**Optimal Research Strategies when Innovation is Combinatorial**

Matthew S. Clancy*

**Abstract**

I develop a knowledge production function where new ideas are built from combinations of pre-existing elements. Parameters governing the connections between these elements stochastically determine whether a new combination yields a useful idea. Researchers are Bayesians who update their beliefs about the value of these parameters and thereby improve their selection of viable research projects. This approach provides a micro-foundations for knowledge spillovers, knowledge accumulation and fishing out effects. I use a combination of special cases and computer simulations to show that this model generates many stylized features of the research process. In particular, the optimal research strategy is a mix of harvesting the ideas that look best, given what researchers currently believe, and performing exploratory research in order to obtain better information about the unknown parameters. Initially, it may be optimal to perform relatively myopic, rather than exploratory research.

**Key Words:** R&D, Research, Innovation, Knowledge, Bayesian, Science

JEL codes: O31

This Version: May 19, 2015

* Matthew S. Clancy is a Ph.D. candidate, Department of Economics, Iowa State University, Ames, IA 50011. Clancy was supported by a USDA NIFA National Needs Fellowship grant.

**0 - Introduction**

This paper develops a new model of knowledge production, based on two premises.

First, all knowledge is composed of pre-existing parts. There is no *creatio ex nihilo*, wherein new ideas or technologies spring into existence fully formed from out of a void. Look deep enough and even the most creative ideas reveal themselves to be complex structures and arrangements of parts that were already there. These "parts" may be methods, techniques, concepts, mental models, designs, relationships, conventions, symbols, materials, facts, and so forth. What makes an idea new and creative is not *what* it is built from, but *what combination* of parts it is built from.

Second, knowledge is not *just* a combination of pre-existing parts. It is the connections between parts that matters, rather than the parts in and of themselves. Components interact, often in unexpected ways, to produce novel effects possessed by none of the components on their own. These effects may be desirable or undesirable, and learning how parts work together in concert or in opposition is crucial.

This paper will show how such a microfounded model of knowledge production, coupled with a simple model of a learning, innovating agent, is consistent with several stylized facts about the knowledge discovery process. For example;

- Ideas can build on each other, but can also be exhausted.
- There is a natural mechanism for knowledge to accumulate and spillover from one application to another.
- Mature knowledge sectors will be primarily characterized by applied, rather than basic, research.
- Incremental innovation will tend to peter out in the absence of radical innovations.

Some of this model's predictions are less intuitive, but I will argue they are consistent with actual patterns of knowledge discovery:

- The early days of a technological paradigm will often be dominated by applied, rather than basic, research.
- The absence of "moonshots" and other attempts at radical innovation is consistent with a healthy research sector.

The plan for this paper is as follows. After a review of some related literature (Section 1), I will first lay out the knowledge production function used in this paper, and embed this function in a simple model of a solitary agent conducting research (Section 2). I then discuss the optimal research strategy when there is no uncertainty about model parameters (Section 3). In general, however, parameter uncertainty is a key feature of this model, and the next section discusses the nature of researcher beliefs about model parameters (Section 4). I then discuss the optimal research strategy in the special case where the model takes the form of a multi-armed bandit problem (Section 5), before presenting the more complex general version of this model (Section 6). At this point, closed form solutions will be impossible, so I will numerically solve the problem for a large set of differing parameter values (Section 7). I will then characterize these numerically generated solutions (Section

8). Finally, I discuss the results and offer some thoughts on directions for further research (Section 9).

**1 - Background**

The combinatorial nature of knowledge is clearest for physical objects like technology. Consider a laptop. Besides being an object with features and behaviors I value – for example, being able to run programs – it is also a configuration of tightly integrated components: a screen, trackpad, keyboard, hard drive, battery, etc. All of these components existed, in some form or another, before the invention of the laptop. In an important sense, the invention of the laptop consisted in finding a configuration of suitable components that could operate in concert. A similar exercise can be performed on other technologies. For example, a car is an integrated system of wheels, guidance systems, engines, structural supports, etc.

Non-physical creations can also be understood as combinations. Works of fiction draw on a common set of themes, styles, character archetypes, and other tropes; musical compositions rely on combinations of instruments, playing styles, and other conventions; and paintings deploy common techniques, symbols, and conventions. Indeed, even abstract ideas can be understood this way. In an essay on mathematical creation, Henri Poincaré, noted, "[Mathematical creation] consists precisely in not making useless combinations and in making those which are useful and which are only a small minority."[1]

Weitzman (1998) is the first to incorporate this feature of knowledge creation into the knowledge production function. In Weitzman's model, innovation consists of pairing "idea-cultivars"[2] to see if they yield a fruitful innovation (a new idea-cultivar), where the probability an idea-pair will bear fruit is an increasing function of research effort. If successful, the new idea-cultivar is included in the set of possible idea-cultivars that can be paired in the next period. Weitzman's main contribution is to show that combinatorial processes eventually grow at a rate faster than exponential growth processes, so that, absent some extreme assumptions about the cost of research, in the limit growth eventually becomes constrained by the share of income devoted to R&D rather than the supply of ideas. Simply put, combinatorial processes are so fecund that we will never run out of ideas, only the time needed to explore them all.

Weitzman's model is echoed in Arthur (2009), who views all technologies as hierarchical combinations of sub-components. Arthur agrees that the laptop is a combination of screen, trackpad, keyboard, hard drive, battery, and so forth, but goes further, pointing out that, say, the hard drive, is itself a combination of disks, a read/write head assembly, motors to spin and move these assemblies, and so forth. These sub-components themselves are combinations of still further subcomponents. Ridley (2010) also proposes a model akin to Weitzman's, arguing the best innovations emerge when "ideas have sex." This suggests ideas are combinations of genes which, when mixed, yield new offspring. This metaphor also anticipates the second premise, since it is the interaction of genes that matters in most instances.

---

[1] Poincaré (1910), p. 325.

[2] So-called because the hybridization of ideas in the model is analogous to the hybridization of plant cultivars.

Because clearly there is more to invention than simple combination. It does not suffice for engineers to bolt together elements at random, nor should musicians compose with a computer program that generates arbitrary lists of instruments, players and themes. Indeed, if we go back to the musings of Henri Poincaré, the full quote is:

*In fact, what is mathematical creation? It does not consist in making new combinations with mathematical entities already known. Any one could do that, but the combinations so made would be infinite in number and most of them absolutely without interest. To create consists precisely in not making useless combinations and in making those which are useful and which are only a small minority. Invention is discernment, choice.*

-Poincaré (1910), p. 324-325

Later Poincaré compares mathematical creation to the jostling of atoms, which are hoped will hook onto each other in stable configurations. He explains "The mobilized atoms are… not any atoms whatsoever; they are those from which we might reasonably expect the desired solution."[3] In this paper I argue that we create by trying combinations similar in their composition to ideas with desirable features.

Several papers have explored this perspective, devising ways to measure the "distance" between ideas. Jovanovic and Rob (1990) represents a technology by an infinite vector, each element of which ranges between 0 and 1. The elements of this vector have an interpretation as methods, and the value of the element indexes how the method is used.[4] Technologies are production functions and agents learn the mapping from technology vectors to productivity via Bayesian updating. Research consists in changing the values of the elements in a vector and observing the labor productivity associated with the new vector.

A related approach is developed by Kauffman, Lobo and Macready (2000) and Auerswald et al. (2000). These papers follow Jovanovic and Rob (1990) in thinking of technologies as a large combination of distinct operations, although here the length of a technology vector is finite and each element can take on one of a finite number of states (rather than ranging over a continuous interval). The mapping between each technology vector and its productivity level is called a fitness landscape. When states are interdependent, the authors show this landscape is characterized by many local maxima. Innovation in such a model consists of exploring the fitness landscape by changing different operations. The roughly correlated nature of the landscape means small changes are likely to result in productivities that are similar to current levels, and large changes are essentially a draw from the unconditional distribution of productivity values. To reach the global maximum from any given position, it may be necessary to first traverse productivity "valleys." Auerswald et al. (2000) uses the framework to show random deviations in a production process (akin to mutations in biology) can replicate many of the features of so-called "learning curves."

The model that I develop in this paper combines the explicitly combinatorial framework of Weitzman (1998) with the vector based learning models of Jovanovic and Rob (1990), Kauffman,

[3] Poincaré (1910), p. 333-334.

[4] For example, element $k$ might be encoding "drill for oil at location $k$" and the value between 0 and 1 encodes some measure of the depth of drilling.

Lobo and Macready (2000) and Auerswald et al. (2000). This framework permits the derivation of many stylized facts about innovation to emerge from a common set of principles.

## 2 – Model Basics

### 2.1 – The Knowledge Production Function

I now describe formally how ideas are created in this model.

> **Definition 1: Primitive Elements.** Let $Q$ denote the set of primitive elements of knowledge $q$ that can be combined with other elements to produce ideas, where $q \in Q$.
>
> **Definition 2: Pairs.** Let $p$ denote a two-element subset of $Q$, or "pair," and $P$ denote the set of two-element subsets of $Q$, where $p \in P$.
>
> **Definition 3: Ideas.** An *idea* $d$ is a set of pairs $p$, satisfying the condition that if $p_0 \in d$ and $p_1 \in d$, then $p \in d$ for any $p \subseteq p_0 \cup p_1$.

This model assumes a fixed set $Q$ of technological building blocks that can be assembled into ideas, where any idea must contain at least two $q$ from the set $Q$. For convenience, however, I define ideas in terms of the *pairs* of elements contained therein. So for example, an idea combining elements $q_1$, $q_2$, and $q_3$ is represented as the set of subsets $\left((q_1,q_2),(q_1,q_3),(q_2,q_3)\right)$. The condition attached to Definition 3 merely insures the pairs between all elements in the idea are included in the idea, so that we do not have ideas such as $\left((q_1,q_2),(q_1,q_3)\right)$, which uses elements $q_1$, $q_2$, and $q_3$ but does not include the pair corresponding to $(q_2,q_3)$.

There are three important concepts in this model.

> **Definition 4: Compatibility.** The *compatibility* of pair $p$ in idea $d$ is $c(p,d) \in \{0,1\}$. When $c(p,d) = 1$ then the pair $p$ is compatible in $d$. When $c(p,d) = 0$ then the pair $p$ is incompatible in $d$.

Note that $c(p,d) = c(p,d')$ is not generally true. The compatibility of a pair may be equal to 1 in one idea and 0 in another.

> **Definition 5: Affinity.** The probability a pair $p$ is compatible defines its *affinity* $a(p) \in [0,1]$.

The notions of compatibility and affinity are related as follows:

$$c(p,d) = \begin{cases} 1 & \text{with probability } a(p) \\ 0 & \text{with probability } 1 - a(p) \end{cases} \tag{1}$$

Essentially, this model assumes pairs of elements have an underlying tendency to be compatible or incompatible, and this tendency is described by the affinity of the pair.

Lastly, ideas are either *effective* or *ineffective*, where an idea is effective if and only if all the pairs of its constituent elements are compatible.

> **Definition 6: Efficacy.** An idea $d$ is *effective*, represented by $e(d)=1$, iff
>
> $c(p,d)=1\ \forall p\in d$. In all other cases, represented by $e(d)=0$, idea $d$ is *ineffective*.

Restated, affinity determines the probability a pair is compatible, and when all pairs in an idea are compatible, the idea is effective. We may imagine ideas as sets of interacting elements that must be mutually compatible for the idea to prove useful. If any two elements are incompatible, I assume the idea suffers a catastrophic failure that renders it unfit for use. Note the probability an idea is effective can be written as:

$$\Pr(e(d)=1)=E[e(d)]=\prod_{p\in d}a(p) \tag{2}$$

This is the joint probability that every pair in the idea is compatible. Ideas are most likely to be effective when they are composed of elements that have a high affinity for each other, and least likely to be effective when composed of elements with a low affinity for each other.

Whereas I believe that this formulation for the structure of ideas strikes the right balance between simplicity and realism, a few caveats are in order, which I address here.

