

# Hanson-Hughes Upload Debate

## Editor's Introduction

"Follow the money" has been the operational rule for historians and investigative journalists since at least the Watergate era, if not earlier. Futurists do not have a money trail to follow, but instead must predict the trajectory of economic relations based on assumptions of what technological and social developments the future may hold. Many futurists assume that nanotechnology in combination with Artificial Intelligence (AI) will yield a world of material abundance with little or no need for human labor. The nano/AI cornucopia will rain down wealth upon one and all, giving slackers and solid workaholics equal access to almost anything they could ever need or want. But is this really the most likely scenario?

Economist Robin Hanson thinks not. As he reasoned in his paper "If Uploads Come First: The Crack of a Future Dawn" (1994), if the technology to copy, or upload, human minds is developed before strong AI, then the "result could be a sharp transition to an upload-dominated world, with many dramatic consequences. In particular, fast and cheap replication may once again make Darwinian evolution of human values a powerful force in human history. With evolved values, most uploads would value life even when life is hard or short, uploads would reproduce quickly, and wages would fall. But total wealth should rise, so we could all do better by accepting uploads, or at worse taxing them, rather than trying to delay or segregate them."

In his book *Citizen Cyborg* (2004), bioethicist James Hughes took issue with the social implications of Dr. Hanson's paper. Dr. Hughes objected to Hanson's upload scenario, characterizing it as a "dismal, elitist utopia" that "recapitulates Marx's vision of universal immiseration, but this time in the Matrix."

When Hanson learned of what Hughes had written, on March 29, 2006 he posted a response to the WTA-Talk email list of the World Transhumanist Association. During March and April, a debate ensued. The text that follows is a lightly-edited transcript of that online debate. The original discussion thread, which includes messages from additional participants, can be read at: <http://www.jetpress.org/thread.html>

The debate text here is limited to what Hanson and Hughes wrote online, but with the addition of closing comments from each man which were written specifically for this document.

I would like to thank Robin Hanson and James Hughes for allowing me to assemble their email debate messages into this document. I hope it will prove to be a more convenient format for readers to follow the fascinating and important issues under discussion. All credit for the content of this document belongs to Dr. Hanson and Dr. Hughes. All errors, omissions, or infelicities of language are my responsibility alone.

*Michael LaTorra*

*WTA Publications Director*

## ***The Debate***

Robin Hanson:

I learned last night that pages 169-170 of James Hughes' book *Citizen Cyborg* discuss my paper "If Uploads Come First" <http://hanson.gmu.edu/uploads.html> . In that section Hughes severely misrepresents my positions. He paints me as gleefully advocating having a ruthless pampered ruling class, "not very different from contemporary luxury suburbs," being "set off from a larger terrain of violence and injustice" among downtrodden masses. I am posting a public response here, to a list I know that he reads.

From James Hughes' book *Citizen Cyborg* (excerpted at:

<http://hanson.gmu.edu/PAM/press/CitizenCyborg-excerpt.txt>):

The extropians have also cultivated important allies in libertarian politics such as Virginia Postrel and Ron Bailey, sympathizers with their militant defense of personal liberty and hostility to regulation and environmentalism. ... Postrel has now organized Bailey and other technolibertarians, ... into The Franklin Society. The first project of the Society has been to campaign against attempts to ban embryonic stem cell research. In 2003, one member of the new Franklin Society, extropian economist Robin Hanson, a professor at George Mason University, achieved his full fifteen minutes of fame. ... While I think the experiment had merit and would not have encouraged terrorism, the episode does illustrate some of the moral and political blindness that the unreformed extropian anarcho-capitalist perspective lends itself to.

Putting me in this context suggests that I have a "militant defense of personal liberty and hostility to regulation and environmentalism" and that I am an "unreformed extropian anarcho-capitalist". While I have long associated with people under the flag "extropian" (via mailing lists, conferences, and journals), I deny these other claims.

In 2002 I agreed to sign a petition saying "therapeutic cloning should not be banned," sponsored under the name "Franklin Society," but I otherwise have no knowledge of or association with such a society. I presume that James would also have signed such a petition at the time.

The Policy Analysis Market (<http://hanson.gmu.edu/policyanalysismarket.html>) was a joint project of many people, and I was the only such person with any "extropian" associations. Other people on the project were more directly responsible for the web page examples that caused the furor; those people can reasonably be blamed for "political blindness," though not in my opinion for "moral blindness."

... he published a now often-cited essay "If Uploads Come First - the Crack of a Future Dawn" in *Extropy* magazine. ... He argues that the capabilities of

machine-based person would be so much greater than those of organic humans that most non-uploaded people would become unemployed.

My main argument was that uploads will *cost* much less, not that they would be more capable.

... Eventually the enormous population of uploads would be forced to work at very low subsistence wages - the cost of their electricity and disk space - ruled over by a very few of the most successful of the uploads.

I say nothing about people being ruled over by a successful elite. I talk disapprovingly about wealth inequality among humans, caused by some humans not insuring against an upload transition. I talk about inequalities in the number of copies made of particular uploads, but I do not speak at all about wealth inequalities among uploads.

Hanson dismisses the idea that governments could impose redistribution on uploads since there would be large economic benefits of an unfettered transition to Matrix life.

The only thing I say about government redistribution [in "If Uploads Come First"] is this:

*politicians would do better to tax uploads and copies, rather than forbidding them, and give the proceeds to those who would otherwise lose out. {Note that such a tax would be a tax on the poor, paid to the relatively rich, if one counted per upload copy.}*

This is hardly a dismissal of redistribution. Nor is my claim one I think James would disagree with.

Returning to *Citizen Cyborg*:

The average quality of life of the subsistence upload and the unemployed human would allegedly be higher than before. So the best we future residents of an uploaded society can do is become as versatile as possible to maximize our chances of ending up as one of the lucky rule or employed classes.

The first sentence here is a reasonable summary of my position. But the second sentence here does not at all follow from the first, and I said nothing like it in my paper.

