview with CFO Judy Lewent," *Harvard Business Review*, January-February, pp. 88-99.

Pakes, A. (1981). "Patents, R&D and the Stock Market Rate of Return," NBER Paper 786.

Romer, P. (1990). "Endogenous Technological Change," *Journal of Political Economy*, 98: S71-S102.

Slade, M. (1996). "Multitask Agency and Contract Choice: An Empirical Exploration," *International Economic Review*, 37(2) :465-486.

Stein, J. (1997). "Internal Capital Markets and the Competition for Corporate Resources," *Journal of Finance*, 52(2): 737-783.

Stephan, P. (1996). "The Economics of Science," *Journal of Economic Literature*, 34: 1199-1235.

Stern, S. (1999). "Do Scientists Pay To Be Scientists?" NBER Working Paper No. 7410.

The Wall Street Journal, Rat whisker article, December 2003

TABLE 1
VARIABLES AND DEFINITIONS

Variable Name	Definition	Unit of Observation
FUNDING VARIABLES		
$RESEARCH_{i,j,t}$	Annual expenditure on drug discovery in program *i* by firm *j* in year *t* in \$M 1986, excluding clinical development	program-firm-year
$\Delta RESEARCH_{i,j,t}$	$RESEARCH_{i,j,t}$ - $RESEARCH_{i,j,t-1}$	program-firm-year
FIRM $RESEARCH_{j,t}$	Annual overall expenditure on drug discovery by firm *j* in year *t* in \$M 1986, excluding clinical development	firm-year
PATENTING VARIABLES		
$PATENTS_{i,j,t}$	Annual number of patent applications in year t granted in at least two of U.S., Japan, EU; by worldwide priority date	program-firm-year
FIRM $PATENTS_{j,t}$	Annual overall number of patent applications in year t granted in at least two of U.S., Japan, EU; by worldwide priority date	firm-year
ORGANIZATIONAL DESIGN VARIABLES		
$PROPUB_{j,t}$	Likert scale variable between 1 and 5, where higher values indicate that the firm promotes individuals on the basis of their standing in the scientific community	firm-year
"HIGH" PROPUB $DUMMY_{j,t}$	Dummy equals 1 for firm j for all years after "switch" to higher level of PROPUB by firm j, 0 else	firm-year
MEASURES OF SHOCK TO APPLIED RESEARCH OUTPUT		
PATENT $SHOCK_{i,j,t}$	$\frac{PATS_{i,j,t-1} - E[PATS_{i,j,t-1}]}{E[PATS_{i,j,t-1}]}$	program-firm-year
SHOCK $(x^S)_{i,j,t}$	PATENT $SHOCK_{i,j,t} * I_{i,j,t-2}$	program-firm-year
MEASURES OF TECHNOLOGICAL ACTIVITY		
COMPETITOR $PATENTS_{i,j,t}$	Annual number of patent applications granted to 29 competitor firms	program-firm-year
RELATED $PATENTS_{i,j,t}$	Annual number of patent applications granted in classes related to a given program	program-firm-year

TABLE 2
MEANS AND STANDARD DEVIATIONS

Variable	N	Mean	Standard Deviation
RESEARCH BUDGET VARIABLES			
RESEARCH	2417	1.58	3.07
ΔRESEARCH	2417	0.10	1.03
FIRM RESEARCH	2417	38.20	26.86
# OF RESEARCH PROGRAMS	2417	10.10	4.37
PATENTING VARIABLES			
PATENTS	2417	3.30	4.60
FIRM PATENTS	2417	90.56	60.29
ORGANIZATIONAL DESIGN VARIABLES			
PROPUB	2417	3.35	1.46
"HIGH" PROPUB DUMMY	2417	0.26	0.44
MEASURES OF SHOCK TO APPLIED RESEARCH OUTPUT			
PATENT SHOCK	2417	0.08	1.30
SHOCK (x^S)	2417	0.26	3.22
MEASURES OF TECHNOLOGICAL OPPORTUNITY			
COMPETITOR PATENTS	2417	40.79	40.79
RELATED PATENTS	2417	7.58	7.58

TABLE 3
SENSITIVITY OF RESEARCH BUDGETS TO APPLIED RESEARCH OUTPUT SHOCKS (x^A)

DEPENDENT VARIABLE = ΔRESEARCH	
YEAR	0.011 (0.005)
SHOCK	0.043 (0.007)
TECHNOLOGICAL OPPORTUNITY CONTROLS	
COMPETITOR PATENTS	0.008 (0.008)
RELATED PATENTS	-0.007 (0.012)
SCALE & MOMENTUM CONTROLS	
$RESEARCH_{t-1}$	-0.018 (0.008)
$\Delta RESEARCH_{t-1}$	0.122 (0.022)
CONSTANT	-0.789 (0.380)
N	2417.00
R-squared	0.04
H_0: $\gamma_{i,j} = \gamma_{i,j}$ for all i,j F(13, 2396) = 8.23, rejected at 1% level	

TABLE 4
RESEARCH FUNDING CHANGE
BY PATENT SHOCK & BASIC RESEARCH INCENTIVE INTENSITY

	LOW PROPUB (PROPUB = 1, 2, or 3)	HIGH PROPUB (PROPUB = 4 or 5)
PATENT SHOCK < 0	-0.02	0.03
PATENT SHOCK > 0	0.16	0.38
"Boost" in Research Funding for Positive Shock	0.18	0.35
Difference in "Boost" by Organizational Form	.17 = 94%	

