Robin Hanson Statement on Teaching

Long ago I was one of the top teaching assistants at the University of Chicago physics department. I had to learn to translate this ability into teaching economics here.

For example, I learned that students are far more willing to believe physicists about physics, than to believe economists about economics. To help overcome this skepticism, I run some classroom experiments in my microeconomics courses. These experiments give students a good initial hook for the concepts of marginal value and cost, and they show that standard theory predicts experimental outcomes reasonably well.

One physics practice that translates well to economics is "back of the envelope" analysis. For example, I ask why poets are paid less than romance novelists, and teachers are paid less than garbage men (hint: note supply side effects). And I ask whether a subsidy or tax would best correct for law mower externalities (hint: neighbors can negotiate easily).

While many undergraduates prefer teachers to give them the one true answer on policy questions, I try to avoid taking sides. In class discussions, I ask students to take and defend policy positions, and then I defend any undefended major positions, and mention any unmentioned major criticisms of these positions. The important thing is for students to understand the sorts of arguments that an economist would respect.

I am big on normative analysis. I review how normative economics differs from normative analysis elsewhere, and the basic types of market failure. I cover counter-arguments that most teachers don't cover, such as consistency checks (e.g., if this were the failure then we should also see that) and minimal interventions (e.g., this milder alternative should work just as well). I pay special attention to the many subtleties of - "stupid public" arguments, such as ignorance vs. irrationality, and consumers vs. voters.

Students are often quite capable of applying economic analysis to emotionally neutral products such as apples or video games, but then fail to apply the same reasoning to emotionally charged goods to which similar analyses would seem to apply. I make a special effort to introduce concepts with the neutral examples, but then to challenge students to ask wonder why emotionally charged goods should be treated differently.

While I have a strong mathematical training, and use math heavily in my research, I have learned that I do not need much math to communicate basic concepts to undergraduates. My math abilities are an important resource for our graduate students, however, who need to quickly transition to reading the journals and trying to write research papers.

As befits a junior faculty, I have so far taught courses as I've been needed. Like most professors, I would love to develop a new graduate course close on my research topics, such as "The Economics of Information." Also, I would love to lead broad ranging class discussions with honor students.

Robin Hanson Statement on Research

I here summarize my core research, research history, other research, and future plans.

Core Research

The question that animates my core research is this: what goes wrong when people debate important questions in science and policy, preventing them from making the best possible estimates, and what can we do to fix these problems? I study the rationality of disagreement in general and general institutions for improving debate, especially idea futures. I also focus on some particular areas in health policy, politics, and law.

Idea futures Otherwise known as prediction markets or information markets, idea futures are in my opinion one of our best hopes for better science and policy consensus. People are more honest with themselves and others when they bet things they value on their claims.

I was the first to publish proposals to widely apply speculative markets to better aggregate information for science and policy, to subsidize them via automated market makers, and to lower judging cost via audit lotteries. My leadership has been acknowledged by mentions in one hundred press/media articles, including a profile in *Fortune*, interviews in *Wired* and *The Chronicle of Higher Education,* and mentions in six *New York Times* articles, in two *Washington Post* articles, and in *Time, Science, Nature*, *Science News*, *Business Week*, and *Harvard Business Review*.

My general proposals, together with many replies to critics, appeared in *Social Epistemology*, *Wired*, *Foresight Update*, and *Extropy*. I also published a lab experiment design to test my proposal in *Social Epistemology*. My *IEEE Intelligent Systems* paper proposed using conditional markets to directly advise policy, a concept I explored further in a paper and in a forthcoming book chapter, and a concept which three research and three commercial teams are now pursuing.

I created the first internal corporate idea futures at Xanadu in 1989, and in 1994 my design and inspiration led to one of the first web markets, then called Idea Futures, for which I received the Prix Ars Electronica Golden Nica, a prestigious electronic art prize. Now called the Foresight Exchange, it is still the only web market where users can add claims to trade. My papers inspired the NewsFutures market, and the Hollywood Stock Exchange uses my proposed market makers.

