Numbers are Weightless

I. Introduction

Should we believe in abstract objects? If our theories commit us to belief in abstract objects, how do we know whether our commitment to them is justifiable? We know some ways of figuring out what exists. But, I will argue, those ways don't seem to apply to abstract objects, or at least the results seem inconclusive.

In what follows, numbers will be my central example of abstract objects. But I think that much of what I say generalizes to other kinds of abstract objects. I will also approach this problem from a naturalistic standpoint. I think we should use the methods of science until we have good reason to believe they won't work for our purposes. My conclusions in this paper could be interpreted as evidence that the methods of science won't work for determining whether abstract objects exist. If they won't work, then I think the next step should either be to come up with another method, or to conclude that the question of whether abstract objects exist is ill-formed.

II. How do We Find Out What Exists?

Here are some tried-and-true methods for determining that some object exists:

a. *We detect it with our senses (possibly aided by some apparatus)*

b. *We detect its effects only*

c. *Experts testify that it exists*

d. *There is a logically sound argument with the conclusion that it exists*

e. *It plays a certain theoretical role in our best theories of the world*

(a) and (b) won't work for numbers, since they aren't physical or causal.

As for (c), the testimony of mathematicians won't establish that there are numbers until we settle what mathematicians' statements mean, which means we need to settle whether nominalism is true or false, which is exactly the question we were trying to answer in the first place. If nominalism is true, then mathematicians' statements are all false, or they're true within a certain fiction or practice, or they're true but not about numbers, or they're merely about what numbers would be like if they existed. According to nominalists, it takes more than testimony to prove that there are numbers. They might well be wrong about that, but at this stage we can't just assume that there is no something else — that's what we're trying to figure out.

Tests of type (d) face a similar problem. Consider this argument: "The number of planets is eight. Therefore there are numbers." Whether that's sound depends on what the logical form of "The number of planets is eight" really is. If nominalists are right, then either it doesn't actually imply that there are numbers, or it and the conclusion are both false. Since we're trying to find a test that will tell us whether nominalism is true or false, we can't assume that this is a sound argument.

In this paper, I will investigate some tests of type (e). I think these tests won't work for numbers either, for a couple of reasons. Firstly, the following argument makes me worry that no amount of information about the physical world will tell us whether there are numbers:

1. If there are numbers, then no matter what the physical world was like, there would still have been numbers.

2. If there are numbers, then finding out more information about the physical world won't help us figure out whether there are numbers. (From 1)
3. If there are no numbers, then no matter what the physical world was like, there would still have been no numbers.
4. If there are no numbers, then finding out more information about the physical world won't help us figure out whether there are numbers. (From 3)
5. Either there are numbers or not.
6. Finding out more information about the physical world won't help us figure out whether there are numbers. (From 5, 2, 4)

If you don't think that numbers exist necessarily if at all, then you won't accept 1 and 3, and I'm not sure if the connection between 1 and 2 and between 3 and 4 is as tight as I think it is, but this is my best attempt at putting my worries in premise-conclusion form.

Another reason to be skeptical of type (e) tests is that scientists don't seem concerned about postulating numbers, or abstract objects in general. Burgess (1998) investigates several examples of scientists rejecting extraneous ontology, and none of that ontology turns out to be abstract. If scientists don't think that commitment to abstract objects is intellectually extravagant, then naturalistic philosophers probably shouldn't either. Still, I think that if we're going to find a test that tells us whether numbers exist, this is the best place to start looking. So the rest of the paper will be spent investigating tests in this category.

Here's a famous one: Quine's indispensability criterion. Something exists if it's in the range of the variables of our best theory of the world. That test does seem applicable to numbers — we use nouns to refer to them, so they're probably the kinds of things that variables can range over — but it's unclear whether they pass or fail[1]. It does seem really difficult to eliminate numbers from our theoretical talk, but it's not definitively impossible. It's also unclear whether passing the test means that numbers do exist, or failing the test means that they don't. In the next section, I'll discuss "easy road" nominalism, which says that eliminating number-words from our vocabulary isn't necessary, because whether or not they're indispensable to our best theories, we needn't be committed to them. Easy-roaders think that numbers should pass a stricter test than indispensability; we'll look at some proposed tests in that section.

If those tests don't satisfy us, Occam's razor gives us another potential route to determining whether numbers exist. Occam's razor, as invoked by philosophers, requires objects to pull their own weight. If postulating abstract objects doesn't make our theory any better, then there's no need to postulate them (or at least, there's no need to be committed to them, even if we keep talking as if they exist), so we needn't bother looking for existence tests that apply to them. But Occam's razor is less specific than we might like. What exactly makes one theory simpler than another? If scientists are right not to be concerned about postulating abstract objects, then maybe they're a kind of object that doesn't burden a theory with extra complexity. In section IV, we'll look at Occam's razor more closely, and see what "simpler" might mean. Before we ask whether

1. Quine's criterion also seems backwards, if it's a test for what exists. Once we've figured out what exists, then we know what our quantifiers range over. But Quine's criterion goes the other way.

numbers can pull their own weight, we need to see whether they have any weight to pull.