Hanson dismisses the idea that people will fight the division of society into a mass of well-fed plebes and a superpowerful elite since the growth in the gross domestic product is the sole measure of his utopia,

I never mentioned anything like "gross domestic product" and so certainly didn't endorse it as a "sole measure" of value. The division I talk most about is humans and uploads, not "well-fed plebes and a superpowerful elite," and to the extent I take sides it is with the uploads, who are poorer.

Hanson: ...and the elimination of the weak will select for "capable people willing to work for low wages, who value life even when life is hard."

With a dismal, elitist utopia like this who needs a Luddites's dystopia?

My paper was mainly a *positive*, not a *normative* analysis. That is, I mainly tried to forecast what would happen under such a scenario, and only make limited comments along the way about what private actions or public policies we might prefer. I tried not to shy away from describing the negatives along with the positives.

Even after all of Hughes' strong language, I'm not sure I can identify any particular claim I made in my paper that he would disagree with. And while he favors redistribution, it is not at all clear to me who he wants to take from, and who to give to under the scenario I describe. After all, given the three distinctions of human/upload, rich/poor, and few/many-copied, there are eight possible classes to consider.

---

James Hughes:

Thanks for taking the time to respond Robin.

And doing so in a comradely, academic exchange even though my description of your views was polemical and, in your analysis, incorrect.

However, I just was in Oxford with you at the James Martin Institute's Future Forum ("Tomorrow's People: The Challenges of Technologies for Life Extension and Enhancement" <http://www.martininstitute.ox.ac.uk/jmi/forum2006/>) and saw you give another version of this very "Crack of a Future Dawn" scenario in which you did not mention any regulatory or political solution possible to this scenario of general unemployment that you foresee being created by a proliferation of uploaded workers, so I don't think my analysis of your views needs much revision. When a member of the audience asked, as I have in the past, whether we might not want to use some kind of political method to prevent general unemployment and wealth concentration in this Singularitarian scenario your response was, as it has been in the past and was in that paper, that no one will want to prevent this coming to pass since we will all own stock in this explosive economy and will therefore all be better off than we were before.

In the essay "If Uploads Come First" you say:

...imagine the potential reaction against strong wage competition by "machine-people" with strange values. Uploading might be forbidden, or upload copying might be highly restricted or forbidden...If level heads can be found, however, they should be told that if uploading and copying are allowed, it is possible to make almost everyone better off. While an upload transition might reduce the market value of ordinary people's human capital, their training and ability to earn wages, it should increase total wealth, the total market value of all capital,

including human capital of uploads and others, real estate, company stock, etc. Thus it can potentially make each person better off.

I'll say again: I think the scenario is a scary one, in ways that you don't appear to recognize, because most people have little confidence that they would actually be better off in a world in which all "human capital" is radically devalued by the proliferation of electronic workers. That includes me; although I do own stocks in mutual funds today, and those stocks might benefit from a Singularitarian economic boom, I still feel like my world and my future is being determined by unaccountable elites who control my political institutions, elites quite content to see vast numbers of people immiserated as inequality grows.

The scenario you describe is one where it appears these inequalities of wealth and power would just get a lot more extreme and far less politically ameliorable.

If Singularitarianism wants to paint a truly attractive future, and not one that simply fans the flames of Luddism, then it has to put equality and social security in the foreground and not as a dismissive afterthought. To his credit Moravec, in *Robot*, argues for a universalization of Social Security as a response to human structural unemployment caused by robot proliferation. Marshall Brain (<http://marshallbrain.com>) reached the same conclusion, and several of the principals at the Institute for Ethics and Emerging Technologies (<http://www.ieet.org/>) and I are supporters of the concept of a Basic Income Guarantee (<http://www.usbig.net/>). But since this would require state intervention I suspect you don't favor such a proposal, which is why you advocate(d) minarchist solutions like universal stock ownership in the Singularity.

Perhaps the most troubling parts of the essay are:

As wages dropped, upload population growth would be highly selective, selecting capable people willing to work for low wages, who value life even when life is hard. Soon the dominant upload values would be those of the few initial uploads with the most extreme values, willing to work for the lowest wages.

And then later

Those who might want to be one of the few highly copied uploads should carefully consider whether their values and skills are appropriate. How much do you value life when it is hard and alien?...Those who don't want to be highly-copied uploads should get used to the idea of their descendants becoming a declining fraction of total wealth and population....

How is this different from a radical Social Darwinism, arguing that this Pac-man world will eliminate all the uppety prole uploads, the ones who might want minimum wage laws or unions, and just leave the good hard workers willing to work for subsistence?

You say:

I talk disapprovingly about wealth inequality among humans, caused by some humans not insuring against an upload transition.

Which I assume refers to this passage, the only one that mentions inequality in the essay:

Would there be great inequality here, with some lucky few beating out the just-as-qualified rest?...Computer technology should keep improving even if work on uploading is delayed by politics, lowering the cost of copying and the cost to run fast. Thus the early-adopter advantage would increase the longer uploading is delayed; delaying uploading should induce more, not less, inequality. So, if anything, one might prefer to speed up progress on uploading technology, to help make an uploading transition more equitable.

So yes, you did argue against inequality, but only in passing as one reason why of a rapid transition to general unemployment in an upload-dominated economy should not be hampered by political regulation. If we try to slow this transition, a minority of uploads will just become even richer. So we should speed the transition to give more uploads a piece of the pie.

But you are right that you do not explicitly describe a concentration of wealth, only mention it as a possibility in order to discourage regulation, and you do describe mechanisms that might spread wealth out among the uploads and humans. But then how is that consistent with the scenario "As wages dropped, upload population growth would be highly selective, selecting capable people willing to work for low wages"?

Doesn't that imply that humans would be unemployed, most uploads working for upload-subsistence, and some very few uploads will be raking in the big bucks? Or is the scenario one of truly universal and equal poverty among all the uploads, with no wealthy owners of capital anymore in the equation?

You note that we might progressively tax wealth accumulators in this economy, but then in the last sentence of the paper's abstract you say:

...total wealth should rise, so we could all do better by accepting uploads, or at worse taxing them, rather than trying to delay or segregate them.

And then later:

If forced to act by their constituents, politicians would do better to tax uploads and copies, rather than forbidding them, and give the proceeds to those who would otherwise lose out.

