Mark Balaguer

Why the Market Value of Free Will is $99.99

1. Introduction: In Ted Chiang's story "What's Expected of Us," a machine is invented that predicts with 100% accuracy what people are about to do. When this happens, people become convinced that they don't have free will, and as a result, a third of the population descends into akinetic mutism—a kind of "walking coma". Maybe that would happen if we discovered that we don't have free will. But I'll argue in this chapter that if it did happen, it would involve a wild over-valuing of free will. Descending into akinetic mutism because you discovered that you didn't have free will would be like descending into akinetic mutism because you lost $100. Or so I'll argue.

2. Determinism: Let's start by asking how we could discover that we don't have free will. The most obvious answer to this question is that we could discover that *determinism* is true—i.e., that every event is *causally necessitated* by prior events together with the laws of nature.[1]

We can get a better understanding of determinism by thinking about pool balls. Imagine that you and I are playing pool and that I hit the cue ball and it goes into the corner pocket. It seems that, given the laws of nature, and given all the facts about the situation—how hard I hit the ball, how much spin I put on it, and so on—there was only *one* thing that the cue ball could have done. In other words, it seems that the way I hit the cue ball *determined* the path it would follow. It seems that the cue ball *couldn't have done anything else*. It seems that the laws of physics, together with all of the specific details about the situation, *forced* the cue ball to move in the exact way that it *did* move.

[1] The characters in Chiang's story discover not that determinism is true, but that *eternalism* is true, which implies that the future already exists. But I think that eternalism is, perhaps surprisingly, compatible with full-blown free will, and so I think we should focus on determinism rather than eternalism.

Determinism is the view that *all* events in the physical universe are like this. It's the view that every physical event is *completely causally necessitated* by prior events together with the laws of nature. Or to put the point differently, determinism is the view that every event has a cause that *makes* it happen exactly as it *does* happen.

If determinism is true, we get the following surprising result: as soon as the Big Bang occurred, it was already determined that the entire history of the universe would unfold exactly as it *has* unfolded. For instance, it was already determined 13 billion years ago that there would be a global pandemic in 2020. Why? Because if determinism is true, then every time something happens, there's only one next thing that can happen. So once the Big Bang happened, the next event was *forced on us* by the laws of physics; and then the next event after that was forced on us as well—and so on, all the way through history. So according to determinism, once the Big Bang occurred, it was an inevitable step-by-step 13-billion-year march to the 2020 pandemic.

And what goes for the 2020 pandemic goes for human decisions as well. Last night I went to an ice cream parlor, and there were 31 flavors to choose from, and I chose chocolate. It *seemed* that I chose of my own free will. But if determinism is true, then once the Big Bang occurred, it was already determined that I was going to choose chocolate. And that doesn't sound like free will. So it seems that if we discovered that determinism is true, then it would follow that we don't have free will.

3. Do-What-You-Want Free Will: There are actually *some* kinds of free will that are compatible with determinism. Consider, e.g., *do-what-you-want free will*, which is just the ability to choose and act in accordance with your desires—or, more simply, to do what you want. This sort of free will is compatible with determinism. To see why, notice that the following two claims could both be true:

(i) Once the Big Bang occurred, it was already causally determined that 13 billion years later, I would have a *desire* to eat chocolate ice cream.

(ii) My desire to eat chocolate ice cream caused me to order chocolate ice cream.

If these claims are both true, then my choice was causally determined by the Big Bang and the laws of nature. *But I still did what I wanted to do*. Indeed, my *want*—my *desire*—straightforwardly *caused* my choice. So, again, do-what-you-want free will is compatible with determinism.

But do-what-you-want free will is not the kind of free will that we could *discover* that we don't have. We know for certain that we do have this kind of free will because we've all had the experience of doing what we want. So if we're worried about discovering that we don't have free will, we're presumably thinking about a *different* kind of free will.

4. Not-Pre-Determined Free Will: Let's say that a person has *not-pre-determined free will* if and only if they make at least *some* decisions with the following two traits:

(i) The decision was made *by the person in question* (in other words, the decision didn't just *happen* to the person; rather, the person *controlled* which option was chosen); and

(ii) the decision wasn't causally determined by prior events.

So, for example, my decision to order chocolate ice cream was free in this not-pre-determined sense if (a) *I* made the decision, and (b) the decision wasn't *forced on me* by prior events—or, more succinctly, if (a) *I* did it, and (b) nothing *made* me do it.