First, this model assumes every pair is equally important. In reality, technology is modular, with some elements tightly coupled and others only weakly interacting. A desktop computer, for example, consists of a monitor and the computer. The elements that make up the monitor are tightly interacting, but only loosely impacted by the elements in the computer. One way to capture this feature would be to assume, as in Weitzman (1998) that new combinations become elements available for combination in the future. However, as economists have long emphasized, unintended consequences are a prevalent feature of reality. Just because an inventor did not expect two elements to interact with each other does not mean they will not do so in unanticipated ways. Suppose the probability that two elements are incompatible proceeds in two steps. First, there is some probability that the two elements interact with each other. Then, if they do, there is a second probability that they interact poorly, causing the entire assemblage to break down. This is a perfectly valid way to understand what is being captured by the single parameter affinity.

Second, the model assumes that only pairwise interactions between the building blocks of ideas matter. There are obvious counter-examples. Suppose elements $q_1$ and $q_2$ are chemical compounds that do not react unless in the presence of a catalyst $q_3$. Or imagine that a stabilizer regulates the interaction between two components that would otherwise interact in a calamitous manner. Higher order interactions are equally plausible. One impact of allowing for interactions above the pair level would be to make learning more difficult, since observing the behavior of a single pair of elements is less informative if the presence or absence of a third element is crucial. Moreover, if we determined, for example, that only interactions between sets of three elements matter, many of the results could be derived with appropriate redefinitions (for example, affinity would now apply to sets of three elements).

### 2.2 – The Conduct of Research

This production function is used by a researcher who is trying to discover effective ideas. The researcher is a risk-neutral infinitely-lived profit maximizer with a discount factor $\delta \in (0,1)$. She knows every element in the set $Q$, and in each period may choose to conduct a research project on some idea $d$ built from the elements in $Q$.

> **Definition 7: Possible Ideas.** The set $D_P$ is the set of all possible ideas that can be made from elements in $Q$. It contains all subsets of $P$ that satisfy the condition in Definition 3.

> **Definition 8: Eligible Ideas.** A set of eligible ideas $\tilde{D}$ is a subset of $D_P$. It is only sensible to conduct research projects on eligible ideas, and when a research project is attempted, the idea is removed from $\tilde{D}$ at the end of the period.

The set $\tilde{D}$ is primarily intended to indicate the set of *untried* ideas, and so it shrinks as research proceeds. I add to this set an additional element, the null set $d_0 \equiv \{\varnothing\}$, which represents the option not to conduct research in a period.

> **Definition 9: Available Actions.** The researcher's set of available actions is $D \equiv d_0 \cup \tilde{D}$.

Note that because $d_0 \notin \tilde{D}$, if the researcher chooses not to conduct research, then this option is not removed from its action set in the next period.

In principle, the researcher "knows" every idea that can be built from elements in $Q$, in the same sense that I "know" every economics article that can be written with words and symbols in my repertoire. However, just as I do not know whether any of these articles are good until I think more about them, or actually write them out, the researcher does not learn if an idea is effective until she decides to conduct research on it.[5] Indeed, research is costly, requiring investments of time and other resources. I assume that research on any idea has cost $k(d)$, known to the researcher, and that the option $d_0$, to do nothing, has $k(d_0) = 0$.

The return from conducting research is a reward $\pi(d)$, also known to the researcher, which is received if the idea is eligible and discovered to be effective. This reward could indicate a prize for innovation from the government, or the sale of patent rights over the idea to a firm, or some other incentive for innovation. Because each chosen idea is removed from the set of eligible ideas $D$ at

[5] Jorge Luis Borges tells a parable of an infinite library containing books with every combination of letter and punctuation mark. In this library, there is a book resolving the basic mysteries of humanity, since every possible book exists, but finding the book and verifying it is true amongst all the gibberish and babel is a daunting task for the library inhabitants. See Borges (1962).

the end of a period, the researcher cannot claim a reward for the same idea multiple times. I assume the reward value of the outside option $d_0$ is always zero.

Hence, a researcher who chooses to conduct research on idea $d$ expects to receive a net value of:

$$\pi(d)E[e(d)]-k(d) \tag{3}$$

The idea is successful with probability equal to the expected efficacy $E[e(d)]$, in which case the researcher obtains a reward $\pi(d)$. Whether the idea succeeds or not, the researcher pays up front research costs $k(d)$. This formulation of the innovator's problem is not unusual, except for the term $E[e(d)]$, which is determined by the knowledge production function described earlier.

Recall that $e(d)=1$ (an idea is effective) if and only if all of its pairs are compatible. This implies the probability distribution of $e(d)$ is:

$$e(d)=\begin{cases}1 & \text{with probability } \prod_{p\in d} a(p) \\ 0 & \text{with probability } 1-\prod_{p\in d} a(p)\end{cases} \tag{4}$$

Therefore:

$$E[e(d)]=\prod_{p\in d} a(p) \tag{5}$$

Hence, if the researcher knows the affinity of each pair, she can compute the expected efficacy of every idea. In general, I will assume the researcher does *not* know the true affinity of each pair and must infer its likely value from the outcomes of research projects.

Before proceeding to this more complex and realistic case, I discuss the special case where the agent knows the true affinity of each pair with certainty.

**3 – Special Case 1: Affinity is Known**

Suppose there is a researcher with perfect knowledge of $a(p)$. In any period, the researcher's problem is to determine which idea, if any, to attempt. Define the expected present discounted value of the optimal strategy in period $t$ as:

$$V_t=\sum_{\tau=0}^{\infty}\delta^{\tau}\left(\pi(d_{t+\tau})E[e(d_{t+\tau})]-k(d_{t+\tau})\right) \tag{6}$$

where $d_t$ denotes the optimal decision in period $t$. Note that, because no relevant information is revealed by the outcome of research, the optimal choice in period $t+\tau$ depends only on information available in period $t$. Recall also that researchers can always choose $d_0$, to obtain a payoff of zero with certainty. Because the set of eligible ideas is finite, the researcher will always resort to choosing $d_0$ in the end.

The optimal strategy in the absence of learning is simple:

> **Remark 1: Optimal strategy with certainty.** In each period, the optimal strategy is to choose the eligible idea with the highest $E[e(d)]\pi(d)-k(d)$ so long as $E[e(d)]\pi(d)-k(d)\geq 0$. If no idea satisfies this, then choose $d_0$.

Because there is no learning in this special case, the researcher's problem collapses to choosing the order in which to consume a set of lotteries. Because the researcher is risk-neutral and discounts the future, she orders these lotteries in descending expected (net) value. Furthermore, the researcher never attempts a lottery where cost exceeds expected value. Because there is no learning, and because the researcher uses up the best lotteries first, the following remark holds.

> **Remark 2: Value decreases over time.** The anticipated value of research declines over time, i.e.,
>
> $$V_t \geq V_{t+\tau} \quad \forall t,\tau > 0 \tag{7}$$

In the absence of learning, the best research ideas are used up ("fished out") first, so that the value of conducting research falls over time. This stands in contrast to the prevalent view that knowledge is cumulative, and that today's researchers can accomplish more than their forebears by building on their accomplishments ("standing on the shoulders of giants"). This result is not general, of course, but a consequence of the absence of learning.

In general, the researcher does *not* know the affinity of a pair with certainty. However, the problem faced by a researcher becomes close to this limiting special case when they already possess very good information about most pairs, so that firms are *nearly* certain about the true affinity of each pair. Thus, mature technology sectors should be more characterized by myopic research strategies than young ones. Alternatively, it may be difficult for firms to capitalize on information they learn, if the outcomes of research is difficult to conceal from other firms. If there are many firms operating in a sector, and if knowledge diffuses rapidly, then competitors can expropriate the value of information. If it is difficult to conceal information (for example, if discoveries are products that can be easily reverse engineered), then firms may behave more or less myopically, even if the sector is young. Finally, firms may also behave myopically if learning is very costly or difficult. In any case, a discussion of firm learning is necessary. I turn to this topic next.

## 4 - Research and Learning

In general, I assume the affinity $a(p)$ between a pair of elements is *ex ante* unknown to researchers. Instead, researchers are Bayesians with prior beliefs over the possible distribution of $a(p)$. Though the researcher does not observe $a(p)$ directly, she can make educated guesses based on the tendency of $p$ to be compatible or incompatible. Using her beliefs about the affinities between all pairs in an idea, she can compute the probability an idea will be effective. In more formal terms, a crucial part of the discovery process is the inference of likely affinity values from the compatibility or incompatibility of component-pair interactions.

I impose the following assumption on the researcher's beliefs:

**Assumption 1: Independence of Affinity.** The researcher believes $a(p)$ is independently distributed for all $p$.

As long as this assumption stands, the updating of beliefs about any $a(p)$ depends only on observations on the pair $p$ alone. If $a(p)$ were not independently distributed, it would be necessary to also take into account the observations on correlated pairs, greatly complicating the problem.

Each observation of compatibility is the outcome of a Bernoulli trial governed by the pair's true affinity, with the two possible states being compatibility (probability $a(p)$) or incompatibility (probability $1-a(p)$). Given $s$ instances of compatibility ("success") and $f$ instance of incompatibility ("failure"), the researcher updates her beliefs according to Bayes law under the Bernoulli distribution:

$$\Pr(a(p)=\tilde{a} \mid s,f) = \binom{s+f}{s}\tilde{a}^s(1-\tilde{a})^f \frac{\Pr(a(p)=\tilde{a})}{\int_0^1 \binom{s+f}{s} a^s (1-a)^f \Pr(a(p)=a)\,da} \tag{8}$$

where $\binom{s+f}{s}\tilde{a}^s(1-\tilde{a})^f = \Pr(s,f \mid a(p)=\tilde{a})$.

In the remainder of the paper, I impose the following assumption:

**Assumption 2: Beta Distributions.** The researcher believes $a(p)$ follows a beta distribution.

The beta distribution has the useful property of being the conjugate family of a Bernoulli distribution. The conjugate family for a distribution defines a class of distributions such that, if the prior distribution belongs to the conjugate family, then the posterior distribution will as well. In this case, if the prior distribution for an affinity belongs to the beta distribution, then after updating the researcher's beliefs by observing a set of Bernoulli trials, the posterior distribution will also be a (different) beta distribution. This structure maintains a constant form for the beliefs of the researcher as her information varies. Moreover, the form of a beta-distribution is sufficiently flexible to enable the exploration of many kinds of assumptions about the prior beliefs of the researcher.

The probability density function of a beta distribution takes the following form:

$$f(a) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} a^{\alpha-1}(1-a)^{\beta-1} \tag{9}$$

where $\Gamma(x)$ is the gamma function. The distribution's support is over the [0,1] interval with its shape governed by the parameters $\alpha > 0$ and $\beta > 0$. Changing $\alpha$ and $\beta$ can yield a centered bell shape, highly skewed distributions, and U-shaped distributions. It can be shown that, given a beta distribution: [6]

$$\int_0^1 \binom{n}{s} a^s (1-a)^{n-s} f(a) da = \binom{n}{s} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(s+\alpha)\Gamma(n-s+\beta)}{\Gamma(n+\alpha+\beta)} \tag{10}$$

Combining equations (8), (9), and (10), the updated beliefs of the researcher given $s$ instances of compatibility from $n$ total observations is:

$$\Pr(a(x) = \tilde{a} \mid s,n) = \frac{\Gamma(n+\alpha+\beta)}{\Gamma(s+\alpha)\Gamma(n-s+\beta)} \tilde{a}^{s+\alpha-1} (1-\tilde{a})^{n-s+\beta-1} \tag{11}$$

Note that this is equivalent to a beta distribution with $\alpha' = s + \alpha$ and $\beta' = n - s + \beta$. Hence, defining $\alpha(p)$ and $\beta(p)$ to be the initial parameters governing the prior beliefs about of $a(p)$, after observing $s(p)$ instances of compatibility and $n(p) - s(p)$ instances of incompatibility, the researcher believes $a(p)$ to be governed by a beta distribution with parameters $\alpha(p) + s(p)$ and $\beta(p) + n(p) - s(p)$. The expected value of such a distribution[7] is given by:

$$E[a(p) \mid n(p), s(p)] = \frac{\alpha(p) + s(p)}{\alpha(p) + \beta(p) + n(p)} \tag{12}$$

Note that as the number of observations grows large, the expectation converges to $s(p)/n(p)$, which is simply the proportion of observations where compatibility is observed. The sum $\alpha(p) + \beta(p)$ determines the relative weight put on new observations and the initial beliefs, and is a measure of initial certainty.