Which pretty clearly implies that you only grudgingly accept Social Security and redistributive taxes on uploaded wealth accumulators as a concession to political unrest, and not as an obvious and essential step in maintaining an egalitarian polity.

That said, the reason I devoted the attention to the essay that I did was because I think it is a very smart and foresightful scenario of a future that could come to pass. But I do think the piece illuminates a techno-libertarianism that most people will find scary, and which our movement needs to contextualize in proactive social policies, precisely in order to defend the possibility of uploading from bans. As you note, in such a future I would recommend (fight for) redistribution from the wealthy - uploads or human - to the majority, to ensure some form of rough equality, and some form of Social Security more egalitarian and universal than stock ownership, such as a Basic Income Guarantee. (Did you have in mind the distribution of mutual fund shares to everyone in the developed and developing world? If so, I think that would be a welcome addition to the scenario.)

And if the economy and world start to change with the rapidity that you forecasted at Oxford—doubling every couple of weeks, with a proliferation of uploads—I would also favor strong regulatory action to slow and temper the transition. A rapid take-off Singularity is both dangerous and anti-democratic, and we should say so and say what kind of policies we think are necessary to make sure it doesn't happen, and how we can slow it down if it starts. You don't really endorse redistributive, Social Security or regulatory policies in the essay, but rather argue against them, and you didn't even mention them at Oxford. Clearly, you consider them suboptimal, counter-productive concessions to Luddites. So I do think we have a difference of opinion that I have not mischaracterized.

However, I apologize again for the polemical tone of the passage since we are friends, and for not more fully describing your views.

---

Robin Hanson:

James, you are acting more like a politician than a scholar here. I tried to focus attention on how the specific words of your summary differ from the specific words of my paper that you purport to summarize, but you insist on trying to distill a general gestalt from my writings, based on a simple one-dimensional redistribution-based political axis. Apparently in your mind this axis consists of good people on the left who support redistribution, employment, and high wages in the service of equality, and evil people on the right who seek inequality, unemployment, and low wages in the service of social Darwinism. Since I predict that the technology of uploads will lead to unemployment for humans and low wages and Darwinian selection for uploads, and I only mention and endorse one possible redistribution, apparently not enthusiastically enough for you, I must be one of the evil people. Come on!

With cheap uploads there is pretty much no way to escape "unemployment" for most humans. That is, while you could give people make-work jobs, and/or pay them lots more than the value of their work, the truth is that for most people the value of their labor to others would be little, and if that were all they were paid

they would not work. Also, unless we are willing to impose population controls on uploads far more Draconian than those in China today, we could not escape uploads getting low wages and undergoing Darwinian selection. The only way to induce upload wages far above the cost of creating uploads would be to prevent the vast majority of uploads from existing at all. And the only way to avoid Darwinian selection among uploads would be, in addition, to limit severely the number of copies made of individual uploads. These are not statements of advocacy; they are just the hard facts one would have to deal with under this scenario. So are you criticizing me for not endorsing Draconian upload population control?

I repeat again the conclusion of my last message:

*while he [Hughes] favors "redistribution," it is not at all clear to me who he wants to take from, and who to give to under the scenario I describe. After all, given the three distinctions of human/upload, rich/poor, and few/many-copied, there are eight possible classes to consider.*

To elaborate, the key reason I hesitate to more strongly endorse redistribution is that it is not clear who are really the "deserving poor" to be aided in this scenario. In dollar terms the poorest would be the uploads who might be prevented from existing. If one only considers the per-capita wealth of existing creatures, the poorest would be the many copies of those "who value life even when life is hard." But these would be the richest uploads in clan terms, in that such clans would have the most copies; counting by upload clans identifies a different poor. Humans would have far larger per-capita income, but many would be poorer if we talk in terms of income relative to their subsistence level, since the subsistence level for uploads would be far lower than that of humans. Should their not taking advantage of the option to convert from human to upload be held against the "poor" humans? Finally, a few humans will have rare abilities to make substantial wages; does that make them "rich" even if they do not own much other wealth? If you are going to criticize me for not explicitly supporting the redistribution you favor, I think you should say more precisely who you would take from and who you would give to.

Now for a few more detailed responses:

If Singularitarianism wants to paint a truly attractive future, and not one that simply fans the flames of Luddism, then it has to put equality and social security in the foreground and not as a dismissive afterthought.

My purpose is *not* to paint a truly attractive future; my purpose is to paint as realistic a picture as possible, whatever that may be.

... in Oxford with you ... When a member of the audience asked, as I have in the past, whether we might not want to use some kind of political method to prevent general unemployment and wealth concentration in this Singularitarian scenario

This did not happen. One person asked "What does your economic model predict people will do?" This was in response to the idea of improving robots, but he said nothing specifically about politics, employment, or wealth concentration.

Hughes wrote:

your response was, as it has been in the past and was in that paper, that no one will want to prevent this coming to pass

I never said that no one would try to stop uploads.

I'll say again: I think the scenario is a scary one, in ways that you don't appear to recognize, ... although I do own stocks in mutual funds today, and those stocks might benefit from a Singularitarian economic boom, I still feel like my world and my future is being determined by unaccountable elites who control my political institutions, elites quite content to see vast numbers of people immiserated as inequality grows.

I am well aware that the scenario I describe is scary, and also that many people do not trust political elites to act in their interest. I do not argue that people should trust political elites.

[Hughes quoted Hanson] "As wages dropped, upload population growth would be highly selective, selecting capable people willing to work for low wages."

Doesn't that imply that humans would be unemployed, most uploads working for upload-subsistence, and some very few uploads will be raking in the big bucks? Or is the scenario one of truly universal and equal poverty among all the uploads, with no wealthy owners of capital anymore in the equation?

My scenario is consistent with both high and low concentration of ownership of capital, and with high or low inequality of wages among uploads. I make no prediction about there being a few very rich uploads.

