Transcription of a talk by Eric Lander
National Math Festival. Gala dinner at the Library of Congress, 4/16/2015

(Introduction by David Eisenbud)

First, thanks to Jim Simons, to David Eisenbud, to Robbert Dijkgraaf, to Bob Tjian, and everyone who put together this first-ever National Math Festival. It's about time!

My assignment, from David, was to make the case for basic research. Now, looking around the room, I have to confess, I'm not sure to whom I need to make the case. But that was the assignment. And so, I'm going to do it. Because the truth is, we ***all*** have to be able to make the case for basic research. It is very important for us to be able to articulate the case for basic research.

We have in this room people who have been tremendous forces for the importance of basic research. I want to thank former Representative Bart Gordon, who was Chairman of the House Committee on Science and Technology from 2007 until 2011, and former Representative Rush Holt, who is a research physicist and former professor and current CEO of the American Association for the Advancement of Science. We have Representative Jerry McNerney, who has the distinction of being the only mathematician serving in Congress. We had, until a moment ago, Leader Nancy Pelosi, who has also been a great supporter of research. So, even within the Congress, I preach to the converted and, indeed, from some of them, I have learned much about how to make the case for basic research. But, let's get down to it.

My title is, The Miracle Machine. I'll come back to what that means in a moment.

Because we are in Washington, DC, though, I want to start by talking about responsible public policy. As David says, I co-chair the President's Council of Advisors on Science and Technology for the White House, and I've come, over the six years I've been doing that, to appreciate the importance of responsible public policy. I've been aware that, since 2008, we've been living through a great recession. We're now climbing out, but it's been slow. As you heard from President of the European Central Bank Mario Draghi, Europe is climbing out much more slowly, if at all. Budgets are tight everywhere. We have not the surplus of the late 1990s, but deficits. Understandably, our elected officials want to ensure that with respect to public spending, we're getting the maximum bang for the buck – or euro, as the case may be. Before committing money, they want to see clear objectives, short-term outcomes, and avoidance of waste. No line item in the national budgets has escaped this scrutiny. And that includes research.

Looking to the north, Canada, in 2013, decided to radically shift the focus of its National Research Council. It announced the transformation of a 98-year-old agency, which had been the leading force for basic research in Canada, into essentially a one-stop concierge service to bolster technological innovation by industry and generate jobs. To increasingly focus the Canadian National Research Council on industry issues, they describe themselves now as a strategic R&D organization for industry that aims to be directly responsive to those short-term goals. I can understand why a government might choose to do that.

In Europe, we've seen similar things. In 2011, Ireland's government substantially changed how it thinks about funding science – to focus predominantly on research with clear potential to grow the economy and create jobs. The European Commission recently announced a plan to cut nearly 3 billion euros from the Horizon 2020 program – which includes funding for the European Research Council, which is the premier, frontier research organization directed at fundamental research – in order to move money into a new Strategic Investment Fund for Europe, a short-term focused stimulus program. They plan to cut about 5% of the European Research Council's budget.

Here in the United States, we are not immune from this kind of thinking. There are bills in Congress to ensure the accountability of taxpayer dollars invested in science, to ensure that dollars are stretched efficiently and effectively – I quote from websites here – to ensure "accountability" and "transparency". There are proposals that the National Science Foundation be required to publish a justification of each and

every grant's scientific merits and relevance to the broad national interest – that is, to the economy or defense. Now, how could you possible disagree with such prudent investing of American dollars? How could you possibly disagree with accountability, transparency? If I'm investing in real estate, I want to see a clear path to a return on my investment. If I'm investing in a start-up company, I might be willing to wait for a few years before I get my return, but I want to see a clear business plan. If investing in building roads and bridges for a country, I'd like a clear connection between social investment and social return. So what can possibly be wrong with wanting to have a clear case for precisely how investments in basic research will pay off?

The answer is: absolutely everything! Everything is wrong with it. Applying this kind of filter to basic research is a terrible strategy. It'll guarantee that you will have ordinary returns – projects that pay off, at an ordinary rate. In the short term, you'll get outputs. But you will miss the extraordinary returns. Fundamental research is fundamentally different than any other kind of investment. We're all still struggling to understand it, but it is a remarkable thing: because time and time again, we've found that basic research can pay huge, out-of-proportion returns. Basic research can completely transform for the better our society, our economy, our defense.