TABLE 5
γ^A AND PROPUB
THE "REGIME" LEVEL

DEPENDENT VARIABLE = $\hat{\gamma}^A$ **N=14**		
	(5-1)	**(5-2)**
PROPUB	0.034 (0.016)	0.034 (0.019)
YEAR		-0.0006 (0.0065)
CONSTANT	-0.092 (0.058)	-0.138 (0.489)

TABLE 6
RESEARCH BUDGET SENSITIVITY TO APPLIED RESEARCH OUTPUTS SHOCKS: INTERACTION WITH PROPUB
PROGRAM-FIRM-YEAR "POOLED" SAMPLE

DEPENDENT VARIABLE = ΔRESEARCH N=2417				
	(6-1)	**(6-2)**	**(6-3)**	**(6-4)**
SHOCK	0.059 (0.020)	0.006 (0.025)	0.008 (0.025)	
SHOCK INTERACTION TERMS				
SHOCK*PROPUB ($\rho^{S,L}$)		0.015 (0.004)	0.015 (0.004)	0.017 (0.005)
SHOCK* YEAR	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)	
SHOCK*[Year Fixed Effects]				Significant
DIRECT EFFECTS OF PROPUB AND YEAR				
PROPUB		0.018 (0.014)	0.021 (0.015)	0.021 (0.015)
YEAR	0.011 (0.004)	0.0092 (0.0047)	0.0087 (0.0048)	
[Year Fixed Effects]				Significant
TECHNOLOGICAL OPPORTUNITY CONTROLS				
COMPETITOR PATENTS			0.009 (0.008)	0.008 (0.008)
RELATED PATENTS			-0.007 (0.012)	-0.010 (0.012)
PROGRAM SCALE AND MOMENTUM CONTROLS				
$RESEARCH_{t-1}$			-0.020 (0.008)	-0.009 (0.008)
$\Delta RESEARCH_{t-1}$			0.122 (0.022)	0.112 (0.022)
CONSTANT	-0.028 (0.050)	-0.070 (0.062)	-0.048 (0.062)	-0.033 (0.056)
R-Squared	0.022	0.028	0.041	0.071

TABLE 7
RESEARCH BUDGET SENSITIVITY TO APPLIED RESEARCH OUTPUTS SHOCKS: INTERACTION WITH PROPUB. PROGRAM-FIRM FIXED EFFECTS

DEPENDENT VARIABLE = ΔRESEARCH N=2417				
	(7-1) Program-Firm FEs and Controls	**(7-2)** (7-1) w/ PROPUB CHANGE	**(7-3)** Program-Firm FEs, Program-specific trends, and PROPUB CHANGE	**(7-4)** (7-3) w/ Controls
[Program-Firm FE]	Insignificant	Insignificant	Insignificant	Insignificant
SHOCK INTERACTION TERMS				
SHOCK*[Program-Firm FE]	Significant	Significant	Significant	Significant
SHOCK*HIGH PROPUB DUMMY ($\rho^{S,L}$)		0.151 (0.040)	0.115 (0.045)	0.121 (0.044)
SHOCK* YEAR	0.001 (0.002)	-0.004 (0.003)		
SHOCK*YEAR*[Program FE]			Significant	Significant
DIRECT EFFECTS OF PROPUB AND YEAR				
"HIGH" PROPUB DUMMY		0.007 (0.088)	-0.113 (0.089)	-0.012 (0.088)
YEAR	0.022 (0.005)	0.023 (0.007)		
[Year Fixed Effects]			Significant	Significant
TECHNOLOGICAL OPPORTUNITY CONTROLS				
COMPETITOR PATENTS	0.013 (0.012)	0.013 (0.012)		0.015 (0.012)
RELATED PATENTS	-0.004 (0.015)	-0.002 (0.015)		-0.008 (0.014)
PROGRAM SCALE AND MOMENTUM CONTROLS				
$RESEARCH_{t-1}$	-0.156 (0.016)	-0.159 (0.016)		-0.171 (0.016)
$\Delta RESEARCH_{t-1}$	0.145 (0.024)	0.138 (0.024)		0.131 (0.024)
CONSTANT	0.701 (0.957)	0.656 (0.956)	0.572 (1.027)	0.191 (1.000)
R-Squared	0.291	0.296	0.289	0.331

TABLE A1
"ADAPTIVE" MODEL FOR EXPECTED PATENT PRODUCTION

Poisson regression:

$$\ln E(PATENTS_t) = \lambda_0 + \lambda_P \mu_{t-1}^P + \lambda_R \ln(RESEARCH_{t-1})$$

Where μ_t^P is given by $\mu_t^P = \frac{1}{t - T_0} \sum_{s=T_0}^{t-1} PATENTS_s$

DEPENDENT VARIABLE= PATENTS$_t$ N=3446	
μ_{t-1}^P	0.146 (0.002)
ln(RESEARCH$_{t-1}$)	0.121 (0.007)
CONSTANT	0.666 (0.017)
Log-Likelihood	-8511.30

TABLE A2
SUMMARY STATISTICS FOR ALTERNATIVE MEASURES OF EXPECTED PATENTS

Expected Patent Production Measure	μ_t^M	μ_t^P	μ_t^A
Definition	PATS$_{t-1}$	$\frac{\sum_{s=0}^{t-1} PATS_s}{t-1}$	$e^{\hat{\lambda}_0 + \hat{\lambda}_P \mu_{t-1}^P + \hat{\lambda}_R \ln(RESEARCH_{t-1})}$
Mean Expectation:	3.11	2.43	2.89
Std. Deviation of Expectation	4.34	2.95	3.47
N	2417	2417	2417.00