Moravec, in *Robot*, argues for a universalization of Social Security as a response to human structural unemployment caused by robot proliferation. ... since this would require state intervention I suspect you don't favor such a proposal, ... You don't really endorse redistributive, Social Security or regulatory policies in the essay, but rather argue against them, and you didn't even mention them at Oxford, and clearly consider them suboptimal, counter-productive concessions to Luddites. ... Which pretty clearly implies that you only grudgingly accept Social Security and redistributive taxes on uploaded wealth accumulators as a concession to political unrest, and not as an obvious and essential step in maintaining an egalitarian polity.

You keep jumping to conclusions. Just because I take no position does not mean I am against your position.

---

James Hughes:

Robin Hanson wrote:

Since I predict that the technology of uploads will lead to unemployment for humans and low wages and Darwinian selection for uploads, and I only mention and endorse one possible redistribution, apparently not enthusiastically enough for you, I must be one of the evil people.

I don't think you are evil. I just think you share the worldview of many American economists, and most of the 1990s transhumanists, who prefer a minarchist, free-market oriented approach to social policy, and do not see redistribution and regulation as desirable or inevitable. My book was a critique of that point of view, and I used your article as a brilliant paradigmatic example of it. Empirically, the people who are most attracted to libertarianism, neo-liberalism (or whatever) are those who are most likely to benefit from those policies: affluent men in the North. My challenge to you, and all of us, is that we need to break out of those blinkers. Try to see the world from the perspective of the billions who live on dollars a day. And from the perspective of those who are quite suspicious of emerging technologies because these are used to bomb them or exploit them. For such people, the benefits of technology are often inaccessible.

As to your assertion that your piece is merely descriptive and not normative, I leave that to the reader to judge; see <http://hanson.gmu.edu/uploads.html>

To me it is clear that you are excited about this future (a "Crack of a Future Dawn" after all) and see it as a desirable one with universal advantages, a future that should not be slowed or regulated by state intervention. So you are about as non-normative as Karl Marx in *Das Kapital*—here is how the system works, here is our inevitable future, here is how people will react, and here is how we will end up in paradise. No, there is no normative analysis needed in techno-utopian determinism—we either get with the program, or end up on the dustbin of history.

...unless we are willing to impose population controls on uploads far more Draconian than those in China today, we could not escape uploads getting low wages and undergoing Darwinian selection.

Rights do not exist in isolation. Reproductive rights have to be balanced against others, such as the right to life, liberty and the pursuit of happiness. Aubrey de Grey, for instance, has been quite clear in emphasizing that we will inevitably need to consider limits on reproduction if we have unlimited life expectancy. Uploading and space exploration only moves out the necessity.

In addition, potential future people, uploads or human, do not have rights; only existing people do. So I do think reproductive control on uploads would make perfect sense, and would be one of the policies that should be pursued if we were faced with your scenario.

In effect, your scenario is one version of the runaway AI scenario, with individual viral egos instead of one monolithic AI, and I see both as existential risks that we need transhumanist policies to prevent, not to facilitate.

The only way to induce upload wages far above the cost of creating uploads would be to prevent the vast majority of uploads from existing at all.

Then why isn't population control the only way to induce human wages to rise? Yes, labor supply does affect wages, but so do government policies like worker safety laws, taxation and minimum wages. The fact that these policies are completely off your radar is part of the problem.

And the only way to avoid Darwinian selection among uploads would be to in addition would be, in addition, to limit severely the number of copies made of individual uploads.

Again you reveal a Social Darwinist view without any acknowledgement that there can be collective solutions to social problems. Of course, we can prevent the forces of social selection from killing off all the beings who don't want to work for low wages, and selecting for the diligent subsistence drones. If there is such a population pressure, we create new selection parameters to encourage or require other population traits. But again, the notion of social engineering is apparently anathema.

An example: clearly employers already prefer human workers who work long hours, are perfectly loyal, and never organize for collective benefits. To the extent that there are psychopharmaceuticals and cybernetics that allow employers to "perfect" their workers there will be efforts to apply them.

So we pass laws that, even if we all get to take Modafanil, no one can work more than 50 hours a week. We pass laws against loyalty drugs/chips, just as we once outlawed serfdom and company towns. We pass collective bargaining laws that mandate that all uploads need to use at least 30% of their CPU cycles for personal, non-remunerative enrichment.

Without these kinds of policies we could drift toward hive-mind drone existences, losing individual subjective agency, which is one of the existential threats pointed to by World Transhumanist Association Chairman Bostrom.

...while he [Hughes] favors "redistribution," it is not at all clear to me who he wants to take from, and who to give to, under the scenario I describe. After all, given the three distinctions of human/upload, rich/poor, and few/many-copied, there are eight possible classes to consider.

Rich to Poor will do nicely thank you, regardless of their number or instantiation.

To elaborate, the key reason I hesitate to more strongly endorse redistribution is that it is not clear who are really the "deserving poor" to be aided in this scenario.

Yes, "deserving poor" is part of the problem. The desirability of rough social equality does not depend on any notion of "deservingness".

I do not argue that people should trust political elites.

No, only the unfettered market. Is there any form of law, state or collective action other than market exchange in your imagined Dawn?

My scenario is consistent with both high and low concentration of ownership of capital, and with high or low inequality of wages among uploads. I make no prediction about there being a few very rich uploads.

Sadly, reality is not consistent with the notion that there will be a new era of equality with radical technological change. The winners/owners will change, but any equality to be achieved is something we have to fight for, not something to be fervently wished for.

Just because I take no position does not mean I am against your position.

Robin, I don't think you have ever taken my position(s) seriously enough to reject them - they simply are alien to the kind of economic analysis that you do. I wish you would take them seriously enough to explicitly reject them so we could have that conversation.

---

James Hughes:

Quoting a message from Russell Wallace:

*I agree with you that this is a potential problem, but rather than rely on a monolithic government to legislate our way out of them (which has well known problems of its own), I will suggest that this is exactly the sort of thing my Domain Protection idea is designed for:*  
<http://homepage.eircom.net/~russell12/dp.html>

As I understand your proposal Russell, it is that we would ask the world-controlling Friendly AI to set up regions that are not allowed to interfere with one another, one for uploads and one for ur-humans.

This of course broaches the problems that we face today with the enforcement of international agreements that countries should not invade one another.