From a mathematical point of view, you'd say that the distribution of returns from basic research has a very fat tail. That is, you can get 10:1 returns, and 100:1 returns, and 1000:1 returns, and 10,000:1 returns out of that tail. It doesn't fall off like an exponential; it has very fat tails. And, indeed, as the mathematicians here know, there are even distributions with fat tails where the *expected* return is infinite. Basic research is probably closer to that than to some normally distributed return for typical kinds of investments.

The problem is that transformative ideas and discoveries often come out of left field. Try as you might, you cannot predict where they come from. We're dealing with something pretty remarkable here. I've started to call this thing the Miracle Machine. It is miraculous, and it is a machine, because it's quite reproducible; it works again and again. More and more, I want to understand how this Miracle Machine works. It is a challenge to explain it to people. Maybe the best way to explain it is through example. So I'm going to pick some examples, and tell you about them. Some of them you'll all know, because it's a highly mathematical crowd. Some of them you're not going to know very much about, because I'll pick some biological examples; they'll be kind of fun. These examples are at the heart of the case we need to make.

So, let's start with math. As you know, I am by training a pure mathematician: a Princeton undergraduate, Oxford graduate student. I studied algebraic number theory and topology at Princeton, and then wrote my Ph.D. thesis at Oxford on group representation theory and algebraic coding theory. What got me most excited about math first in high school and in college was number theory. I love number theory – how can you not love number theory at that age?

In high school, I read and loved G.H. Hardy's Introduction to the Theory of Numbers. The book conveyed the essential beauty of number theory, especially of prime numbers. Hardy loved number theory, precisely because it was both beautiful and completely useless. In his famous essay, "A Mathematician's Apology", Hardy likened pure mathematics to painting and poetry; he was proud that it had no practical applications. In the essay, he wrote that no one had found any practical applications – he actually said, any "war-like" applications; that was what was on his mind at the time – to be served by number theory or by general relativity. And, he said it seemed unlikely that anyone will do so for many, many years. Well, Hardy must be turning in his grave. Because number theory – prime numbers – lie at the heart of national security and communications. Public-key cryptography is fundamentally based on the question of decomposing numbers into their prime factors. So, number theory is central to commerce and defense.

The other thing that Hardy said had no foreseeable application was relativity. [Laughter]

Relativity itself has its roots in mathematics. You know, if you were explaining to, let's say, a Congressional committee in the 1800s why you were funding a grant on non-Euclidean geometry, you'd be saying: yes, yes, Mr. Congressman, I realize that for 2000 years we've known that parallel lines never meet, but let's just imagine that what we know isn't true, and let's fund someone to study what would happen in a counterfactual world where all lines intersected and there were no parallel lines.

You could imagine that this would not be looked on with great favor. It would be viewed as some kind of mind game, because it didn't describe the real world, and it therefore wasn't practical. Except, of course, it turns out that it *does* describe the real world. It just was that the math was ahead of our understanding of the real world – to the curved space-time of general relativity.

Now, to return Hardy's question, is relativity practical? Well, anybody who used their iPhone to get here tonight was using GPS, and anybody who's using GPS is using general relativity, because you actually need to correct for the time dilation of the satellites in order to get accurate GPS positioning. Your iPhone has general relativity built into it. Hardy was a great number theorist, but, I think, a very bad predictor of the ultimate impact of mathematics.

Let me switch to physics. When Charles Townes invented lasers, he was advised by many people that it was a solution looking for a problem. What were you going to do with these lasers? It was a curiosity. People had masers and he wanted to know whether he could do a similar thing at the frequency of light, I don't think I would have wanted to defend its utility in advance based on that rationale. But lasers did turn out to be useful – for cutting, welding, printing, CDs, bar codes, scanners, treating acne, treating kidney stones, eye surgery, dentistry, fingerprint analysis, holograms, and laser light shows.

What about huge particle accelerators? Those giant machines built to study obscure sub-atomic particles that you will never meet in your everyday life: quarks and leptons, and things like that. What excuse do we have for investing public money to build large particle accelerators to pursue this curiosity at this scale?

Well, it kind of turns out that, when you build those things, it ends up giving us the technology to make synchrotron light sources that are used for x-ray crystallography, which used to study the structures of human proteins and which are central to all drug development efforts. Any good drug-development effort needs to have the structures of protein targets, and they have solved these structures using synchrotron light sources. It's very hard to predict that things like quark investigations will be useless – because they have this uncanny way of turning out, just when you'd least expect it, to be enormously useful.

And then there's Andre Geim, the physicist who had Friday sessions in the lab where they deliberately worked on wacky things. They studied how geckos can climb up and stick to very smooth surfaces; it might not sound very important. But it led to the development of a super-sticky adhesive able to stick to the smoothest surfaces.