I teamed with the firm Net Exchange on the DARPA FutureMAP program, and was a chief architect of our Policy Analysis Market (PAM). I picked our application area of military instability. We did lab experiments with difficult environments (six subjects estimated 256 states in three minutes) and found that my combinatorial betting mechanism aggregated information better than three other mechanisms. So my mechanism was used in PAM, and was elaborated in my other funded FutureMAP project. My mechanism is described in my *Information Systems Frontiers* paper, is the

basis for software now sold by Net Exchange, and is elaborated in two theory papers, on book orders and on modularity.

FutureMAP was cancelled one day after two senators (falsely) accused PAM of being a market for betting on terrorist attacks. Media Coverage of PAM was initially negative, but then the coverage of idea futures turned positive, inspiring many corporations to begin projects to forecast things like sales, product delivery dates, and bug rates.

Many critics said that bad guys might mislead such markets with their trades. I now have a related theory paper, and a forthcoming experimental paper in the *Journal of Economic Behavior*. When other traders suspect that manipulators may be present, the net effect is to *increase* average price accuracy, compared to markets without such manipulators.

Disagreement Our strongest clue that something goes very wrong in policy discussions is Aumann's classic result that rational agents do not knowingly disagree. After all, we knowingly disagree all the time. Theorists mostly treat this result as a curiosity, at first because it required strong assumptions. So I have worked to generalize Aumann's result.

My first *Theory and Decision* paper showed that knowledge of exact opinions is not needed – knowing who is an extremist will do. My *Economics Letters* paper generalized this result to every point during a conversation, instead of a mythical endpoint. Another paper tests this prediction in lab experiments, and my *Econometrica* paper corrects some mistakes in a related experiment. My second *Theory and Decision* paper generalized to the case of Bayesian "wannabes," who make errors, but also make a few simple corrections.

All these results depend on the common prior assumption, and so my third *Theory and Decision* paper shows how common priors are required by reasonable assumptions about the causal origins of priors. A paper in preparation also shows how common priors are implied by widely accepted assumptions regarding indexical counterfactuals. Another paper shows how to create common priors and then elicit probabilities, even with arbitrary state-dependent payoffs and risk aversion.

I summarize and interpret this whole literature in a paper where I suggest that human disagreements are typically dishonest, because we have other goals when choosing beliefs besides seeking truth. I explore this idea further in a chapter in a best-selling academic crossover book on the movie *The Matrix*. With Tyler Cowen, I co-organized a small workshop on self-deception attended by the main leaders of this field.

Better Science How can we improve this situation? I have pursued several other general approaches besides idea futures. My forthcoming book chapter considers how enhanced humans could be more truth-seeking. As prizes seem an attractive way to improve science funding, my empirical paper looks at why science funding has been mostly via grants. My recent theory paper suggests that a key to keeping experts honest is having enough participants with amateur motivations but expert knowledge.

Long ago I worked to develop the web in the hope that it would improve our ability to find good criticism. My first publication, on hypertext publishing in *ACM SIGIR Forum*, influenced the design of Xanadu, a well-known precursor to the World Wide Web.

Health Policy I have identified several specific information problems, and problem fixes, in health policy. My very first attempts at institution design led to a *Cato Journal* paper proposing to improve incentives by bundling health and life insurance. Ian Ayres and Barry Nalebuff of Yale recently revived this suggestion in their Harvard Business School Press book *Why Not?*

My *Journal of Public Economics* paper shows how regulators might rationally ban products rather than label them, for fear that consumers would interpret warnings as endorsements. So in many cases customers are better off when regulators are not able to ban products. My *Economics of Governance* paper explains why unions choose insurance, by showing how democratic choice mitigates adverse selection. Insurance chosen by random juries would do even better.

During my Robert Wood Johnson Foundation health policy post-doctorate at UC Berkeley, I sought an integrated explanation of diverse puzzles in health behavior and policy, based on evolutionary psychology and the idea that medicine is more about showing that you care than improving health. A summary of my longer paper appears in my *Social Philosophy and Policy* paper, where I also introduced a Bayesian style approach to treating intuition errors in meta-ethics. My related survey paper shows that people pick more medicine for loved ones than they would pick for strangers or themselves in the same situation. Related ideas also appear in my note on why cryonics isn't popular, and in my paper on fairness norms as way to get clearer fitness signals.