When a researcher knows the affinity of all pairs with certainty:

$$E[e(d)] = \prod_{p \in d} a(p) \tag{13}$$

and given Assumption 1, this can be expressed as:

$$E[e(d)] = \prod_{p \in d} E[a(p)] \tag{14}$$

[6] See Casella and Berger (2002) p. 325.
[7] See Casella and Berger (2002) p. 325.

where $E[a(p)]$ is given by equation (12). This expression captures the core notion of this model: ideas are more likely to be effective ("useful") when composed of elements that have worked well together frequently in the past.

Under some special assumptions, this learning framework, coupled with our model of innovation, can be modeled as a multi-armed bandit problem, which I discuss in the next section.

**5 – Special Case 2: Research as a Multi-armed Bandit Problem**

To isolate the effects of learning in the researcher's problem, I reformulate the problem as a multi-armed Bernoulli bandit problem. While this imposes strong restrictions on the model, the advantage is that there exists a large literature on such problems.[8] In such problems an agent must choose between *n* options, each of which offers a reward with a fixed probability unknown to the agent. Over time, as the researcher observes the frequency with which she receives a reward from any given option, she obtains a progressively better estimate of the underlying probability of receiving a reward from that option. The agent's problem is to balance a myopic strategy that selects the option currently believed to be most favorable, and a far-sighted strategy which seeks to gather information on other options so that the true best option can be found. Essentially, she must make a trade-off between exploitation and exploration.

Such models have a well-known solution technique, called a Gittins index (discussed below). My model takes the form of a multi-armed Bernoulli bandit problem under the following conditions:

1. There are $n$ pairs $p_i$, where $i = 1,...,n$. These pairs have no elements in common.
2. The researcher correctly believes $a(p_i)$ follows a beta distribution with parameters $\alpha_i$ and $\beta_i$.
3. Each pair may be combined with any number of element from the set $Q$, which is infinitely large. These ideas:
    a. Have prize value $\pi(d) = 1$.
    b. Have cost $k(d) = k$.
    c. Have expected efficacy $E[e(d)] = E[a(p_i)]$ (which implies the researcher knows the other affinities are equal to 1 with certainty).
4. No other ideas are eligible.

In this setting, the researcher's problem collapses to choosing which pair $p_i$ to combine with elements in $Q$ (the identity of the elements does not matter, since they all have the same prize value, cost, and expected affinities). The researcher must balance the expected net reward:

$$E[a(p_i)] - k \tag{15}$$

[8] See Gittins, Glazebrook, and Weber (2011).

against the value of learning more accurately the true value of each $a(p_i)$. Because the researcher believes all affinities except for $a(p_i)$ are equal to one, any time the attempted idea is effective, the researcher learns the pair $p_i$ is compatible, and vice-versa.

Specifically, the researcher's problem is characterized by a Bellman equation of the form:

$$V(B) = \max\left\{ \max_{p_i} \left[ \frac{\alpha_i}{\alpha_i + \beta_i} \{1 + \delta V(B_{i+})\} + \frac{\beta_i}{\alpha_i + \beta_i} \delta V(B_{i-}) - k \right] , 0 \right\} \quad (16)$$

where

$$\begin{aligned} B &= ((\alpha_1, \beta_1), \dots, (\alpha_i, \beta_i), \dots, (\alpha_n, \beta_n)) \\ B_{i+} &= ((\alpha_1, \beta_1), \dots, (\alpha_i + 1, \beta_i), \dots, (\alpha_n, \beta_n)) \\ B_{i-} &= ((\alpha_1, \beta_1), \dots, (\alpha_i, \beta_i + 1), \dots, (\alpha_n, \beta_n)) \end{aligned} \quad (17)$$

Taking this equation from left to right, the researcher's problem is first to choose whether or not to conduct research. If she does not, she obtains 0 with certainty. Moreover, if the researcher ever chooses to quit research in one period, she will do so in all subsequent periods, since her information and action set will be unchanged in the following period.

If the researcher does choose to conduct research, she must decide the best pair $p_i$ to select. With probability $\alpha_i / (\alpha_i + \beta_i)$, the idea is effective and the researcher obtains a reward equal to 1. Moreover, in the next period, she will update her beliefs in accordance with equation (12), so that her beliefs are described by the vector $B_{i+}$. Therefore, in the next period, she obtains $V(B_{i+})$, discounted by $\delta$. Conversely, with probability $\beta_i / (\alpha_i + \beta_i)$ the idea is ineffective and she obtains no reward this period. Moreover, if an idea is ineffective, the researcher learns pair $p_i$ is incompatible and the researcher will update her beliefs to vector $B_{i-}$. In the next period, she will obtain $V(B_{i-})$, discounted by $\delta$. Finally, in either case, she pays $k$ to conduct research.

This formulation is equivalent to a standard multi-armed bandit problem with expected affinity of each pair corresponding to the fixed Bernoulli probability of receiving a reward. Using the Gittins Index approach, we can obtain some clear insights, summarized in the following remarks.

> **Remark 3: Optimal strategy with learning.** The optimal strategy in every period is to choose the option with the highest Gittins Index $\lambda_i(\alpha_i, \beta_i)$ where
>
> $$\lambda_i(\alpha_i, \beta_i) \equiv \sup\{\lambda_i : v_i(\alpha_i, \beta_i, \lambda_i) = 0\} \quad (18)$$
>
> and $v_i(\alpha_i, \beta_i, \lambda_i)$ is equal to:
>
> $$\max\left\{ \frac{\alpha_i}{\alpha_i + \beta_i} (1 + \delta v_i(\alpha_i + 1, \beta_i, \lambda_i)) + \frac{\beta_i}{\alpha_i + \beta_i} \delta v_i(\alpha_i, \beta_i + 1, \lambda_i) - k - \lambda_i , 0 \right\} \quad (19)$$

A proof is presented in the appendix. See Gittins, Glazebrook, and Weber (2011) for more discussion.

Note that $v(\alpha_i, \beta_i, \lambda_i)$ only depends on the beta parameters of one pair, $p_i$. This decomposes the $n$-dimensional choice problem into $n$ one-dimensional problems. The Gittins index $\lambda_i(\alpha_i, \beta_i)$ can be thought of as a riskless payment the researcher can receive in lieu of the reward from choosing $p_i$, chosen so the researcher is exactly indifferent between the two options. It accounts for the expected value in this period, equal to $\alpha_i / (\alpha_i + \beta_i) - k$, plus the prospects of achieving a better or worse outcome in subsequent periods, as the researcher's beliefs are updated. Choosing the highest Gittins index in every period therefore accounts for the rewards in the current period, plus the potential gains from better information.

Note, however, that the researcher's expected payoff is not equal to the Gittins index. This is because a Gittins index is computed with reference only to one pair. The index tells the researcher what to do, but it does not say how much she should expect to make. However, the value of research still follows some principles:

> **Remark 4: Value rises with affinity (value of learning).** The expected value of research is nondecreasing in $\alpha_i$.

In this model, value only comes from obtaining rewards, which occur with probability $\alpha_i / (\alpha_i + \beta_i)$. If the researcher fixes an optimal strategy, and then one $\alpha_i$ is increased, the researcher cannot be worse off. Neither will she be worse off if we let her re-select the optimal strategy.

Multi-armed bandit problems also have the following feature:

> **Remark 5: Stick with the winner.** The optimal strategy follows a "stick-with-the-winner" formulation, i.e.,

$$\lambda_i(\alpha_i + 1, \beta_i) > \lambda_i(\alpha_i, \beta_i) > \lambda_i(\alpha_i, \beta_i + 1) \qquad (20)$$

A proof is presented in Bellman (1956).

Once an idea has been found successful, in this model, the probability it will be successful in the next period as well is increased, which makes the choice still more favorable in the next period and the opportunity cost of trying something else higher. Therefore, the researcher always sticks with a winning pair, at least until it stops working (although this is not sufficient for her to switch his strategy either).

Besides preferring winners, the researcher also prefers the pair about which less is known:

> **Remark 6: Favor uncertainty.** When the expected reward is the same, an optimal strategy chooses the option where more is learned, i.e.,

$$\lambda_i(\alpha_i, \beta_i) > \lambda_i(m\alpha_i, m\beta_i) \qquad (21)$$

if $m > 1$.

A proof is presented in Gittins and Wang (1992).

Note that the expected value of a beta distributed variable with parameters $(\alpha, \beta)$ and $(m\alpha, m\beta)$ is the same, since:

$$\frac{m\alpha}{m\alpha + m\beta} = \frac{\alpha}{\alpha + \beta} \tag{22}$$

However, a pair with $(\alpha + 1, \beta)$ has a higher expected value than one with $(m\alpha + 1, m\beta)$. In the next period, the potential benefits are higher for the more uncertain idea. At the same time, the potential downsides also looms larger for the more uncertain choice. However, since the researcher always has the option to quit research, downside risks are capped at 0. This leads agents to prefer ideas from which they can learn more, that is ideas where they have less certainty about the affinity of their pairs.

Finally, the above results imply the following.

> **Remark 7: Value is rising and concave in success:** The expected value of research is increasing in the number of times any given pair is found to be effective, denoted $s(p)$, and bounded from above by $(1-k)/(1-\delta)$.

That $V(B)$ is increasing in $s(p)$ is simply a reformulation of Remark 4, since the researcher's $\alpha_i$ parameter is updated to $\alpha_i + s(p)$ after observing $s(p)$ instances of compatibility. Moreover, since researchers stick with winners, as $s(p)$ continues to increase, the expected payoff begins to resemble the payoff from simply playing the same pair in each period. This is a concave function of $s(p)$, bounded from above by $(1-k)/(1-\delta)$.

To summarize, in the multi-armed bandit formulation of the researcher's problem, the optimal strategy is a mix of myopic and far-seeing strategies, since the Gittins index takes into account both the immediate payoff and the distant future payoffs. Researchers, somewhat paradoxically, prefer both proven research paths (they stick with winners) and unproven *and* untested research projects (they favor uncertainty). The value of research increases when research is successful, so that research is cumulative and has a standing-on-the-shoulders-of-giants effect. However, eventually, the payoff from success stops increases as it approaches a ceiling, given by the present discounted value of a successful research project in every period.

However, to derive these results, I had to rely on some extreme assumptions that simplified the combinatorial dynamics of this model. This simplified special case is nonetheless a close approximation of the researcher's problem in some settings. It bears similarities to the literature on general purpose technologies. General purpose technologies are those like steam power, electricity, or computers, for which the new technology has many applications (Helpman 1998). These applications are well understood, and can be relatively easily developed, if the general purpose technology is developed. In the terms of this model, the (potential) new general purpose technology is like the pairs whose affinity is unknown, and the applications are a pool of technological building blocks where the affinity between them is very high, but where the technologies built from them alone have either been exhausted or are of low value.

Imagine, for example, several different ways to generate electrical power. If a successful platform for generating electrical power can be discovered, the researcher is confident it can be easily combined with a functionally limitless set of existing equipment. The different ways of producing electricity can be modeled as different pairs of elements (say, magnets and wires), each of which has an unknown affinity. The different technologies awaiting a new power source can be modeled as sets drawn from a set $Q$. Such a problem is very similar to the multi-armed bandit approximation discussed here, and the solution concept would be similar. In particular, researchers could proceed by independently weighing the prospects of each alternative method of generating electricity, as in the Gittins index approach. If the first approach turned out to work all the time, the researcher would never bother trying others. The value of her research agenda would asymptote at its maximum, with the researcher successfully extending her electrical system to a new application in every period. If the approach fails, however, in choosing the next method, she will put more weight on methods that are comparatively less well understood.

In the next section, I will incorporate uncertainty about the true affinity into a more general setting.