He was also very interested in magnetic levitation and famously developed in his lab a system to levitate a frog. For this he won, in my fair city of Cambridge, Mass., an Ig Nobel prize for frog levitation. The Ig Nobels, if you've not been to the ceremony (I've actually been a speaker at the Ig Nobels) are a parody of the Nobel Prize.

Geim, also in his lab, used Scotch tape to peel off very thin layers of graphite, then used tape on those layers to peel off thinner layers, and then thinner layers and thinner layers – until he got to layers of graphite that were just one atom thick. That is to say, *graphene* – a single monolayer of carbon atoms. For

this work, he received not an Ig Nobel prize, but a Nobel Prize. Graphene is the thinnest material, incredibly strong, incredibly transparent, totally impervious, you can't even get helium through it. The same curiosity that levitated the frog produced the graphene. It's very hard to figure out in advance which ideas will turn into floating frogs and which will turn into graphene.

Now these stories are just not a fluke. In the United States, we have this thing that I am increasingly thinking of as the Miracle Machine: a reproducible system for making miracles come to pass, for making dreams into reality.

Vannevar Bush, the famous dean of engineering at MIT who directed research for the United States government during World War II, wrote a famous essay after the war about an imaginary device that he called a Memex. The essay was about the idea of a device that would give you access to all human knowledge. He imagined it with lots of microfilm and ways of accessing all that microfilm, whereby people could trace paths through all of human knowledge – a very appropriate image for the building in which we sit today, the Library of Congress. The Memex would contain all of the Library of Congress, all of human knowledge. It was…kind of crazy in 1945 to dream about such a thing. Well, along the way, as more people thought about such ideas, computers were developed. The field of computer science, built very much on mathematics, was developed. Hundreds of playful projects and algorithms, all sorts of things; the Internet was born. All of these things flowed one after another, after another.

In 1993, the National Science Foundation gave a 3.6 million-dollar grant to some investigators at Stanford. According to the abstract of the grant application, the purpose was to work on an integrated virtual library that will provide uniform access to the large number of emerging networked information sources and collections that will link everything to personal information collections to collections found today in conventional libraries to large data collection. Two students working on this . . . you know where this is going . . . were Sergey Brin and Larry Page. This work gave rise to Google.

On my iPhone, I just checked the market cap of Google today. It is 360 billion dollars, which is a very, very precise 100,000-fold return on NSF's investment in a 3.6 million dollar grant. One hundred thousand-fold is a good return, if you can get it. Of course, the problem is, you don't know in advance precisely ***where*** to get it.

Now, to get Google took more than just that public investment. You had a public investment from the NSF in basic science, and then, as these ideas proved more and more practical a few years later, the other part of the great Miracle Machine swung into action – private investment, from venture capitalists, from capital markets, to launch and sustain businesses, like Google. What we have in this country is an amazing partnership between a public good – publicly-funded, basic research, where we don't know precisely what it will deliver, or on what time frame – and then private investment. It's a fabulous partnership. Now, America is not the only country that does this. The Miracle Machine is not unique to the United States, but nowhere is it operating better than in the United States. And no place has reaped greater return from this Miracle Machine. So, it's important to understand how the pieces work together.

What about biology? While I am trained as an algebraic combinatorialist, I have migrated over the years into biology (although I still use mathematics throughout biology). Biology is very fertile territory for some crazy-sounding studies. Let me tell you about some of the truly wacky things that have been done in biology.

Biologists went and used federal dollars to go to Yellowstone to look at bacteria that grow in hot springs. Pretty obscure. Who cares? Sounds like just butterfly collecting. Why hot springs…don't we have enough bacteria everywhere else? Well, it turns out that bacteria that grow in hot springs are remarkable, because all their enzymes are stable against extremely high temperatures. And a particular enzyme in these

bacteria, a DNA polymerase, turned out to be a crucial component in the DNA amplification technology called polymerase chain reaction, PCR, that is used millions of times around the world every day. We have PCR in its current working form because somebody *did* go to the hot springs, although they didn't go for the purpose of enabling PCR.

Biologists also wondered why jellyfish glow green. It's a cool question, but why should you waste public funding figuring out why they glow green? Well, it turns out the answer is a certain protein, now called Green Fluorescent Protein (GFP). It can be attached, via genetic engineering, to other proteins that will now be marked by glowing green. With some modifications, you can also change the protein to glow red, and yellow, and other colors. And, with these tutti-frutti tags on proteins, you can study the internal architecture of cells in exquisite detail. It turns out that wondering why jellyfish glow green gave rise to tools that are used in tens of thousands of labs around the world and was worth a Nobel Prize.