A) There are sometimes good reasons for countries to be invaded, as when they pose a threat to the rest or are violating human rights.

B) There needs to be a legitimate, accountable global authority to enforce those agreements, and unfortunately the US Presidency is not such an authority

I don't see how a Friendly AI gets us there. If it has the kind of power necessary, it is clearly monolithic. If it is legitimate, but not accountable, it's a benevolent monarchy (cross your fingers). If it is legitimate and accountable (replace-able, control-able) then it is a part of global democratic governance.

---

Robin Hanson:

James Hughes wrote:

I don't think you are evil. I just think you share the worldview of many American economists, and most of the 1990s transhumanists, who prefer a minarchist, free-market oriented approach to social policy, and do not see redistribution and regulation as desirable or inevitable. ...

Hanson: I do not argue that people should trust political elites.

No, only the unfettered market. Is there any form of law, state or collective action other than market exchange in your imagined Dawn?

You keep making these false statements about me, which I deny. I teach economics and in most lectures I make statements about the desirability and inevitability of regulation and redistribution. Really.

... Yes, labor supply does affect wages, but so do government policies like worker safety laws, taxation and minimum wages. The fact that these policies are completely off your radar is part of the problem.

I am well aware of such policies. But my claim is that in this context they would "prevent the vast majority of uploads from existing at all" if they raised wages a lot remains true. I wrote: "...while he favors 'redistribution,' it is not at all clear to me who he wants to take from, and who to give to, under the scenario I describe. After all, given the three distinctions of human/upload, rich/poor, and few/many-copied, there are eight possible classes to consider."

James responded:

Rich to Poor will do nicely thank you, regardless of their number or instantiation.

I gave a long analysis showing how there were at least five different ways to conceive of who are the "poor" in such a scenario, and I have twice now asked you to clarify which of these groups you want to favor with redistribution. You complain that I have not supported "redistribution" but without clarification this can only be a generic slogan.

---

James Hughes:

Robin Hanson wrote:

You keep making these false statements about me, which I deny.

I'm sorry you think I'm misrepresenting you. Of course you know about the political side of political economy, and I'm sure you teach about it. What I keep wanting is more realistic application and advocacy of the legitimate role of democratic deliberation and law in your writing.

You are associated, for instance, with "ideas futures" and market-based approaches to aggregating social preferences as a way to replace democratic mechanisms. As I said, I think your proposals are interesting and I would love to see the results of the experiments. But they do indicate a directionality in your work over the last fifteen years, arguing for a shift from reliance on democratic deliberation to market mechanisms.

Isn't that the case? Isn't it fair to characterize you as a libertarian economist?

worker safety laws, taxation and minimum wages... would "prevent the vast majority of uploads from existing at all" if they raised wages a lot remains true.

Yes, we agree about that. If we regulated uploads in certain ways it would restrict the incentive to clone/bud/build more of them. Just like passing laws that you have to send your kids to school instead of work them to death in the fields or factories changes kids from exploitable labor into luxury consumables, reducing the economic incentive to have them.

I gave a long analysis showing how there were at least five different ways to conceive of who are the "poor" in such a scenario, and I have twice now asked you to clarify which of these groups you want to favor with redistribution. You complain that I have not supported "redistribution" but without clarification this can only be a generic slogan.

Your examples are interesting, and worthy of additional discussion, but I really don't have to parse them before I can advocate a general principle that I want to live in a roughly equal society.

But I'll make a stab: in other writing I've pointed to the fact that liberal democracy is founded on the consensual myth of the discrete, continuous, autonomous individual. To the extent that neurotechnology erodes that consensual illusion, it fundamentally problematizes liberal democracy (and "the market"). I call that the "political Singularity," and I don't mean that in a *whoopee!* way.

So the problem you pose of whether a "clan" of upload clones, all sharing core identity software, should be treated as one—very rich—individual or a bazillion very poor individuals is a really serious problem for the future. Perhaps we will need a bicameral legislature, like the US Senate and House, one based on personalities and the other on bodies.

I don't know and I find the prospect very troubling. I would like to live in a world, like Brin's *Kiln People*, where I could send a copy of myself to work while the base-unit me stays home to read and cook. But in Brin's world, even though the clones only last 48 hours, they still have existential crises about whether they are the same as the base person, or a separate person with a right to life. We have yet to come up with a good solution to these dilemmas, which may be another reason to phase them in cautiously.

---

Robin Hanson:

Marcelo Rinesi wrote:

*The notion that -devoid of legal, societal or other restrictions; assuming that they will be possible and cheap; assuming that they will behave roughly as Von Neumann-Morgenstern utility maximizers, etc- uploads will eventually displace humans from most of the economic system and then compete fiercely between themselves, seems reasonable under the light of what we know of economics (substitute for "game theory" if you will or even "what **I** would do if I woke up uploaded"). The qualifications "devoid of legal, etc." are critical in this paragraph, of course. Change the parameters and the model results change; to some degree, the polemical question is not that the model is wrong, but what end results would be desirable, which ones of those end results would be possible, and what parameters would take us there.*

Yes, that is just how economic theorists like myself work. We first create a baseline model, the simplest one we can come up with that describes the essence of some situation, and then we vary that model to explore the effects of both various outside influences and of possible policies one might choose. The simplest model of most situations tends to be a low regulation model, but that does not mean that we are recommending no regulation. That is just usually the best starting point for considering variations.

---

Robin Hanson:

James Hughes wrote:

Hughes: I just think you ... do not see redistribution and regulation as desirable or inevitable.

Hanson: You keep making these false statements about me, which I deny.

Hughes: I'm sorry you think I'm misrepresenting you....You are associated, for instance, with "ideas futures" and market-based approaches to aggregating social preferences as a way to replace democratic mechanisms.... But they do indicate a ... shift from reliance on democratic deliberation to market mechanisms. Isn't that the case? Isn't it fair to characterize you as a libertarian economist?

No, it is not fair to characterize me as a libertarian economist. Some of my colleagues perhaps, but not me. You have been so far complaining that since I did not talk much about regulation in my uploads paper, that I must be hostile to the idea and unaware of the regulatory issues you hold dear. I have been trying to explain that I am aware of such issues and remain open to regulation, but that a low regulation analysis is usually the best first analysis step in economic

analysis. I had thought a bit about upload regulation, but it is a messy situation and I felt uncertain, so I choose not to say anything in that twelve-year-old paper.