## 6 – The General Case

### 6.1 – Learning in the General Setting

To implement the Bayesian belief updating presented in section 4, I summarize the researcher's beliefs by the vector $B$ where:

$$B = \left( (\alpha_1, \beta_1), \ldots, (\alpha_i, \beta_i), \ldots, \left(\alpha_{m(N)}, \beta_{m(N)}\right) \right) \quad (23)$$

and $m(n) \equiv n(n-1)/2$. These beliefs are now updated via a stochastic vector $\omega(d)$ of information revealed by a research project over idea $d$. This vector has the same number of elements as $B$ and is defined so that after conducting a research project on $d$, the updated beliefs vector $B'$ is given by:

$$B' = B + \omega(d) \quad (24)$$

For example, if a research project on some $d'$ reveals pair $p_1$ is compatible and pair $p_{m(n)}$ is incompatible, and does not reveal any other information, then $\omega(d)$ takes the form:

$$\omega(d') = ((1,0),(0,0),\ldots,(0,0),(0,1)) \quad (25)$$

In Section 5, I restricted attention to research projects where $E[e(d)] = E[a(p_i)]$. In this setting, it was obvious that a research project would reveal the compatibility $c(p_i, d)$: since the only uncertainty pertained to this pair, if the idea was effective, it must have been that $c(p_i, d) = 1$, and if the idea was ineffective, it must have been that $c(p_i, d) = 0$.

In the general setting, where *every* pair in an idea may be compatible or incompatible, what is learned from research is more complicated. Whenever an idea is effective, it must be that $c(p,d)=1$ for all $p \in d$ (since this is the definition of efficacy). Accordingly, whenever an idea is effective, the researcher must know that every $p \in d$ was compatible and update her beliefs accordingly.

However, whenever an idea is ineffective, any configurations of compatibilities where at least one $c(p,d)=0$ is capable of generating the same result. Which pair compatibilites are revealed in this case is not a trivial matter, and can easily make the model intractable or embed unrealistic features. The following procedure has two main virtues, discussed more below. First, I believe it has realistic features about what can be learned from successes and failures. Second, it keep beliefs independent, which maintains the model's tractability.

The information about the compatibilities $c(p,d)$ of each pair $p \in d$, is generated by the following stochastic process.

> **Assumption 3: Learning From Research (Frustrations of Failure).** The information revealed about the compatibility of pairs is determined according to the following procedure:
>
> 1. A pair $p \in d$ is randomly drawn with equal probability from among the pairs whose compatibility has not already been selected.
> 2. The compatibility of this pair is added to the research project's revealed information.
> 3. If $c(p,d)=1$ and unselected pairs remain, return to step one and repeat the above procedure. If $c(p,d)=0$ or if no unselected pairs remain, do not add any more compatibilities to the research project's revealed information.

When this procedure is completed, the researcher observes a packet of information $\omega(d)$. If an idea is effective, this revelation procedure will reveal that all of its pairs are compatible. If the idea is ineffective, it will reveal some, but possibly not all, of the compatibilities of the pairs that make up $d$. Specifically, the above revelation procedure will never reveal more than one pair is incompatible.[9]

This revelation mechanism is meant to capture the frustrations of failure. Researchers often have some indication of where things began to go wrong – for example, a proof step that does not go through, or an engine part that overheats – but a full understanding of why the idea failed is often elusive. This partial revelation of information about what part of the innovation failed is captured by the fragmentary knowledge of which pairs are compatible and incompatible when an idea is ineffective. If information is not fragmentary when an idea fails, then researchers could in principle learn about the affinity between every idea simply by performing research on an idea that draws on every element in the set $Q$.

[9] This assumption could be made more general with the introduction of a parameter $\eta \in [0,1]$, so that when a pair with $c(p,d)=0$ is revealed, the revelation procedure stops with probability $\eta$. The model presented here is then the special case with $\eta=1$.

Note also the researcher is unlikely to learn much from a research project with many incompatible pairs, because the probability of encountering an incompatibility that stops the revelation process early is high. This is meant to capture the notion that we do not, on average, learn as much from a project that is wrong on many levels. When an idea has only one or two incompatible pairs, the researcher may learn a lot or a little by trying the research project. If she is lucky, it is the kind of idea where, although the idea is ineffective, she can get a long way before hitting a roadblock. Such a research project might be represented by one where the first incompatible pair is only reached after a long series of compatible ones. If she is unlucky, the idea is the kind in which it is very difficult to make any headway until a certain problem is cracked. This would be represented by an idea where the revelation of compatibilities is quickly stopped by an incompatible pair.

Lastly, not that $\omega(d_0) = ((0,0),\ldots,(0,0))$ by assumption: agents learn nothing when they choose not to do a research project.

**6.2 – The Researcher's Problem in the General Setting**

In this paper I limit attention to the case where the researcher has no competition, thereby evading strategic considerations. Hence the researcher's problem is to find an optimal policy function $d^*(D,B)$ mapping from available actions $D$ and beliefs $B$ to an optimal choice.

The policy function $d^*(D,B)$ maximizes:

$$V(D,B) = E\left[\sum_{t=0}^{\infty} \delta^t \left( e\left(d^*(D_t,B_t)\right)\pi\left(d^*(D_t,B_t)\right) - k\left(d^*(D_t,B_t)\right)\right)\right] \tag{26}$$

where

$$\begin{aligned} D_{t+1} &= d_0 \cup \left(D_t \setminus d^*(D_t,B_t)\right) \\ B_{t+1} &= B_t + \omega\left(d^*(D_t,B_t)\right) \end{aligned} \tag{27}$$

The payoff function in (26) is the discounted sum of expected per-period returns from choosing idea $d^*(D_t,B_t)$ in period *t*. The researcher obtains $\pi\left(d^*(D_t,B_t)\right)$ if $d^*(D_t,B_t)$ is effective, which occurs with probability $E\left[e\left(d^*(D_t,B_t)\right)\right]$, but pays $k\left(d^*(D_t,B_t)\right)$ either way. As noted earlier, if the researcher chooses $d_0$, then $\pi(d_0) = E[e(d_0)] = k(d_0) = 0$.

It is instructive to write equation (26) as a Bellman equation:

$$V(D,B) = \max_{d \in D}\left\{E[e(d)]\pi(d) - k(d) + \delta E[V(D',B')]\right\} \tag{28}$$

where

$$\begin{aligned} D' &= d_0 \cup (D \setminus d) \\ B' &= B + \omega(d) \end{aligned} \tag{29}$$

This formulation makes clear that the choice of idea has a payoff in the current period, but also an impact on the future, through the $E\left[V\left(D',B'\right)\right]$ term. Agents may prefer to take a loss in the current period, in order to learn and increase their payoffs in the future. This formulation also makes clear that if it is ever optimal for the researcher to choose $d_0$ at some stage, then it is optimal for her to do so in every subsequent period because such a choice ensures $D' = D$ and $B' = B$.

**7 – Numerical Simulation**

**7.1 – Solution Methodology**

There is no general closed-form solution for this problem, but it can (in principle) be solved by backwards induction. Because the set of possible ideas is finite, and (as noted above) the researcher never pauses then restarts research, it is certain that from period $|D|$ on the researcher will choose $d_0$ in every stage. Thus, $V\left(D,B\right)=0$ in period $|D|$, no matter what the researcher does in period $|D|-1$. Since the researcher knows the next period payoff with certainty, she can find the best choice in period $|D|-1$, and work backwards toward period 0.

**7.2 – Simulation Assumptions**

Although a closed form solution is not feasible, I would like to highlight some common contours of an optimal solution. Unfortunately, this model is also beset by the "curse of dimensionality," so that even a general numerical approximation is difficult to obtain.[10] The curse of dimensionality is a name applied to problems where the computational resources needed to solve them grow very quickly.[11] For example, consider the number of parameters needed to characterize the state-space for a problem with 3, 4, and 5 elements. Given 3 elements, there are 3 pairs and 3 affinities, the beliefs about which are described by 2 parameters each, so that $B$ has six dimensions. Three elements also implies there are 4 possible ideas. Since each of these can be available or unavailable, there are $2^4$ different sets $D$ associated with every $B$. If I increase the number of elements to 4, then $B$ has 12 dimensions and there are $2^{11}$ possible sets of $D$ for each vector of beliefs. If I increase the number of elements to 5, then $B$ has 20 dimensions and there are $2^{26}$ different sets $D$ associated with each one. Obtaining a good approximation of a state space with 20 dimensions and $2^{26}$ discrete action states is very challenging.

Given the foregoing, my approach is to instead solve a manageable (small) version of the problem 100 times, and then to use a regression analysis to see if characteristics of Remarks 1-7 hold for these optimal solutions. Essentially, I will project a linear approximation onto a highly complex and non-linear solution, to check for the validity of Remarks 1-7 outside of the special cases for which they were derived. It will be important that the choice of problem to solve is sufficiently rich that it

[10] Of course, if a numerical simulation is difficult to achieve for a computer, then it is likely to be even more difficult for a human brain. The following section may also be interpreted as one set of heuristics that computationally-constrained researchers might use to approximate the optimal research strategy.

[11] Powell (2011) provides a good overview of the problem and possible approximation methods.

captures the complexity of the model, but remains solvable. The basic structure of the problem I will solve takes the following form:

1. There are four elements $q \in Q$ that can be combined into ideas.
2. There are ten eligible ideas. Six ideas are composed of pairs of elements, and four ideas are composed of three elements.[12]
3. The cost of ideas is normalized to 1.
4. The researcher's discount factor is 0.95.
5. The researcher's beliefs about the affinity of each pair follow a beta distribution. Each pair has unique beta parameters. I discuss this more below.
6. Each idea has a unique prize value $\pi(d)$, known to the researcher. I discuss this more below.

These six conditions capture several key features. First, I model the beliefs of the researcher (as in Special Case 2), while allowing ideas to depend on multiple pairs and to have different reward values (as in Special Case 1). Second, each pair can be used up to three times (once in a two-element idea, and twice in a three-element idea), so that updating of beliefs can happen more than once. Third, information revealed from one idea spills over to other ideas. Fourth, some ideas are subsets of others.

Most importantly, this problem can be solved in a reasonable amount of time. I have written a computer program in python to solve this problem. To begin, I define the possible sets $D$ in which the researcher may find herself. Since there are 10 ideas, and each idea may be either eligible or ineligible, there are $2^{10} = 1{,}024$ distinct sets of eligible ideas. For each of these sets, I define the potential belief vectors of interest. Since I know the initial beta parameters of every $a(p)$, the program can exhaustively list the vectors $B$ that might be attained in any given set.

Once I have a set of $(D,B)$ states, I work backwards. The program begins by evaluating the null set $D = \{\varnothing\}$, where the only available option is to quit research and earn zero with certainty (for every belief vector $B$). Next, using this result, it evaluates the best action for each set with just one eligible idea remaining. When there is just one eligible idea, the problem simplifies to:

$$V(D,B) = \max\{E[e(d)]\pi(d) - k(d), 0\} \tag{30}$$

Using these results, the program evaluates the best action for each set with two eligible ideas remaining, which has the form of equation (28). At each stage, it uses the researcher's beliefs and the revelation procedure to compute the probabilities associated with each state the researcher may find herself in next period.

[12] Including the 11th idea, composed of all four elements, dramatically increases the computational time to solve, without adding much insight, so I omit it. Alternatively, we might assume this idea has prohibitively high costs, but is theoretically eligible.

After working backwards, the program obtains a mapping from every $(D, B)$ state to a best action. With four elements and ten ideas, this program still takes approximately an hour to solve depending on computer processing power. See the appendix for a more detailed description of this program.