Here's another one: there are weird bacteria that can grow in amazingly salty conditions, such as in incredibly salty lakes in Egypt and Sudan. Scientists were curious about how bacteria can grow in five-molar salt solutions. Somebody decided to investigate this seemingly ridiculous question, and figured out that they have a light-driven ion pump that uses sunlight to transport ions. Now, this was just a curiosity, until some investigators realized that if you put light-driven ion pumps into neurons, you could get the neurons to fire whenever you shined a light on them. This gave rise to optogenetics – the optical control of neurons – which is now used all over neurobiology. By the way, one of the applications that people are working on is inserting those optically-responsive ion channels into the retina to restore the sight of people with certain kinds of blindness. I could go on and on with examples from bacteria.

But, let's move from bacteria to fruit flies. Politicians love to beat up on fruit fly research. Geneticists collect strange mutant fruit flies with weird defects; there's one called hedgehog, and it's got all these little funny spikes on it. It sounds like a silly thing to do – except that the FDA just approved a hedgehog inhibitor in humans that treats basal cell carcinoma. It turns out that the same pathway the fruit fly uses for early development is a pathway operative in your skin cells. Evolution has conserved it over the course of hundreds of millions of years.

Now, some politicians, including Republicans and Democrats over time, have sought to sound reasonable by ridiculing federal grants that sound silly – at least if you only read their abstracts. William Proxmire, a Democrat of Wisconsin, famously gave the Golden Fleece awards in the 1970s for projects that he thought were frivolous, a waste of the taxpayers' money. In 1975, he gave the Golden Fleece to a study of the sex life of screwworms. You couldn't pick a better title than that: the sex life of screwworms. The US Department of Agriculture gave a quarter of a million dollars to study the sex life of screwworms! Ridiculous! Except it turns out, the sex life of screwworms is pretty important. The screwworms are parasites. And scientists found out that females lay their eggs immediately after the first time they mate with a male. And it occurred to them: if one were to release a very large number of sterile males, they would overwhelm the fertile males and mate with the females, who would lay eggs that weren't fertilized, and one would be able to control the pest. This strategy was implemented and it has saved about twenty billion dollars in livestock costs and reduced the cost of beef by five percent. To his enormous credit, Senator Proxmire apologized, acknowledging that he blew that one. It tells you that you ought to read more than the abstract.

Now, not all basic research in life sciences is bizarre-sounding or curiosity-driven. Sometimes critical research is about collecting lots of information, because you have a hunch that it's going be useful, but you don't exactly know how.

To explain this, let me turn to another great example of the Miracle Machine. In 1953, Crick and Watson discovered the double helical structure of DNA and inferred how life encoded information in these two

strands of DNA – with a sequence of DNA bases on one strand and a complementary sequence on the other strand. At the time, they had no idea that you would ever be able to read those sequences. I know, because I asked each of them. They said, not a chance, we never imagined that we'd be able, in our lifetime, be able to read out all that information.

But the information was there, and maybe you could find a way to read it. So the Miracle Machine got to work, with tens of scientists, and then hundreds of scientists, and then thousands of scientists, chipping away at bits of this really great puzzle. And they began to work out how DNA is copied in cells. And how it's turned into RNA, and it's interpreted to make proteins, and then how we might be able to take some of those tricks and do it in the laboratory to actually clone DNA and sequence DNA. First a few letters, then hundreds of letters, then thousands of letters, then millions of letters and then this idea of a Human Genome Project, which started about 1985 and was, at first, considered pretty nuts. Some thought it was a total waste of money, because it was going to cost billions of dollars. But the scientific community debated it, and the Congress decided to fund it. By 1990, there was work underway. And by April 25, 2003, we had a finished sequence of the human genome. April 25, 2003 – we actually chose that date on purpose. It was the 50th anniversary to the day of the publication of the Watson-Crick paper. That's what you can do in a half a century: you can go from a crazy idea that you might be able to somehow read out all that information to actually accomplishing it.

So, what did we get out of it? Well, we got many of the things we expected. But, the most interesting things we got were the things we *hadn't* expected. The Human Genome turned out to be an amazing font of information for understanding the genetic basis of disease. For example, this past week, I was talking to a colleague who has been working out the genetics of early heart disease. We all know that lipids play an important role in heart disease – for example, LDL, the bad cholesterol, and HDL, the good cholesterol. But, he's using the human genome to discover entirely *new* pathways – for example, a pathway that has nothing to do with lipids at all; it looks like that's going to be very important. Similar things are happening for diabetes. And, for schizophrenia, genetics is starting to tell us that the disease is being driven by excessive pruning of neuronal outgrowth in adolescence. For autism, we have a great collaboration with Jim Simons that is helping to unravel the biological basis of autism. For cancer, the genome is revealing a tremendous amount information about the cellular basis of the disease, and has led to hundreds of drugs that are in clinical development.