Unlike do-what-you-want free will, not-pre-determined free will is incompatible with determinism. If our decisions are all causally determined by prior events, then we don't have not-pre-determined free will. So we could *discover* that we don't have not-pre-determined free

will. And the question I want to ask is *how bad it would be*, if we discovered that we don't have not-pre-determined free will.

5. When Should We *Want* Free Will?: Let's start by asking *when*, during the course of our lives, we should want to have not-pre-determined free will. It might seem that we should want to have it whenever we're in a decision-making situation—whenever we need to choose one option from a set of possible options. But I don't think that's right; I think we should want to have not-pre-determined free will in only *some* decision-making situations. To bring this point out, let me start by defining two kinds of decisions:

> 1. *No-brainer decisions*: These are decisions in which it seems to you, in your conscious thinking, that your beliefs and desires and reasons-for-choosing pick out a *unique best option*, so that you're *certain* which option you want to choose.
>
> 2. *Torn decisions*: These are decisions in which (a) you feel completely torn between two or more live options—i.e., you have two or more tied-for-best options, and you have no conscious belief about which of them is best (they seem equally good to you)—and (b) you decide *while feeling torn* (because *just choosing* is better for some reason than remaining undecided).

Torn decisions are pretty rare. We usually go through life doing things without really thinking about it. For instance, if you're driving home from work and talking on the phone, you'll probably do all sorts of things without putting any thought into them at all—like putting your seat belt on, and starting your car, and turning into your driveway from your street. But this isn't what *all* of your actions are like. Every once in a while, you come to a "fork in the road," so to speak, and you stop and think about whether you want to pursue one course of action or another.

And in some of these cases, you become completely torn about what to do and then you just *choose*—that is, you make a *torn decision*.

I think we make torn decisions a few times a day. These decisions are usually pretty unimportant—they're usually about things like whether to order chocolate or vanilla ice cream, or whether to meet a friend for a drink or stay home and watch TV. But once in a while, we have to make torn decisions about important things. For example, you might have to choose between (a) accepting a great job offer that would require you to move to a city you hate and (b) passing on the job and remaining in the city you love. And you might have to decide while feeling completely torn—because you might have a deadline that forces you to choose while you still feel torn.

No-brainer decisions, on the other hand, are the exact opposite of this. These are decisions in which you feel completely certain which option is best. Suppose, for example, that I tell you that one minute from now, I'm going to give you either a million dollars or a poke in the eye with a sharp stick—and that you get to choose between these two options. If you're anything like me, this will be a no-brainer decision for you.

Now here's an observation: no one should want no-brainer decisions like this to be undetermined. We should all want the choosing of a million dollars to be straightforwardly caused by our reasons. Wanting this choice to be undetermined boils down to wanting there to be some chance—some non-zero probability—that you will choose the eye poking (or that an eye-poking choice will *occur in your head*, since it's hard to see how it could be the case that *you* make such a choice), *despite the fact that you don't* ***want*** *to choose the eye-poking option*. And no one should want that. What you should want, in cases like this, is for your choices to be caused by your reasons-for-choosing.

You might object here as follows: "We'll get a different result if we choose a less ridiculous case. Suppose, e.g., that Jane is trying to quit smoking but that she badly wants to smoke a cigarette right now. She has to decide whether to smoke a cigarette, and this seems like a no-brainer decision (because Jane's reasons-for-choosing pick out a unique best—namely, not smoking). But surely we want not-pre-determined free will when we're in situations like this."

My response: This is *not* a no-brainer decision—because Jane has desires pulling her toward both options. Indeed, if Jane's current *felt* desire to smoke *feels exactly as strong to her* as her current felt desire to refrain from smoking, then this is a *torn* decision. A no-brainer decision is a decision in which your *whole self*—not just your "best self"—feels certain that one option is best. And my claim here is that no one should want decisions like *that* to be causally undetermined. (I'm sure that some people *do* want to have not-pre-determined free will in connection with no-brainer decisions; but they *shouldn't*—because it's not worthwhile.)

So when should we want not-pre-determined free will? Answer: when we're making *torn* decisions. Why? Because it's in torn-decisions that we find ourselves at a *crossroads*—having to choose between multiple possible paths that feel equally good to us. And so it's here that we want to be able to choose our *own* path, without anything *making* us choose one of the possible paths.[2]

6. What We'd Be Discovering, If We Discovered That We Don't Have Free Will: I just argued that we should want to have not-pre-determined free will in our torn decisions. But it's not clear that we *do*. Here are two possibilities:

[2] I'm simplifying a bit. What I actually think is that (a) there's a continuum of possible decision types here (torn decisions and no-brainer decisions are at opposite ends of the spectrum, and in between, there are decisions in which you're *leaning toward* one option but aren't entirely certain that that option is best); and (b) as we move down the spectrum from no-brainer decisions to torn decisions, our desire for not-pre-determined free will should increase, and it should max out in the limiting case of torn decisions. But we needn't worry about this complication here.