The subject of "idea futures" as applied to government policy is about *how* we should choose regulation. It is not itself pro- or anti-regulation. Yes, I've advocated trying out markets to choose regulation, but that doesn't make me against democratic deliberation. For example, I am a fan of James Fishkin's experiments in deliberative democracy mechanisms.

As I said previously:

I gave a long analysis showing how there were at least five different ways to conceive of who are the "poor" in such a scenario, and I have twice now asked you to clarify which of these groups you want to favor with redistribution. You complain that I have not supported "redistribution" but without clarification this can only be a generic slogan.

To which James replied:

Your examples are interesting, and worthy of additional discussion, but I really don't have to parse them before I can advocate a general principle that I want to live in a roughly equal society.

Well, that is a key difference in our styles. "Equal society" is too vague a slogan for me to endorse. ("Equal in what?" my internal critic screams.) I would rather not take a public position if I cannot find something clearer to endorse. But please do not mistake my lack of many positions on upload regulation in my first uploads paper for not my caring about or being aware of regulatory issues.

For your information, regarding the questions I posed, my current leanings are that creatures who might exist should count in our moral calculus, that upload copies will diverge quickly enough that they should mostly be treated separately, instead of as clans, that the ability of humans to earn substantial wages should not matter much beyond its contribution to their income, and that while the fact that the human subsistence levels are higher should be a consideration, that consideration is greatly weakened when humans reject the option to convert into cheaper-to-assist uploads. Your intuitions may differ, but I don't think anyone should feel very confident about such opinions.

---

James Hughes:

Robin Hanson wrote:

...it is not fair to characterize me as a libertarian economist.

Excellent. Delighted to hear it.

I am a fan of James Fishkin's experiments in deliberative democracy mechanisms.

Excellent. Me too. I think they complement the idea markets mechanism nicely in our promotion of participatory models of future governance.

...my current leanings are that creatures who might exist should count in our moral calculus,

Hmm. A long-standing debate in utilitarian theory, as you know. Clearly, we want to make policy that will ensure the greatest possible happiness for all the beings that exist in the future, even though we are not obliged to bring them into existence. It seems like your model in "Dawn", if we interpret it as normative rather than descriptive, would fit with "the repugnant conclusion" of utilitarianism that we should create as many beings as possible, even if each of them might have less happy lives, because we will thereby create a greater sum of happiness than by creating fewer, happier beings. Is that what you mean?

...that upload copies will diverge quickly enough that they should mostly be treated separately, instead of as clans....

I would agree, but it depends on how much they are extensions of the primary subjective "parent." One can imagine one consciousness shared across many bodies or upload clones, tightly networked, where separate self-identity never arises. The Borgian possibility.

...that the ability of humans to earn substantial wages should not matter much beyond its contribution to their income....

Not sure what you mean there.

...and that while the fact that the human subsistence levels are higher should be a consideration, that consideration is greatly weakened when humans reject the option to convert into cheaper-to-assist uploads.

I make the same argument about human enhancement and disability. I'm happy to have the Americans with Disabilities Act urge accommodation of the disabled in the workplace. But to the extent that disability becomes chosen in the future (refusal of spinal repair, sight replacement, cochlear implants and so on) it weakens the moral case for accommodation.

If neo-Amish future humans refuse to adopt technologies that allow them to be faster and more enabled, or refuse to upload, any case they might argue for accommodation of their disadvantage is weak. But framing all humans who decide to remain organic as undeserving, self-crippers in a brave new uploaded world is part of the political challenge your essay points us to. We need to come up with a more attractive frame for the co-accommodation of organic and upload life.

---

Robin Hanson:

James Hughes wrote:

Hanson: ...my current leanings are that creatures who might exist should count in our moral calculus,

Hmm. A long-standing debate in utilitarian theory as you know.

Yes.

Clearly we want to make policy that will ensure the greatest possible happiness for all the beings that exist in the future, even though we are not obliged to bring them into existence.

I know many disagree on this point, but it seems to me that bringing creatures into existence with lives worth living should count as a moral good thing, just as I appreciate others having created me and I think they did a good thing worthy of praise. If so, the prevention of vast numbers of uploads must weigh against policies to greatly increase per-upload wages. But this need not be decisive of course.

It seems like your model in Dawn, if we interpret it as normative rather than descriptive, would fit with "the repugnant conclusion" of utilitarianism that we should create as many beings as possible, even if each of them might have less happy lives, because we will thereby create a greater sum of happiness than by creating fewer, happier beings. Is that what you mean?

The "repugnant conclusion" has never seemed repugnant to me, which is another way I guess I disagree with others in population ethics. But, yes, this upload scenario offers a concrete application of such issues.

In that sense, if neo-Amish humans refuse to become faster, more able uploads their case for accommodation of their decision is weak. But framing all humans who decide to remain organic as undeserving, self-crippers in a brave new uploaded world is part of the political challenge your essay points us to. We need to come up with a more attractive frame for the co-accommodation of organic and upload life.

I don't know if a better frame can be found, but I'd be happy to hear of one.

---

Robin Hanson:

Eugen Leitl wrote:

*AI and whole body/brain emulation is the mother of all disruptive technologies. You may want to regulate them -- but you won't be able to, if they don't want to be regulated.*

That is of course another good reason for first analyzing low regulation scenarios.

As one tries to use regulation to move further and further away from those scenarios, the stronger become the incentives to get around the regulation, and so the more Draconian the monitoring and enforcement process must become.

For that reason it seems hard for me to imagine successfully raising upload wages to be more than ten times what unregulated wages would be.

Unregulated wages could be, say, \$1/yr, putting the upper limit of regulated at say \$10/yr.

So there seems no way to escape upload wages being very low by today's standards.

---

James Hughes:

Robin Hanson wrote:

As one tries to use regulation to move further and further away from those scenarios, the stronger become the incentives to get around the regulation, and so the more Draconian the monitoring and enforcement process must become.