Beyond these applications to disease, there are many amazing fundamental discoveries about how genes are regulated, about how DNA is folded in the nucleus of the cell, about evolution, and about human populations and how they spread about the world. And, a complete surprise, about how humans interbred with Neanderthals – which we can read out from the DNA, using a heavy dose of mathematical analysis.

In fact, all of this discovery has depended on a heavy dose of mathematical analysis. For example, suppose we scan the human genome to find genes for heart disease or schizophrenia. How do we know if a correlation is significant? If you're going to do only one test, you can use the classic significance level of $p = 0.05$. But, what if you're looking across an entire continuous genome of three billion bases? What's a significant result? Only math can tell us. It turns out you need to know the extreme valuable distribution of certain stochastic process called an Ornstein Uhlenbeck diffusion. In simpler terms, it turns out that figuring out whether a genetic finding is significant is equivalent to considering the behavior of a particle undergoing Brownian motion while coupled to the origin by a Hookean spring and asking how far it can get from the origin in a certain amount of time. This is just one of many unexpected connections between math and biology.

Here's another: We study how the genome folds up in a cell. We recently studied a paper showing that – to a first-order approximation – the genome folds up as a fractal. Actually, we now have enough data to say that it's not quite a fractal. It's almost a fractal and we can now see how it differs from a fractal. We

learn this by combining molecular biology and mathematics to study the distribution of distances between points in the genome.

So, what's been the economic impact of the Human Genome Project? Well, it's only been a little while so far, but the Battelle Institute calculated that the billions of dollars of investment that the federal government put into the project has returned approximately 140:1, so far. That's a pretty good return. They estimate the total return is in the neighborhood of about a trillion dollars.

All right, you get the point. When economists try to figure out the return on investment of basic research, the numbers vary because it's hard to measure precisely. But, the estimated ROI ranges from 20 percent per annum and 60 percent per annum. Now, as my friend and PCAST colleague Bill Press points out, you might say: "With returns like that, I should invest my whole retirement fund in basic research!"

The problem is, of course, as the economists say, the returns are not appropriable – they can't be captured by a private investor. The investor in basic research can't fully capture the economic returns, because the fruits are largely knowledge that accrues to society at large. You can file patents, but this applies to only a very small fraction of knowledge. You can't patent laws of nature, and a good thing, that.

So this Miracle Machine has these two components. The first component is ***public*** investment in basic research. It has to be public. Private investors won't invest if they can't own the fruits, but the public can invest because it gets its return in the form of benefits for society and in greater tax revenues from the economic activity. The second component is private investment that comes along to enable commercialization.

We need every person in Congress, we need every American to understand how this Miracle Machine works – that it has these (and other) essential parts and they work together to produce social good.

If you think you can improve social returns by moving money from long-term basic research to short-term investments, you fundamentally misunderstand the Miracle Machine. You will end up with modest returns, at best.

Now, basic research investments pose a problem for our elected officials. For other kinds of investments, they expect that ***they*** should be allocating the funds to specific projects. But, we're saying: No, you shouldn't pick the projects. And, we're saying: Please don't ask researchers to explain how their work is directly connected to the national interest. Please don't ask them to be explicit, because it will cause them to focus on the short-term – in which case the game is lost.

So, how should funds be allocated? I hate to say it, because it sounds elitist, but, you need to rely on expert taste and expert judgment as to what's a great question to study. In a democratic society, it is hard to say that the right answer is to rely on experts. You don't want to have to say that. But it ***is*** the right answer.

Now, we shouldn't ask for unbounded trust. As Reagan said, "trust, but verify". The right way to monitor basic research, instead of asking what the payoff is ***going*** to be over the ***next*** 25 years, is to look backward and ask what my payoff has been over the ***last*** 25 years. If you continue to see that you've been doing amazingly well over these last 25 years (which has been the case), you should continue to invest in basic research! I suppose there is always the risk that, at some point the high returns on basic research will cease, but it hasn't happened yet!

We need to keep the Miracle Machine functioning, we need to train amazing young minds, we need to fund the best ideas, because, at least so far, no one has invented any system that has had a greater impact on human welfare. Thank you very much.