Law and Politics My first economics research led to an article on the economics of wiretaps in *Communication of the ACM*, a top computer science journal. I estimated wiretap benefits to be well below the costs to telecom companies to preserve them, and argued that the best tradeoffs come when police agencies pay for such preservation.

Another paper of mine shows that while voter incentives to be informed can be weak after candidates position-taking, they can be strong beforehand, and using random juries makes them stronger at all times. My book chapter comments on game theory in politics.

I also published twenty-eight short book chapters on institution suggestions, for which I received a *Global IdeaBank* award. Another paper of mine considers private law enforcement. And my theory paper shows how spatial competition and efficiency varies with price rules and the spatial dimension.

A Short History

To make sense of my other research areas, it helps to know a bit of my history.

The National Academy of Sciences recently listed me as one of "100 of the country's leading scientists, under the age of 45," and the reporter James Pethokoukis titled his recent profile of me "Chief economist at Starfleet Command - The Big Ideas of Robin Hanson." I may not merit these descriptions yet, but I have long aspired to them.

An avid science fiction reader, I first studied engineering, then physics, then philosophy of science, always focusing on foundations. Silicon Valley then seduced me to pursue artificial intelligence research as a day job. On the side I pursued hypertext publishing (i.e., the web) and future studies with people like Douglas Engelbart and Eric Drexler.

I began to try to invent new institutions, such as "idea futures," my proposal to use speculative markets to improve consensus and fund research for science and policy. Feeling my proposals were ignored, and excited to learn that economists were testing new institution ideas in lab experiments, I decided to study social science at Caltech.

At Caltech I was disappointed to find that economists had less interest than I had hoped in designing new institutions. But I was delighted to find that economists know far more than physicists realize. Upon learning the economic theory of information, I saw the social world anew.

The world made more sense, but many puzzles remained, to which I turned my attention. I studied puzzles in politics, health policy, and elsewhere, but paid the most attention to the puzzle of why people disagree. I was also at times drawn back to topics in future studies, and fundamental questions in philosophy and physics.

Finally, my attention has returned to idea futures. Some computer scientists implemented my first design to make the Foresight Exchange. I studied how such markets could advise policy, and be run more efficiently, and then joined the DARPA FutureMAP project, while continuing to research related issues.

More Research

I have also done research on physics, artificial intelligence, and futures studies.

Physics My interdisciplinary work led once to a *Physics and Computation* publication on reversible agents, and recently to a *Foundations of Physics* paper trying to reconcile the many worlds interpretation with the Born probability rule via a certain mangling process producing a selection effect. A related paper models drift-diffusion of worlds.

Artificial Intelligence My work as an artificial intelligence researcher specializing in Bayesian statistics led to my *Applied Optics* paper and four conference proceedings publications, in *Maximum Entropy*, *IJCAI*, an *IJCAI* workshop, an *AAAI* symposium. Conference proceedings are the dominant publication form in many computer science areas, and *IJCAI* is the most prestigious conference in artificial intelligence.

Future Studies I have published some analyses of future technologies and their economic implications, six in *Extropy* and two in the *Journal of Evolution and Technology (JET)*. I edited *JET* for over two years (but not when I published). My *Extropy* analysis of uploads is still widely cited, and eleven press articles covered my *JET* article on how to live in a simulation. Eight other articles covered my future studies.

Recently I wrote a book chapter arguing that futurists too often implicitly assume autarky, and a book chapter on how nanotechnology might change the economy. Various papers of mine give data suggesting a more peaceful future, a game theoretic model of interstellar colonization, a probability analysis of whether the early appearance of life on Earth indicates that life appears easily, a growth theory analysis of the upload transition scenario, and an empirical analysis showing that long term growth trends are roughly consistent with the predicted upload transition.

Future plans

I will probably spend the next few years continuing to work on the research and papers already in my pipeline. As more time becomes available, I hope to write a book, as I often feel limited by the narrow scope that a journal paper allows.   Possible book topics include disagreement, idea futures, and the economics of future technologies.

I also plan to continue to pursue my core areas of trying to understand better why we disagree, and trying to construct idea futures markets that can give us a better consensus on important science and policy questions.