Right now, around the world, there are many countries that have slavery/involuntary servitude, and within the North there are many employers who evade minimum wage laws by paying in cash, or who have unsafe working conditions, or who coerce workers to do illegal things. Lots of people evade paying taxes, and lots of people commit crimes. But the solution is not to simply give up on the notion of law and the regulation of the labor market. It's to strengthen the regulatory capacity and efficacy of the state.

The limits on making the state stronger in a democracy are the willingness to pay for the costs of the regulation, and the tolerance for the impositions on liberty and privacy. This is where I think we should be creatively imagining - and I'm sure many already are - ways that the cybernetics and information tools, and eventually AI's, can detect crime without imposing high regulatory costs. The balance between law enforcement and liberty will still be a problem, however.

For instance, like most Connecticut residents, I exceed the speed limit every day driving back and forth to work. But I've only gotten two speeding tickets in the last ten years. To actually enforce speed laws effectively would, with cops, take an order of magnitude more traffic cops hidden behind berms on the side of the road. No one can afford that. If we had a smart highway and smart cars, or even if each car had a GPS tracker, we could easily detect speeding and automatically impose fines, and some states have experimented with auto-speed tracking lasers that capture license plates and mail out fines.

So, if truly effective traffic law enforcement was cheap, the question before the public would be whether they really wanted to have those laws enforcing those speeds. I suspect that if we really enforced traffic laws we would raise the speed limit to the usual 80 mph on the CT highway. Or we would keep it the same, the state coffers would fill with fines, and there would be fewer highway deaths. Either way, it's a democratic choice.

This is the situation we face now with all the potentially apocalyptic threats. For example, are we willing to create the regulatory and police apparatuses to ensure that we don't end up cracked in a future dawn by runaway AI's and uploads? If the kinds of surveillance and prevention it will take to prevent apocalyptic risks are "Draconian" then hopefully we can have a public debate about what the trade-offs are between security and risk. At least the cost of surveillance and enforcement should come down though, making the consideration of effective surveillance and enforcement fiscally acceptable.

Of course, I say that after the US has just bankrupted itself and weakened domestic liberty on the pretext of suppressing terrorism and chasing chimerical weapons of mass destruction, while actually generating terrorism and seeing nuclear proliferation continue unchecked. So I grant the capacity of democracies to destroy liberties and spend inordinate sums on law enforcement unwisely. Maybe a Friendly AI-on-a-leash would help us make better decisions.

---

Robin Hanson:

James Hughes wrote:

Hanson: As one tries to use regulation to move further and further away from those scenarios, the stronger become the incentives to get around the regulation, and so the more Draconian the monitoring and enforcement process must become.

Hughes: ... This is the situation we face now with all the potentially apocalyptic threats - e.g. are we willing to create the regulatory and police apparatuses to ensure that we don't end up cracked in a future dawn by runaway AI(s) and uploads. If the kinds of surveillance and prevention it will take prevent apocalyptic risks are "Draconian" then hopefully we can have a public debate about what the trade-offs are between security and risk. At least the cost of surveillance and enforcement should come down however, making the consideration of effective surveillance and enforcement fiscally acceptable.

Imagine that the hardware cost of supporting another upload is \$1/yr, but that regulation has increased the legal wage to \$100/yr. Upload John Smith is thinking of starting a new business whose main expense is 10,000 employees. The costs of this business are then \$1,000,000/yr if done by the book. John could instead create 10,000 copies of himself to run the business, in which case his costs would be \$10,000, plus whatever it takes to hide the computers running his uploads. This would clearly be extremely tempting to John.

Presumably John's copies of himself are not going to complain about the arrangement. So to prevent this one might need to inspect every computer capable of running an upload at anything close to the efficiency of computers designed to run uploads, to make sure they aren't running hidden uploads.

Alternatively one might need accurate ways to estimate the number of people that must be needed to produce any given product or service. And one would

have to prevent the existence of "free wage zones," so global governance would be required.

---

James Hughes:

Robin Hanson wrote:

...one would have to prevent the existence of "free wage zones," so global governance would be required.

Here we agree.

---

### **Concluding Remarks by Robin Hanson**

I have long been interested in the social implications of future technologies, participating in mailing lists and conferences that seemed to take such technologies seriously. And I entered graduate study in economics in part because future speculation usually reflects more expertise in physical or biological science, or psychology, than in social science. Over Christmas break in 1993, halfway through my first year of economics graduate school, I first tried to correct that deficit by applying simple economic theory to an important future scenario. I choose the scenario I judged then, and still judge today, to be the most important analyzable future scenario that I can imagine – cheap brain simulators, also called uploads.

The usual economic practice is to first model a relatively low regulation scenario, both because such scenarios tend to be easier to analyze, and because actual regulation is usually light. The next usual step is to compare this baseline scenario against various possible regulations, ideally considering not only the potential to correct market failures, but also costs of monitoring, enforcement, and evasion. The standard economists' criterion for evaluating such alternatives is quite capable of, and often does, recommend regulation and/or redistribution. (U.S. economists are on average politically to the left of the U.S. public, though to the right of U.S. academics.)

My resulting paper, "If Uploads Come First," followed this standard practice. Having found dramatic implications in a simple baseline scenario, I used my limited time to make those implications clear to a wider audience, mentioning only a few possible regulations. As my paper was relatively informal, I did not try to publish it in an economics journal. A few years later I tried some formal modeling in related areas, but learned that economists have little interest here. I was advised that pursuing this would hurt my academic career, even at a relatively eclectic place like George Mason University. Heeding this advice, I focused my efforts elsewhere, and recently received tenure.

Imagine my surprise, when over 12 years later I happened across a book that James Hughes had published seventeen months earlier, wherein he described my paper as a “brilliant paradigmatic example” of “libertarian ... unreformed extropian anarcho-capitalist” and “techno-utopian determinism,” for “1990s transhumanists, who ... do not see redistribution and regulation as desirable or inevitable.” Supposedly, I celebrated a future where “most non-uploaded people would become unemployed,” with “a radical Social Darwinism ... [to] eliminate all the uppety prole uploads, the ones who might want minimum wage laws or unions,” thereby achieving an ideal “division of society into a mass of well-fed plebes and a superpowerful elite” to “rule over” them.