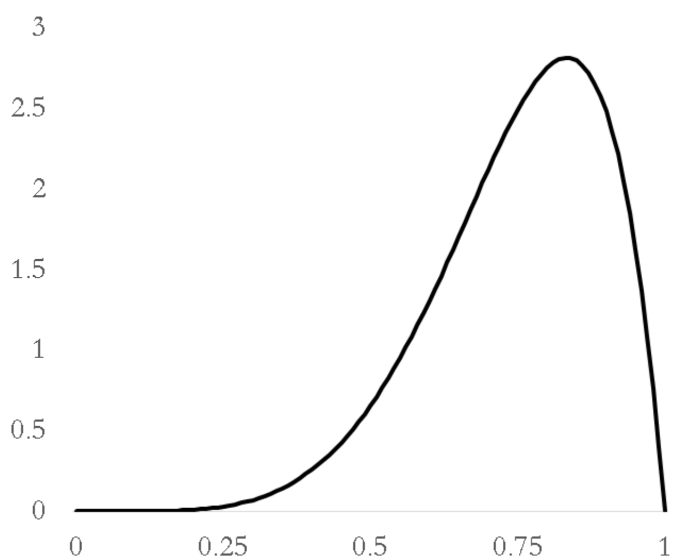
**7.3 – Prizes and Costs**

I solve 100 variations of the above problem which differ in researcher beliefs and rewards $\pi(d)$. To obtain the researcher's beliefs about a given pair, I derive $\alpha_i$ and $\beta_i$ from the following equations, where $\tilde{a}$ is drawn from a beta distribution with $\alpha = 6$ and $\beta = 2$:

$$\begin{aligned} \tilde{a} &= \frac{\alpha_i}{\alpha_i + \beta_i} \\ 0.4 &= \alpha_i + \beta_i \end{aligned} \tag{31}$$

I chose to draw the initial expected affinity from a beta distribution with with $\alpha = 6$ and $\beta = 2$ because such a distribution has an expected value of 0.75, but all values above 0.4 occur with relative frequency, as can be seen in Figure 1. This yields a wide array of initial beliefs about pairs.

**Figure 1:** PDF of Beta Distribution with $\alpha = 6$ and $\beta = 2$




I calibrated $\alpha_i + \beta_i = 0.4$ because at this level of certainty, the inherent difficulty of creating more complex ideas can be overcome by learning, at least for a typical case. At this level of certainty, $E[\tilde{a}] = 0.75$, so that $\alpha = 0.3$ and $\beta = 0.1$. Initially, a 3-element idea with $E[a(p)] = 0.75$ for each pair has $E[e(d)] = 0.75^3 \approx 0.42$. Thus, 3-element ideas are initially less likely to succeed than 2-element ideas. However, if the researcher observes one compatibility on each pair, each pair's beta parameters are increased to $\alpha = 1.3$ and $\beta = 0.1$, so that $E[a(p)] = 1.3 / 1.4 \approx 0.93$. This increases the expected efficacy of the idea to $E[e(d)] \approx 0.93^3 \approx 0.8$, so that such ideas are more likely to succeed than a 2-element idea with $E[a(p)] = 0.75$.

This means a researcher investigating simple ideas composed of a single pair can learn enough to make a 3-element idea just as attractive as a 2-element idea with no information, in the typical case. If the certainty was much higher, learning would not convey much information and Special Case 1 (known affinity) would prevail. Nevertheless, variation in the initial draws of $\tilde{a}$ means I will observe many cases where it is not possible to learn enough to make the efficacy of a three element idea higher than a two-element one (with no information).

Next, I select the value of prizes. There is evidence, both theoretical[13] and empirical,[14] that important traits about ideas, such as their value, are Pareto distributed. Other studies, however, indicate the distribution of the value of ideas, while fat-tailed and highly skewed, is not Pareto.[15] To address both possibilities, I use two distributions for $\pi(d)$ with the same mean and variance, but where one is a Pareto distribution and the other a log-normal distribution. In practice, I find neither has a meaningful impact on the optimal strategy.

For half the cases, I draw prizes from a Pareto distribution with:

$$\begin{aligned} x_{\min} &= 1 \\ \alpha &= 2.41 \end{aligned} \tag{32}$$

These values imply the median prize has value approximately equal to 4/3, so that, on average, half of the two-element ideas will satisfy the condition $E[e(d)]\pi(d) > 1$ and therefore be myopically rational to attempt (since the average two-element idea will have $E[e(d)] = 3/4$).

For the other half of the cases, I use a log-normal distribution tuned to have the same median and variance as the Pareto distribution (so that it is primarily the behavior of the tails that differs between the distributions). This implies the log of this distribution follows a normal distribution with $\mu = 0.288$ and $\sigma^2 = 0.633$. Both distributions are plotted in Figure 2.

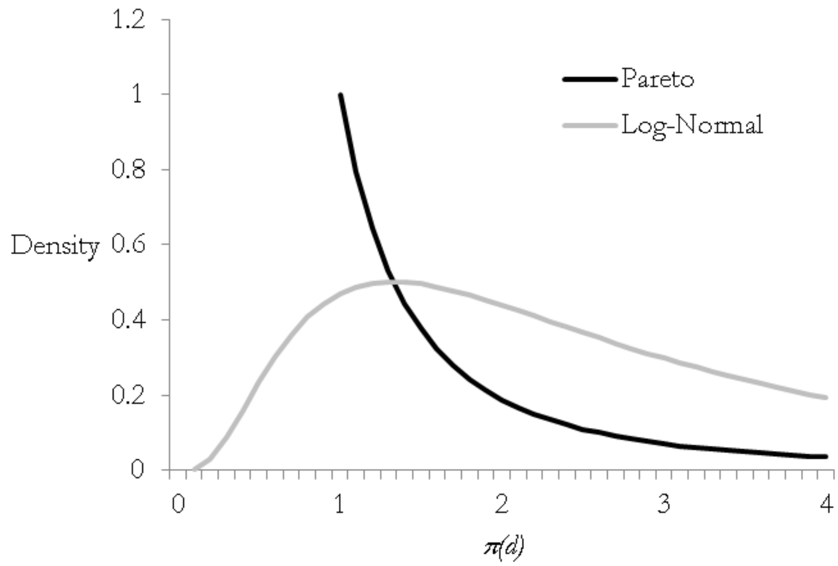
### 7.4 – Simulated Research Decisions

To illustrate features common across many variations, I simulate many actual decisions by a researcher, using the policies that emerge from the solutions to the above problems. To simulate the researcher's problem, I use the researcher's initial beliefs in each model to draw "true" affinities. With the true affinities, I generate the pair compatibilities and efficacies of each idea, as well as the information revealed to the researcher if that idea is attempted. I then have the researcher follow her strategy, observing her choice in each stage. She makes a choice, observes new information, updates her beliefs, and then follows her optimal strategy in the new information and action state. I will

[13] Jones (2005) and Kortum (1997) both show many stylized facts about aggregate R&D and growth can emerge if traits of ideas are Pareto distributed.
[14] See Jones (2005), pg. 533-535 for discussion of this evidence.
[15] See Scotchmer (2004), pg. 275-282 for a discussion.

perform 1,000 such simulations for each model, and in each simulation the researcher makes 10 choices over 10 periods (after the tenth period she always chooses to quit research).

**Figure 2:** $\pi\left(d\right)$ probability density functions




*Note: The two pdfs cross again around* $\pi\left(d\right)=29$.

**8 – Numerical Analysis of the General Case**

So far, I have made a number of remarks about the characteristics of optimal innovation behavior under simplified settings. In one instance, the learning aspect of the model was suppressed, and in the other, combinatorial features of the model were suppressed. In the general setting, optimal innovation behavior has the characteristic of each.

**8.1 – Optimal Strategy: Probability A Pair of Elements Are Combined**

**8.1.1 – Functional Forms and Explanatory Variables**

I first consider the probability a researcher will attempt to combine two elements as part of a new idea (either a 2-element or 3-element idea). My approach is to model the probability a pair is used in any given decision as a function of characteristics of the pair. In this way, I obtain a profile of the features held by the numerically-derived optimal strategies. Specifically, I run a probit regression with the following form over all the simulated researcher decisions:

$$\Pr\left(u_{p,t}=1\right)=\Phi\left(\beta_0+X'_{p,t}\beta\right) \tag{33}$$

where $u_{p,t}$ is a dummy variable equal to 1 if pair $p$ is used as part of the idea attempted in period $t$, when the researcher follows an optimal strategy. Each observation corresponds to one pair in one (simulated) researcher's decision. I omit pairs that do not belong to an eligible idea because the probability that $u_{p,t} = 1$ falls to zero in this case. The explanatory variables are traits for each pair, where I choose what traits to include based on the analysis of Remarks 1-7.

Turning first to Section 3, I showed the optimal strategy when affinities are known is straightforward: always choose the idea with the highest expected net value, so long as it exceeds 0 (Remark 1). Therefore, whether or not a pair belongs to the idea with the highest expected net value is a key explanatory variable. I capture this with the dummy variable $\text{Myopic}_{p,t}$, which is equal to 1 if pair $p \in d^*$ and $\arg\max_{d \in D} \{\pi(d) E[e(d)] - k(d)\} = d^*$ in period $t$.

Of course, when affinities are not known, the optimal strategy in Section 3 is no longer appropriate. Sometimes an idea is selected that is not the myopic best choice, but which provides useful information for future periods. However, when choosing between rival ideas that will provide equally useful information in the future, researchers will still prefer ideas with higher expected net value in the current period, using up the most highly valued ideas first. To account for the fact that the best ideas associated with a pair are fished out over time, I create two variables, $\text{Attempts}_{p,t}$ and $\text{Positive Value}_{p,t}$. The variable $\text{Attempts}_{p,t}$ counts the number of times an idea with pair $p$ has been attempted up to period $t-1$. As $\text{Attempts}_{p,t}$ rises, there ought to be fewer good ideas left that contain pair $p$. The variable $\text{Positive Value}_{p,t}$, conversely, counts the number of remaining eligible ideas that contain pair $p$ and also satisfy $\pi(d) E[e(d)] - k(d) > 0$. The intuition here is that, when the researcher deviates from a myopic strategy, she still wants to minimize her losses. One way to do this is to choose ideas that do not have the highest net expected value, but which still have *high* net expected value. Pairs that belong to many eligible ideas that will be eventually attempted under a pure fishing out strategy are more likely to have *high*, if not the *highest* net expected value.

The optimal strategy for the special case in Section 5 is less straightforward than in the Section 3 case. Researchers adopt a stick-with-the-winner strategy (Remark 7), which I capture by counting the number of times the pair has been observed compatible by the researcher up to period $t$. I denote this variable $\text{Compatible}_{p,t}$. In Section 5, I also showed that researchers prefer pairs with greater uncertainty (Remark 8). I also capture the degree of certainty about a pair with the variable $\text{Attempts}_{p,t}$ since researchers obtain better information about a pair (usually) when more ideas with it have been attempted.

Unfortunately, a Gittins index strategy does not work in the general setting. Firstly, all ideas are not equally valued, so that the payoff from any one strand of research is declining over time (because of fishing out effects), unless this is offset by the researcher's continual upward assessment of each remaining idea's efficacy. Secondly, in the general setting, all ideas with more than two elements depend on *multiple* pairs. Knowledge that a pair has a high affinity is useless if all other pairs have low affinity, since the researcher can't combine the pair with any others to generate effective ideas.

Conversely, if a pair is embedded in a network of many of other pairs with high affinity, learning it too has high affinity is very rewarding, since it can be combined with many other pairs. To measure this effect, I again use the variable $\text{Positive Value}_{p,t}$. Intuitively, any idea with $\pi(d)E[e(d)]-k(d)>0$ will be tried eventually, and so learning about pairs contained in the idea will have an impact on the value of research. Pairs with a high value of $\text{Positive Value}_{p,t}$ belong to many ideas that would benefit from learning the idea is effective.

The final regression takes the form:

$$\Pr\left(u_{p,t}=1\right)=\Phi\left(\beta_0+\beta_1\cdot\text{Myopic}_{p,t}+\beta_2\cdot\text{Attempts}_{p,t}+\beta_3\cdot\text{Positive Value}_{p,t}+\beta_4\cdot\text{Compatible}_{p,t}\right) \tag{34}$$

I anticipate the optimal strategy is characterized by the following:

**Conjecture 1: Probability of Pairwise Combination.** When the probability a researcher will optimally combine two elements as part of a research project is modeled by (34), then $\beta_1,\beta_3,\beta_4>0$ and $\beta_2<0$.

### 8.1.2 – Results

This is indeed the case, as Table 1 indicates.

**Table 1:** Probit Regression Characterizing the Optimal Strategy

| | Constant | $\text{Myopic}_{p,t}$ | $\text{Attempts}_{p,t}$ | $\text{Positive Value}_{p,t}$ | $\text{Compatible}_{p,t}$ |
|---|---|---|---|---|---|
| Coefficient | -2.096<br>(0.002) | 2.765<br>(0.003) | -0.459<br>(0.008) | 0.151<br>(0.002) | 0.428<br>(0.008) |
| Observations | 3,160,024 | | | | |
| Psuedo $R^2$ | 0.626 | | | | |
| Akaike Inf. Crit. | 1,183,283 | | | | |

*Note:* *Standard errors are reported in parentheses.*

All the signs are in the anticipated direction, and all parameters are significantly different from zero.