1. *The Libertarian Scenario*: We do have not-pre-determined free will in at least some of our torn decisions. In other words, in at least some of your torn decisions, *you* control which option is chosen, and nothing *makes* you choose in the way that you do.

2. *The Deterministic Scenario*: We don't have not-pre-determined free will. More specifically, whenever you make a torn decision, there are always hidden prior events—events that occur before your choice and that you're unaware of—that completely cause you to choose a specific option. (We can assume that the hidden prior causes are nonconscious neural events that occur in your brain just before you choose; but other deterministic scenarios are possible, e.g., scenarios in which evil neuroscientists send signals into your brain that cause you to choose in specific ways.)

If what I've argued is correct, then (simplifying a bit) to discover that we don't have free will would be to discover that we're living in the deterministic scenario rather than the libertarian scenario. That, I think, would *suck*. I *hope* that we're living in the libertarian scenario. But how *badly* would it suck?

7. How Bad Would It Be?: One point to note here is that we're talking only about torn decisions, and again, we make torn decisions only a few times a day.[3] But the more important point, it seems to me, is that even if we're living in the deterministic scenario, it *feels* like we're living in the libertarian scenario. So even if we don't have not-pre-determined free will, we've got a *perfect forgery*. Suppose that your favorite painting is van Gogh's *Starry Night* and that you have a near-perfect copy of that painting in your house—so perfect that no human being could tell the difference between your forgery and the real painting. And now suppose that

[3] You might want to add here that we make *important* torn decisions even more rarely. But I think the *little* decisions we make are just as important as the big ones—because we make so *many* of them. The courses of our lives are largely determined by the *mass* of little decisions that we make.

someone gave you the option to pay some money to switch your forgery for the real painting—but with the stipulation that you can't profit financially from the switch (e.g., suppose it was stipulated that you can't sell the painting or will it to anyone and that, upon your death, the two paintings will be switched back). How much would you be willing to pay to have the real *Starry Night* in your house, rather than the near-perfect forgery? Maybe *something*. But not *that* much, right? Maybe $100?

Living in the deterministic scenario would be just like having a perfect forgery of *Starry Night* in your house. Every time you made a torn decision, there would be some prior nonconscious neural event that caused you to choose a specific option. But the prior cause would be *hidden* from you. You wouldn't notice it at all. It would *feel* to you like *you* were doing the choosing—and that nothing was causing you to choose any specific option. We *know* that's how it would feel—because that's how it *does* feel when we make torn decisions. And if we discovered that determinism is true—and, hence, that we don't have not-pre-determined free will—nothing would change about how our torn decisions feel. And so in this scenario, if we wished that we were living in the libertarian scenario, we'd be wishing for something that would feel *exactly* the same to us. We'd be wishing for something to be true, even though we wouldn't *notice* that it was true. How much would you be willing to pay for that? Would you be willing to give up a trip to Hawaii for something that you wouldn't even notice? I wouldn't. I'd pay *something*. But not a lot. I'm not an expert at pricing things, but I'd put the market value at about $99.99.

Perhaps you'll object that if we don't have not-pre-determined free will, then we're all victims of an illusion—and, hence, not living *authentic lives*. But we're victims of *many* illusions, and most of them don't undermine the authenticity of our lives. For instance, many of

our ancestors believed that the sun goes around the earth, but this didn't make their lives less authentic. Now, you might respond that this illusion is unimportant—and that it's not linked to our sense of the kinds of beings we are. But if you think that, it's probably because you live in the 21st Century. The belief that the earth is the center of the universe used to be extremely important to people. Indeed, the Catholic Church burned Bruno at the stake for, among other things, denying this claim.

I agree that *some* illusions would undermine the authenticity of our lives. If we were living in a Matrix-type scenario in which everything we experienced was part of a simulated fiction that was being fed into our brains by evil neuroscientists, then we wouldn't be living authentic lives. But I don't think the deterministic scenario—i.e., the scenario in which all of our torn decisions are causally determined by nonconscious neural events that occur just prior to choice—is analogous to the Matrix scenario. It seems more akin to the earth-isn't-the-center-of-the-universe scenario. Perhaps the deterministic scenario lies in between these other two scenarios—partway down the spectrum from the heliocentric scenario to the Matrix scenario. But I don't think it's very far down that path.