Lord, where to begin? Yes, my baseline scenario implied low wages, human unemployment, and upload selection effects. And yes, of course, I understand that such implications are momentous. If this scenario is realistic, truly great things are at stake here, justifying careful consideration. But I am not libertarian, and I did not dismiss redistribution or regulation, nor forecast a ruling upload elite, nor celebrate negative features of the scenario I described. My paper did not even address income inequality within uploads or humans as groups, and so only considered redistribution from uploads to humans, as an unfortunate but perhaps politically necessary transfer from poor to rich.

I posted and explained these denials, but James replied that since in a recent talk I “did not mention any regulatory or political solution possible to this scenario of general unemployment. ... I don’t think my analysis of your views needs much revision.” Since other futurists have loudly put “equality and social security in the foreground” and endorsed a “Basic Income Guarantee,” James took my silence as revealing my true anti-uppety-prole colors. I suppose since I also did not discuss racism, sexism, or animal rights, James must have also concluded that I am a racist, sexist, and torturer of animals.

I explained to James that human unemployment occurs in pretty much any scenario with many uploads, and that it was not clear to me who James wanted “to take from, and who to give to, ... given the three distinctions of human/upload, rich/poor, and few/many-copied.” James replied that he wasn’t ready to be more specific regarding the scenario I described, but that he wanted me to endorse a general transfer from “rich to poor,” even if we aren’t sure what exactly that phrase means in this context.

To me, however, the details are everything. Yes, of course I’m concerned that people could get hurt, and I want to avoid such harm. But I want to find specific policies appropriate for specific situations, instead of parroting ambiguous political slogans. And the situation really is complex. Not only should we

consider the great good that might come from creating many more lives worth living, but our feelings about inequality are far from simple.

Yes, humans today seem somewhat averse to income inequality, and redistribution can be part of an effective response to that aversion. But we seem much less averse to inequality of sexual, sporting, or artistic achievement. And we seem much more concerned about income inequalities between the families of a nation, and less concerned about the larger inequalities between nations or within families. Given how little we understand about inequality aversion today, it is no small task to project inequality aversion, and its optimal policy response, into a future with unemployed humans, upload clans, preference selection, very rapid growth, and much more.

James focused on aspects of my scenario of concern to the political left, but it is worth noting some other aspects that would seem to be of great concern the political right. Not only are there issues of whether uploads are persons, have moral worth, or threaten human dignity, but there is the jarring thought that in my baseline scenario there may well be far more upload men than women, and almost no children. After all, men now dominate the upper tails of achievement likely to be favored for upload copying, and there is little point in taking decades to raise a child when one can copy and train adult uploads.

So is an upload world a heaven or a hell? What we need is more analysis of what our possible futures really are, and less social pressure to jump to premature conclusions about how to deal with those futures. A metaphorical lynching of the only economist in recent years to explore the social dynamics of upload scenarios, for his failure to parrot political slogans, is hardly the way to achieve this.

---

### **Concluding Remarks by James Hughes**

This very stimulating discussion is a core of what the transhumanist movement should be doing: (a) extrapolating the radical options that humanity faces, (b) making an optimistic argument for a particular set of futures, and c) building a proactive movement to ensure that we create a desirable future. In *Citizen Cyborg* I attempted to do all three of these things, and critiqued Robin Hanson for appearing to argue for what I consider to be a wholly unattractive future, (a) and (b), one which I don't want to live in and which I think most people, if they believed that future would be the result of uploading, would work hard to make impossible by banning uploading.

Through our dialogue I've come to appreciate that Dr. Hanson thought he was only working on the first task, the extrapolation of a possible future. I'm still skeptical that the exercise was not normative since there were so many counterfactual possibilities dismissed in the essay, such as democratic

deliberation and state intervention to affect the outcome. I'm also skeptical because the conclusions were argued to be the most profitable of all possible outcomes for everyone, even those who lose out in this free-market version of the Matrix. But I'll take his assurances that his focus was always primarily academic rather than normative.

So I invite Dr. Hanson to think and write more about how to create an attractive and politically plausible future scenario, one capable of illuminating the public policies we require and inspiring public confidence instead of fear. As a model I would offer Nick Bostrom's essay "The Future of Human Evolution." Bostrom asks in that essay whether we can have any confidence that competitive evolutionary pressures will lead human beings toward a future any of us will want to live in. He offers two scenarios in particular that are possible but unattractive. The first is a world in which we have all outsourced aspects of our knowledge, memories and personality until we become shallow "executive-modules." In the other scenario, which he describes as "all work and no fun," he suggests that agents who jettison all features that give life meaning would outcompete the rest. This seems close to the future described in "Crack of a Future Dawn."

But Bostrom doesn't stop with these two possibilities. He asks further what kinds of policies we might create, and what collective action we might undertake, to prevent these futures, and to ensure the flourishing of people who enjoy the kinds of lives we think are valuable. He suggests that we could ban the kinds of technologies that would make possible these outcomes, but that such bans would be too costly for everyone, since we would have to forgo their substantial benefits. The more attractive alternative is to create a global "singleton" to constrain and guide competition, discouraging trends that lead to the unpleasant future scenario, and encouraging trends that lead to a more positive future. That singleton could be a world government, democratic or despotic, a super-AI, or even a hegemonic moral code. It could be minimalist, but it would have to be hegemonic. I don't think Bostrom's essay quite yet contributes to the goal of building public enthusiasm for an attractive and attainable posthuman future, since most people would rather not live under the rule of despotic tyranny or super-AI, nor will they have confidence in a universal moral code to deter defectors, but at least Bostrom's essay is one step closer.

Hanson's essay is cited throughout Bostrom's paper, and Bostrom's essay is certainly a more careful response to the "Crack of a Future Dawn" scenario than the one I offered in *Citizen Cyborg*, which was a more polemical exercise. But whether presented polemically or academically, we need to move beyond simple extrapolations of possible extreme futures that most consider unattractive. We need positive visions and proactive solutions. Hopefully this dialogue has advanced that project.