Note that $\text{Positive Value}_{p,t}\geq 1$ whenever $\text{Myopic}_{p,t}=1$ (otherwise quitting research would be the myopic best choice), that $\text{Positive Value}_{p,t}\leq 3-\text{Attempts}_{p,t}$ (because the maximum number of eligible ideas associated with a pair is 3), and $\text{Compatible}_{p,t}\leq\text{Attempts}_{p,t}$ (since we can only observe a pair is compatible by attempting an idea containing it). With these constraints, and because these variables are discrete over a small range, I can exhaustively list every feasible combination of pair traits, as well as the probability of selection in Table 2.

**Table 2:** Probability of Selection

| Myopic | Attempts | Positive Value | Compatibility | $\Pr\left(u_{x,t}=1\right)$ |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0.018 |
| 0 | 0 | 1 | 0 | 0.026 |
| 0 | 0 | 2 | 0 | 0.036 |
| 0 | 0 | 3 | 0 | 0.05 |
| 0 | 1 | 0 | 0 | 0.005 |
| 0 | 1 | 0 | 1 | 0.017 |
| 0 | 1 | 1 | 0 | 0.008 |
| 0 | 1 | 1 | 1 | 0.024 |
| 0 | 1 | 2 | 0 | 0.012 |
| 0 | 1 | 2 | 1 | 0.034 |
| 0 | 2 | 0 | 0 | 0.001 |
| 0 | 2 | 0 | 1 | 0.005 |
| 0 | 2 | 0 | 2 | 0.0155 |
| 0 | 2 | 1 | 0 | 0.002 |
| 0 | 2 | 1 | 1 | 0.007 |
| 0 | 2 | 1 | 2 | 0.022 |
| 1 | 0 | 1 | 0 | 0.794 |
| 1 | 0 | 2 | 0 | 0.834 |
| 1 | 0 | 3 | 0 | 0.869 |
| 1 | 1 | 1 | 0 | 0.641 |
| 1 | 1 | 1 | 1 | 0.785 |
| 1 | 1 | 2 | 0 | 0.696 |
| 1 | 1 | 2 | 1 | 0.827 |
| 1 | 2 | 1 | 0 | 0.461 |
| 1 | 2 | 1 | 1 | 0.629 |
| 1 | 2 | 1 | 2 | 0.776 |

Clearly choosing the idea with the highest net expected value is usually the preferred strategy, with the probability of selecting a pair that is the myopic best choice typically on the order of 60-85%. Even when the researcher has twice tried the idea, and never observed a compatibility, such a pair is played 46.1% of the time. Note also that the probability of use is declining in Attempts, although this can be almost perfectly offset by observing compatibilities. For instance, the probability of choosing a pair with Positive Value = 1, Myopic = 0, and Compatible = 0 falls from 2.6% to 0.2% as Attempts rises from 0 to 2, but that it only drops to 2.2% if each attempt reveals the pair to be compatible.

In the general case, the optimal strategy is a mix between the two strategies discussed in Section 3 and 5. In this model with just four elements, fishing out effects are very strong – at most, any pair only belongs to 3 eligible ideas. Nonetheless, researchers are more likely to return to a pair of elements that has been compatible in the past and they favor learning about pairs that are connected to other good ideas.

#### 8.1.3 – Knowledge Accumulation and Fishing Out

These results suggest a natural explanation for how knowledge accumulation effects and fishing out effects can coexist. Within any given set of primitive knowledge elements, the set of ideas that can be created is finite, and if knowledge is perfect, fishing out effects dominate. This is one reason why the coefficient on the number of ideas attempted using a pair of elements is negative. However, since knowledge is generally not perfect – especially at the outset – knowledge accumulation effects kick in, since learning that elements are compatible tends to expand the set of ideas that can be profitably attempted. This is why the coefficient on the number of compatible observations is positive. Thus, if the set of elements is fixed, knowledge accumulation effects can increase the value of R&D, but only up to an upper bound, given by the perfect knowledge setting. Thereafter, fishing out effects dominate. The only way out of this long-run trap is to expand the set of primitive elements, something this paper does not address (but see Weitzman 1998 for an optimistic take).

#### 8.1.4 – Path Dependence

Knowledge accumulation in this model is also related to the concept of technological path dependence. Technological path dependence refers to the idea that certain strands of technology obtain market dominance and hence the attention of future innovators. For example, Acemoglu et al. (2012) present a model where two kinds of technology – carbon neutral and carbon emitting – are substitutes, and in the absence of government policy innovators devote most of their attention to whichever technology has greater market share. Through this mechanism, the transition costs from a carbon emitting to carbon neutral production scheme rise over time, as carbon emitting technology improves at a faster rate than carbon neutral. My model exhibits a similar feature derived from learning, rather than market, effects. Since combinations that have worked in the past are more likely to work in the future, researchers optimally base subsequent research on these combinations, rather than searching for alternatives.

#### 8.1.5 – Spillovers

The above model also provides a clear mechanism for how knowledge spillovers might happen. As I have noted above, if the researcher observes some pair $p$ is compatible, this increases the expected efficacy of all other ideas that also include pair $p$. In this way, positive developments in one idea can spill over to related ideas. There is also a second channel of knowledge spillover though. The probability that pair $p$ forms part of a research project is positively related to the number of ideas containing pair $p$ with $E[e(d)]\pi(d)-k(d)\geq 0$ (captured by the explanatory variable $\text{Positive Value}$). Suppose the researcher observes pair $p'$ is also compatible. This raises the expected efficacy of all ideas that contain pair $p'$, including some ideas that *also* contain pair $p$. If the expected efficacy of such an idea rises by a sufficient amount, it may flip the expected net value of the idea from negative to positive. This increases the probability research projects containing pair $p$ will be attempted. For example, suppose researchers invent a new kind of turbine for power plants. It is known that such a turbine can be reconfigured into a jet engine. This may stimulate research on projects related to jets, but not to turbines at all. In this way, entire technological paradigms can be locked in, since the rise in $\text{Positive Value}$ can be temporarily self-reinforcing. Greater knowledge about the affinity of combinations in a subset of elements raises the value of $\text{Positive Value}$ for pairs in the subset, which in turn increases the probability of research that increases knowledge about these same pairs.

### 8.1.6 – Radical vs. Incremental Innovation

The above results can also be interpreted in terms of radical versus incremental innovation. The distinction between innovation that generates radically new types of processes and products, and innovation that makes improvements to existing technologies while leaving the basic framework unchanged has roots in economic history.[16] Examples of radical innovation might include the steam engine, electricity, and the computer, while examples of incremental innovation might be a new model of a car or smart phone. The importance of radical innovation for establishing a platform for subsequent improvement is also emphasized in the general purpose technology literature (see Helpman 1998).

In terms of this model, I identify radical innovations as those which are composed of very novel combinations of elements. Such a combination would use pairs with relatively low values of Attempts and Compatibility. In contrast, I identify incremental innovations as those which are composed of relatively common combinations of elements. Such combinations are characterized by pairs with a high degree of certainty about their true affinity. These combinations would use pairs with high values of Attempts and Compatibility.

When radical ideas turn out to be effective, the researcher's beliefs are significantly impacted, and the expected affinity of each pair rises by a comparatively large degree (since beliefs are most responsive to new information when they are characterized by a lot of uncertainty). This is more likely than an incremental innovation to flip some ideas using the same pairs from negative to positive expected value, and thereby raise Positive Value for pairs. Radical innovation provides one way to temporarily reverse the fishing out of good ideas.

At the same time, it does not follow that a dearth of radical innovation, or "moonshots," signals a poor outlook for innovation. An assortment of recent works advocate a form of "technological pessimism" (Cowan 2011). This notion is generally based on a raft of arguments, including an intuitive appeal to the reader that technology hasn't lived up to our dreams. For example, Gordon (2012) asks his readers to consider the impact on their life of losing either (1) all innovation since 2002 or (2) running water and indoor toilets, to demonstrate the paucity of recent innovation. These complaints can be read as frustration with the current incremental state of technological advance, as against the radical innovations that have occurred in the past or which were expected soon.

However, as we can see above, radical innovation (low Attempts and Compatibility) is not generally a superior strategy to incremental innovation (high Attempts and Compatibility). Indeed, it may be worse. As I have discussed above, the coefficients on Attempts and Compatibility approximately offset each other – pairs with no prior attempts are about as likely to be used as those with multiple attempts, so long as each attempt reveals the pair compatible. This approximately equal effect stems from the small number of elements in these simulations, which means each pair can only be used

[16] See Mokyr (1990), p. 291-292, and Allen (2009), p. 151-155.

three times. If I increased the number of elements available for combination, the coefficient on Attempts would be reduced relative to the coefficient on compatibility.

Radical innovations, because they rely on untried combinations, are typically riskier and do not benefit much from knowledge accumulation effects. They are always out there, as a backstop research agenda, when other avenues are fished out. Innovations that are very different from the kind around them may be more memorable than another iteration on an existing paradigm, and they may herald promising new avenues of research. But it may be a mistake to complain that there is not more interest in attempting radical innovation. When researchers are disproportionately trying for radical innovations, this is a signal that technological opportunity is low.

#### 8.1.7 – Summary

To summarize, researchers overcome the fishing out effect, temporarily, by directing their efforts towards ideas known to succeed. This "stick with the winners" approach is always exhausted in the long run, so that researchers branch out and experiment with combinations that are less studied and which are characterized by uncertainty. If one of these experiments pans out, research along similar lines can continue, as the researcher revises her beliefs about projects she previously thought to be infeasible. By inducing research on related pairs, the benefits of success can spill over past the initial research project, further sustaining the boom. If none of the experiments pan out though, researchers eventually give up on research and consider the useful ideas which can be pulled from the set of elements exhausted.

### 8.2 – Exploration vs. Exploitation Over Time

The optimal strategy in a no-learning model is myopic and purely exploitative – always choose the idea with the highest net expected value. In contrast, in the learning model, the optimal strategy mixes exploration and exploitation. The researcher is forward looking, so that the possibility of good rewards in the future (conditional on success in the present period) may lead her to attempt ideas that are not myopically optimal.

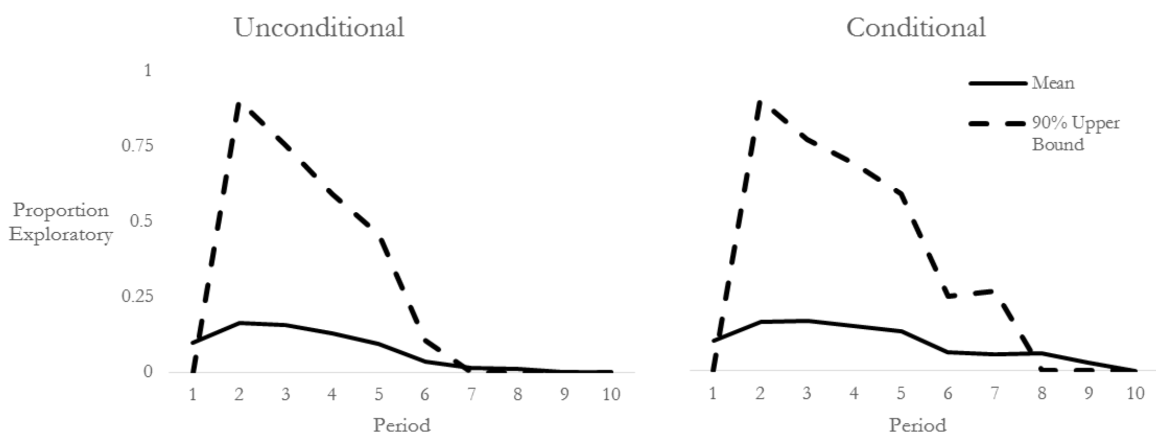
Over time, a researcher gains information and becomes more confident about the value of each pair's affinity. Concurrently, ideas are fished out and there are fewer ideas to which new information can be applied. Both factors should push researchers towards a purely myopic strategy in later periods, as the value of learning wanes.

This suggests a plausible (but incomplete) analogy between research behavior and a lifecycle model of investment. In a simple investment model, consumption is deferred in early periods to invest (for example, in human capital) until some inflection point when the investor begins to draw down their investments, so that there is nothing left after the terminal period. Just so, it might be presumed, researchers should optimally "invest" in learning in early periods, and then enjoy the fruits of their better information in later periods. This would suggest researchers will start with learning strategies and choose the myopic best choice more often in later periods.

Figure 3 indicates this view is incomplete. To generate this figure, I observe for each simulated researcher decision whether he chooses an exploratory strategy, defined as a choice that does not have the highest expected net value ($E\left[e(d)\right]\pi(d)-k(d)$). For each period, I take the fraction of times an exploratory decision is made. I then average these over all 100 variations to generate the chart below. I plot the mean fraction, as well as the 90% upper bound (this gives a boundary below which 90% of variations lie). The unconditional chart includes as observations periods after the researcher has quit research. The other case presents the average fraction of exploratory choices, conditional on the researcher not having quit research in the previous period. Both follow the same general trajectory.

Figure 3 indicates researchers are most likely to make exploratory choices in the *second* period, rather than the first. Some 10% of the time the researcher initially chooses to conduct research on an idea which is not the myopic best choice. This fractions rises to a peak of 16.4% in period 2 before declining to 0% by the final period. The 90% upper bound also follows the same trajectory, rising to a peak in period two and then falling off towards 0% (the 10% lower bounds is always 0%).

**Figure 3: Proportion of Exploratory Choices in Each Period**




*The conditional chart excludes observations after the researcher has decided to quit research. The unconditional chart does not.*

The intuition for this result is as follows. Consider the choice between an exploratory and exploitative strategy in the first period. Choosing the non-exploitative strategy means deferring a higher reward today for the prospect of an even higher reward in the future. The larger the reward offered today, the more unlikely it is the researcher will choose to defer it. And the most valuable ideas are always available in the initial period. Therefore, in the first period, it will tend to be harder for the researcher to adopt a non-exploitative strategy.

In following a myopic strategy, however, the best ideas get used up first. This lowers the penalty of deferring a reward today, and makes an exploratory strategy more attractive. Therefore, after the first few periods, an exploratory strategy is more likely to be adopted.

However, as more research is conducted, an agent becomes more certain about the affinities of different pairs. The importance of learning fades and the optimal strategy should look increasingly like the no-learning strategy again. This suggests the researcher should increasingly favor myopic strategies in later stages.

This discussion is related to the literature on basic versus applied research. While oversimplifying, basic research is associated with projects that expand our knowledge about the world, but which may be farther from commercial application, while applied research is associated with projects that attempt to commercialize well-understood phenomena. Our model is consistent with a research trajectory that begins applied, proceeds through a relatively long period where basic research is important, and then enters a mature phase where applied research dominates. In our model, we may identify what I have called exploratory research with basic research, since both endure costs in order to learn more (for future application). Myopic research may be then be identified with applied research.

In more informal terms, in the beginning the researcher faces a choice between grabbing low hanging fruit and conducting longer-term research projects. In the beginning, some ideas are so valuable it is worth pursuing them even in the absence of good information. Consider, for example, how the steam engine was developed without a theory of thermodynamics. When these valuable ideas prove successful, they teach the researcher about pair affinities. This implies later researcher may learn much from ideas undertaken for more short-sighted goals (just as physics learned a great deal from steam engines). Later, as these ideas are used up, it becomes increasingly optimal to conduct research before attempting ideas, here represented by the decision to defer the best choice from a myopic point of view in favor of an exploratory strategy. In time though, basic research is no longer necessary.

It is possible to apply this framework to an even larger canvas. Up until the late 19th century, technology proceeded without much guidance from science,[17] because there were enough low-hanging fruit about so that long-term research horizons were not optimal. Petroski (1992), for example, traces the development of a large array of everyday pieces of technology, ranging from tableware to carpentry tools to paperclips. These objects were successively improved by a long line of inventors, each of whose goal in conducting R&D was the improvement of the object at hand, rather than the discovery of knowledge to be applied in distant future contexts.

Smil (2005), however, documents a change in this paradigm during the 19th century. During this period, science increasingly became an input into technology, with electricity, internal combustion engines, chemical industries, and communication infrastructure leading the way. Indeed, Hamilton, Narin and Olivastro (1997) shows that citations to scientific papers by patents have increased significantly over the 1987-1994 period, indicating this trend remains underway in the very recent past. In terms of this model, most of the low-hanging fruit has, at long last, been exhausted, and we have entered the phase where exploratory strategies have become competitive with myopic ones.

The tail-end of this process, if it occurs, is the stuff of science fiction. Vinge (1999), for example, tells us about a future where humanity has colonized the stars, but where genuinely new knowledge is extremely rare. In this science fiction story, what we would consider R&D mostly entails the searching of enormous databases for previous discoveries that can be modified for whatever problem is at hand. This is a world where there is nothing fundamental left to learn, but myopic R&D, drawing on the vast knowledge accumulated over human history, is still practiced.

[17] See McCloskey (2010), chapter 38 for a discussion relating to the industrial revolution in Britain and Dorn and McClellan III (2006) for a broader perspective.

## 9 – Conclusions

This paper has shown how a model of innovation where researchers learn about the likelihood different pairs of technological "building blocks" work together can generate a number of stylized facts about the innovation process. Knowledge can spill over from one application to another, as researchers observe how the components that comprise each idea interact with each other. This effect also leads to a "standing on the shoulders" of the giants effect, whereby later researchers can develop technologies that would be seen as impractical and unlikely to succeed by earlier researchers. At the same time, for any fixed set of technological building blocks, the set of ideas can be exhausted. Therefore, an optimal strategy mixes exploratory research, whose benefits stem from better information for decision-making in the future, with myopic strategies. The pool of ideas must be periodically restocked by basic research, or else all the fish worth eating will be gone.

Perhaps surprisingly, it is not necessarily optimal to conduct exploratory research immediately, and then use the information gained in later periods. Instead, it is often optimal to initially grab low-hanging fruit that yield a high expected value payoff immediately. Once these are exhausted, basic research becomes a useful strategy. This prediction appears to be in line with the history of technological development writ large.

Going forward, there are several avenues of research that could build on this approach. This paper only discussed a single researcher, who was endowed with knowledge, acting in a partial equilibrium setting. Relaxing these constraints may show this approach can encompass additional facets of the research process. Moreover, this model suggests a way of empirically measuring the state of knowledge for a given technological field, so long as the fields elemental building blocks and the connections between them can be observed. See Clancy 2015 for one such application.

**References**


Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and Robert N. Stavins. 2012. "The Environment and Directed Technical Change." *American Economic Review 102(1)* 131-66.

Arthur, W. Brian. 2009. *The Nature of Technology: What It Is and How It Evolves.* New York: Free Press.

Auerswald, Philip, Stuart Kauffman, Jose Lobo, and Karl Shell. 2000. "The Production Recipes Approach to Modeling Technological Innovation: An Application to Learning By Doing." *Journal of Economic Dynamics and Control 24* 389-450.

Bellman, Richard. 1956. "A Problem in the Sequential Design of Experiments." *Sankhyã: The Indian Journal of Statistics (1933-1960), Vol. 16, No. 3/4* 221-229.

Borges, Jorge Luis. 1962. *Ficciones (English Translation by Anthony Kerrigan).* New York, NY: Grove Press.

Casella, George and Roger L. Berger. 2002. *Statistical Inference: Second Edition.* Pacific Grove, CA: Duxbury Publishing.

Clancy, Matthew. 2015. *Combinatorial Innovation: Evidence from Two Centuries of Patents.* Dissertation.

Gittins, John, Kevin Glazebrook, and Richard Weber. 2011. *Multi-Armed Bandit Allocation Indices; 2nd Edition.* Chichester, UK: John Wiley and Sons, Ltd.

Gittins, John, and You-Gan Wang. 1992. "The Learning Component of Dynamic Allocation Indices." *The Annals of Statistics, Vol. 20, No. 3* 1625-1636.

Helpman, Elhanan. 1998. *General Purpose Technologies and Economic Growth.* Cambridge, MA: MIT Press.

Jones, Charles I. 2005. "The Shape of Production Functions and the Direction of Technical Change." *Quarterly Journal of Economics 120 (2)* 517-549.

Jovanovic, Boyan and Rafael Rob. 1990. "Long Waves and Short Waves: Growth Through Intensive and Extensive Search." *Econometrica* 1391-1409.

Kauffman, Stuart, Jose Lobo, and William G. Macready. 2000. "Optimal Search on a Technology Landscape." *Journal of Economic Behaviour and Organization 43* 141-166.

Kortum, Samuel S. 1997. "Research, Patenting and Technological Change." *Econometrica* 1389-1419.

McClellan, James E. and Harold Dorn. 1999. *Science and Technology in World History: An Introduction.* Baltimore, MD: The John Hopkins University Press.

McCloskey, Deirdre N. 2010. *Bourgeouis Dignity: Why Economics Can't Explain the Modern World.* Chicago, IL: The University of Chicago Press.

Narin, Francis, Kimberly S. Hamilton, and Dominic Olivastro. 1997. "The Increasing Linkage Between U.S. Technology and Public Science." *Research Policy* 26: 317-330.

Petroski, Henry. 1992. *The Evolution of Useful Things: How Everyday Artifacts-From Forks and Pins to Paper Clips and Zippers-Came to be as They are.* New York, NY: A. Knopf.

Poincaré, Henri. 1910. "Mathematical Creation." *The Monist* 321-335.

Powell, Warren B. 2011. *Approximate Dynamic Programming.* Hoboken, NJ: John Wily & Sons.

Ridley, Matt. 2010. *The Rational Optimist.* New York, NY: HarperCollins Publishers.

Scotchmer, Suzanne. 2004. *Innovation and Incentives.* Cambridge, MA: MIT Press.

Smil, Vaclav. 2005. *Creating the Twentieth Century.* Oxford, UK: Oxford University Press.

Vinge, Vernor. 1999. *A Deepness in the Sky.* New York, NY: Tor Publishing.

Weber, Richard. 1992. "On the Gittins Index for Multiarmed Bandits." *The Annals of Applied Probability 2(4)* 1024-1033.

Weitzman, Martin L. 1998. "Recombinant Growth." *Quarterly Journal of Economics 2* 331-360.

**Appendix A1 – Gittins Index**

The following proof is based on Weber (1992). Suppose we are in the setting described in Section 5 of this paper.

**Optimal strategy with learning:** The optimal strategy in every period is to choose the option with the highest Gittins Index $\lambda_i(\alpha_i,\beta_i)$ where

$$\lambda_i(\alpha_i,\beta_i)=\sup\{\lambda_i : v_i(\alpha_i,\beta_i,\lambda_i)=0\} \tag{35}$$

and $v_i(\alpha_i,\beta_i,\lambda_i)$ given by:

$$\max\left\{\frac{\alpha_i}{\alpha_i+\beta_i}\left(1+\delta v_i(\alpha_i+1,\beta_i,\lambda_i)\right)+\frac{\beta_i}{\alpha_i+\beta_i}\delta v_i(\alpha_i,\beta_i+1,\lambda_i)-k-\lambda_i\ ,\ 0\right\} \tag{36}$$

**Proof:**

1 – A Single Pair

Suppose pair $p_i$ is the only pair available to choose, so that the researcher's problem collapses to the choice between conducting a research project that includes pair $p_i$ or to quit research. Thus, the researcher's problem can be written as a Bellman equation of the form:

$$v_i(\alpha_i,\beta_i)=\max\left\{\frac{\alpha_i}{\alpha_i+\beta_i}\left(1+\delta v(\alpha_i+1,\beta_i)\right)+\frac{\beta_i}{\alpha_i+\beta_i}\delta v(\alpha_i,\beta_i+1)-k,0\right\} \tag{37}$$

Next, suppose there is an additional charge to conduct research, which I will call the *prevailing charge*, denoted $\lambda_i$. The prevailing charge is the same each time the researcher chooses to conduct research on pair $p_i$. The researcher's problem can then be written as a Bellman equation of the form:

$$v_i(\alpha_i,\beta_i,\lambda_i)=\max\left\{\frac{\alpha_i}{\alpha_i+\beta_i}\left(1+\delta v_i(\alpha_i+1,\beta_i,\lambda_i)\right)+\frac{\beta_i}{\alpha_i+\beta_i}\delta v_i(\alpha_i,\beta_i+1,\lambda_i)-k-\lambda_i,0\right\} \tag{38}$$

Define the *fair charge* $\lambda_i(\alpha_i,\beta_i)$ as the maximum prevailing charge selected so that, in expectation, the optimal strategy makes zero profit:

$$\lambda_i(\alpha_i,\beta_i)=\sup\{\lambda_i : v_i(\alpha_i,\beta_i,\lambda_i)=0\} \tag{39}$$

Suppose the researcher is facing a fair charge, so that she is indifferent between conducting research and quitting, since both earn expected profit of zero. If the researcher decides to conduct research, she observes the compatibility of pair $p_i$. If she observe pair $p_i$ to be compatible, her expected profit will be positive going forward, since the charge is fixed but the expected probability of winning a reward in each period is increased. If she observes pair $p_i$ to be incompatible, the prevailing charge will be too high in the next state, so that the researcher would prefer to quit research, earning zero profit in expectation.

Now suppose the prevailing charge is always *lowered* to the fair charge rate, whenever the researcher finds herself in a position where it would be optimal to quit research. This does not affect the researcher's expected profit, since she expects to earn zero under a fair charge, but would have earned zero anyway by quitting research. If the prevailing charge is always reduced in this way, so as to always keep the researcher indifferent when she would otherwise prefer to quit, then the researcher need never stop conducting research. Her expected lifetime profit from such a strategy is zero.

This procedure for reducing the prevailing charge generates a stochastic sequence $\{\lambda_{i,n}\}_{n=0}^{\infty}$ which is nonincreasing in the number of times $n$ the researcher chooses to conduct research.

2 – Many Pairs

Suppose now that there are many pairs available for research, each of which has its own prevailing charge that is periodically reduced in the manner discussed above. Suppose the researcher adopts the following strategy:

> **Gittins Strategy:** In every period, choose the pair with the highest prevailing charge.

The Gittins strategy insures a pair is chosen in every period (since prevailing charges are always lowered when the researcher would otherwise quit research). Such a strategy has zero expected profit. Moreover, there can be *no* strategy that yields strictly positive profit in expectation, since this would require strictly positive profit for at least one pair.

Next, note that the sequences $\{\lambda_{i,n}\}_{n=0}^{\infty}$ associated with each pair are independent of the strategy chosen, since they depend only on the number of times $n$ a pair has been chosen. The Gittins strategy interleaves the many pair sequences into a single nonincreasing sequence of prevailing charges that maximizes the expected present discounted cost of prevailing charge paid. However, this strategy also yields the maximum expected profit of zero, which means the expected present discounted value of net rewards (absent prevailing charges) must exactly equal the expected cost of charges. Since cost was maximized, this strategy also maximizes rewards, and is therefore an optimal policy.

Since the prevailing charge is periodically lowered to the fair charge, and the fair charge only depends on the state of one pair, an equivalent strategy is to always choose the pair with the highest fair charge, which is given by equation (39).

## Appendix A2 – Program Details

This section gives some more details on how I solve the general problem presented in Section 7. The program is in the Python language. For clarity of exposition, I will refer to ideas and pairs by the elements $q \in Q$ that ultimately comprise each. To begin, I define the possible sets $D$ in which the researcher may find himself. Since there are 10 ideas, and each idea may be either eligible or ineligible, there are $2^{10} = 1,024$ distinct sets of eligible ideas. For each of these sets, I next define the potential belief vectors of interest. Since I know the initial beta parameters of every $a(p)$, the program can exhaustively list the vectors $B$ that might be attained in any given set.

For example, suppose we are considering the following set

$$D = \{(q_1, q_2, q_3), (q_1, q_2, q_4), (q_1, q_3, q_4), (q_2, q_3, q_4)\} \tag{40}$$

In this set, all of the ideas composed of three elements are eligible, but all of the ideas comprised of two elements are ineligible. In the model, the only way ideas can become ineligible is if they are tried. Therefore, this state can only be arrived at by the researcher after she has conducted research projects on all the two-element ideas. Specifically, we know the researcher has attempted:

$$attempted = \{(q_1, q_2), (q_1, q_3), (q_1, q_4), (q_2, q_3), (q_2, q_4), (q_3, q_4)\} \tag{41}$$

These are the six ideas composed of two elements. Each of these research projects yields information. For each pair, the researcher now has an observation of either one compatibility, or one incompatibility. Therefore, the beta parameters of each pair can take on one of two states: $(\alpha_i + 1, \beta_i)$ or $(\alpha_i, \beta_i + 1)$. Since there are six pairs, and each can take on two states, there are $2^6 = 64$ potential belief vectors associated with the set of eligible ideas.

Often, I can simplify matters by ignoring some pairs. For example, suppose we are considering the following set

$$D = \{(q_1, q_2)\} \tag{42}$$

In this set, only one idea is eligible – all other ideas have already been attempted. This implies a large number of potential belief vectors. For instance, once every idea has been attempted, any given pair can take on 9 states,[18] implying potentially millions of different $B$ vectors. However, most of this information is irrelevant in this case. The only parameters I care about are the ones that describe pairs in eligible ideas. In this case, there is just one pair left in an eligible idea, so I do not have to compute the millions of different $B$ vectors that apply to irrelevant pairs.

[18] For each pair there are two ideas with three elements containing the pair, each of which may reveal compatible, incompatible, or nothing, and one idea with two elements which may reveal compatible or incompatible. Thus, the potential parameter values are:
$(\alpha+3, \beta), (\alpha+2, \beta), (\alpha+2, \beta+1), (\alpha+1, \beta), (\alpha+1, \beta+1), (\alpha+1, \beta+2), (\alpha, \beta+1), (\alpha, \beta+2), (\alpha, \beta+3)$

Once it has a set of $(D, B)$ states, the program works backwards. It begins by evaluating the null set $D = \{\varnothing\}$, where the only available option is to quit research and earn zero with certainty (for every belief vector $B$). Next, using this result, the program evaluates the best action for each set with just one eligible idea remaining. When there is just one eligible idea, the problem simplifies to:

$$V(D,B) = \max\left[E\left[e(d)\right]\pi(d) - k(d), 0\right] \tag{43}$$

Using these results, the program evaluates the best action for each set with two eligible ideas remaining, which has the form of equation (28). At each stage, it uses the researcher's beliefs to compute the probabilities associated with each state the researcher may find herself in next period. For example, suppose the researcher has:

$$\begin{aligned} &D = \{(q_1, q_2), (q_1, q_2, q_3)\} \\ &B = \left[(0.2, 0.2), (1.2, 0.2), (1.2, 0.2), \ldots\right] \end{aligned} \tag{44}$$

Where the belief parameters apply to pairs $(q_1, q_2)$, $(q_1, q_3)$ and $(q_2, q_3)$ respectively.

If the researcher chooses research project $(q_1, q_2, q_3)$, then her possible outcomes are:

**Table A1: Choosing $(q_1, q_2, q_3)$**

| effective? | $V(D,B)$ | $B' = B + \omega(d)$ | Probability |
|---|---|---|---|
| yes | $\pi(d) + \delta V(D,B') - k(d)$ | $[(1.2,0.2),(2.2,0.2),(2.2,0.2),\ldots]$ | 0.37 |
| no | $\delta V(D,B') - k(d)$ | $[(0.2,0.2),(1.2,0.2),(1.2,1.2),\ldots]$ | 0.05 |
| no | $\delta V(D,B') - k(d)$ | $[(0.2,0.2),(1.2,1.2),(1.2,1.2),\ldots]$ | 0.05 |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| no | $\delta V(D,B') - k(d)$ | $[(1.2,0.2),(1.2,1.2),(2.2,0.2),\ldots]$ | 0.02 |

In fact, there are 12 potential updated belief vectors that may be attained if the idea is ineffective, reflecting the many different ways an idea can be ineffective (compared to the single way it can be effective). To see where these probabilities come from, consider first the probability the idea is effective, given by the first row of Table A1. Since:

$$\mathrm{E}\left[\Pr(e(d) = 1)\right] = \prod_{p \in d} E\left[a(p)\right] \tag{45}$$

And since

$$E\left[a(p_i)\right] = \frac{\alpha_i}{\alpha_i + \beta_i} \tag{46}$$

The probability $(q_1, q_2, q_3)$ is effective is

$$E[e(d)] = \frac{0.2}{0.4}\left(\frac{1.2}{1.4}\right)^2 \approx 0.37 \tag{47}$$

When the idea is effective, each pair is compatible, and so the belief vector next period is given by $[(1.2,0.2),(2.2,0.2),(2.2,0.2),\ldots]$, where I have added 1 to the $\alpha$ parameter of each pair in $(q_1, q_2, q_3)$.

If the idea is ineffective, computing the probability of $B'$ is more involved. Consider the second row, where

$$B' = [(0.2,0.2),(1.2,0.2),(1.2,1.2),\ldots] \tag{48}$$

Equation (48) indicates the researcher observed pair $(q_2, q_3)$ to be incompatible, but did not observe any other pairs. The probability of such a revelation requires using the revelation procedure outlined on page 17, and is the joint product of (1) selecting pair $(q_2, q_3)$, which occurs with 1/3 probability, and (2) finding the pair is incompatible, which occurs with probability $0.2/1.4$. Since $1/3 \cdot 0.2/1.4 \approx 0.05$, the researcher attaches probability $0.05$ to this outcome.

Alternatively, consider the final row, where

$$B' = [(1.2,0.2),(1.2,1.2),(2.2,0.2),\ldots] \tag{49}$$

Equation (49) indicates the researcher observed pairs $(q_1, q_2)$ and $(q_2, q_3)$ to be compatible, and pair $(q_1, q_3)$ to be incompatible. There are two ways the revelation procedure could have generated this particular set of observations.

1. It could have (1) drawn pair $(q_1, q_2)$ and found it to be compatible (probability of being drawn is 1/3, probability of being compatible is 1/2), (2) drawn pair $(q_2, q_3)$ and found it to be compatible (probability of being drawn is ½ and probability of being compatible is 1.2/1.4) and (3) drawn pair $(q_1, q_3)$ and found it to be compatible (probability of being drawn is 1 and probability of being incompatible is 0.2/1.4). The joint probability of this sequence is approximately 0.01.
2. It could have (1) drawn pair $(q_2, q_3)$ and found it to be compatible (probability of being drawn is 1/3, probability of being compatible is 1.2/1.4), (2) drawn pair $(q_1, q_2)$ and found it to be compatible (probability of being drawn is ½, probability of being compatible is ½) and (3) drawn pair $(q_1, q_3)$ and found it to be compatible (probability of being drawn is 1 and probability of being incompatible is 0.2/1.4). The joint probability of this sequence is approximately 0.01.

Taken together, the probability of observing this set of observations is approximately 0.02. Similar calculations are performed for each state.

Conversely, if the researcher chooses research project $(q_1, q_2)$, then her possible outcomes are:

**Table A2: Choosing $(q_1, q_2)$**

| effective? | $V(D,B)$ | $B' = B + \omega(d)$ | Probability |
|---|---|---|---|
| yes | $\pi(d) + \delta V(D,B') - k(d)$ | $[(1.2,0.2),(1.2,0.2),(1.2,0.2),\ldots]$ | $\frac{0.2}{0.2+0.2} = 0.5$ |
| no | $\delta V(D,B') - k(d)$ | $[(0.2,1.2),(1.2,0.2),(1.2,0.2),\ldots]$ | $\frac{0.2}{0.2+0.2} = 0.5$ |

In this case, because there is just the one pair, the researcher observes either the pair is compatible (with probability ½) or that it is incompatible (also with probability ½).

After working backwards, I have a mapping from every $(D, B)$ state to a best action. With four elements and ten ideas, this program still takes approximately an hour to solve depending on computer processing power.