III. The Easy Road

If numbers pass the indispensability test but don't exist, then what does it take to exist? Azzouni (2012) says that the difference between what's real and what's not is that real things are mind- and language-independent. This criterion (the independence criterion) is supposed to explain what we do when we try to find out what exists and what doesn't. We find out what exists by using our senses, or things that assist our senses, like microscopes. *Thickly-accessible* posits are things we can observe. *Thinly-accessible* ones are ones we can't observe, but that we have an "excuse" for not being able to observe. Azzouni doesn't give a characterization of what makes an excuse a good one, but he says "my job is not to give a well-motivated and principled account of what passes for an excuse. Such excuses issue from science"[2]; the exact nature of the excuses turns on the details of the relevant scientific theory and background facts about the entities and our access to them.

What usually happens is that we observe some phenomena (so we have thick access to them), come up with a theory to account for our observations, and then that theory gives us thin access to some other phenomena by providing excuses for why we can't observe them. For example, we might observe the way various elements combine to make compounds and decide that the best explanation is that there are atoms with

2. Azzouni (2012), 962.

electrons arranged in such a way that atoms of different elements will combine in a certain ratio. This theory also tells us why we can't, at this point in our theorizing, observe atoms or electrons directly: they're too small. (Maybe our theory also tells us enough about atoms and electrons that we can start building machines that will let us observe them directly.) We now have thick access to macroscopic bunches of atoms and thin access to atoms and their parts.

So why, according to Azzouni, don't we have thin access to mathematical objects? Because our theories say nothing about how those objects relate to the properties we deduce them to have. The objects don't play a role in mathematical proofs. All those proofs would go through, and we would believe numbers have the properties we in fact believe them to have, even if numbers don't exist. Unlike chemistry and atoms, mathematics doesn't study the nature of our epistemic access to numbers; it can't explain why we don't have thick access to them.[3]

This, I take it, is the existence test Azzouni proposes:

- *If the object doesn't pass tests of type (a) or (b), then we check whether a theory that studies the nature of our epistemic access to this kind of object gives us an excuse for why it doesn't pass those tests, and we find that there is such a theory and that it does give us a good excuse.*

Mathematical objects are supposed to fail that test because there is no theory that studies the nature of our epistemic access to them. But what counts as studying the nature of our epistemic access to a certain kind of object? Species could be considered

3. Azzouni (2012), 963.

abstract objects, and it seems like biology studies them and has something to say about how we get evidence about what they're like. If mathematics features essentially in the explanations science gives, then it seems like science *does* have something to say about our epistemic access to (at least some) mathematical objects, insofar as it has something to say about how we know that those explanations are correct. Why is this not good enough for scientific theories to count as saying something about our epistemic access to numbers? Azzouni says that that kind of indispensability only shows that mathematical *theories* are indispensable to science, not that mathematical *objects* are, since the theories would be the same whether or not the objects existed.

But if mathematical objects can't cause anything, it seems like it's impossible for any theory to be affected by the existence or non-existence of numbers. The nature of mathematical objects entails that no theory can be about our epistemic access to them, which entails that we can't have thin or thick access to them, which entails that they don't exist. Azzouni seems to be saying that mind- and language-independent objects *have* to be either thinly or thickly accessible, that our access to them has to bottom out in sensation, and if they're not the kind of thing that can affect our senses in any way, then they necessarily don't exist. I'm not sure why we can't access things using our minds as well as our senses: we detect some organisms and think of them as a species, we detect the way bodies move in the solar system and explain the movements using operators and eigenvalues. We don't get to just make up facts about species or operators, as we'd be able to make up facts about fictional objects, so at least in that sense they're independent of us. Azzouni would probably reply that this isn't good enough, that according to the methods that scientists use to discover what exists, we need some explanation for why we can't observe (in a way that bottoms out in sensation) these things that we think we can't make up facts about. My main worry about his proposal is that the existence test he proposes may be inappropriate for numbers, in the same way that tests of type (a) and (b) were. Maybe it's also tied too closely to testing for the existence of physical objects and forces to be useful for testing abstract objects. After all, we don't even need to gather any data in order to know that numbers fail this test. Maybe it will turn out that non-physical and non-causal objects necessarily fail to exist, but let's look at some more tests before we give up on them.

Yablo (2012) gives another existence criterion. He says that "The nominalist assumes that how matters stand physically is orthogonal to whether mathematical objects exist. Take any world you like — its physical condition neither demands nor precludes the existence of numbers and sets"[4]. Furthermore, he says, even if numbers are necessary to *state* how things are physically, that doesn't bear at all on whether they exist. But there is one way in which numbers could be indispensable to science that might bear on whether they exist.

Yablo characterizes three "grades of mathematical involvement" in a theory, which are most easily defined in terms of a *covering law*. A covering law explains outcome E as arising from circumstances C according to a generalization G. The lowest grade of mathematical involvement is when mathematics is needed to describe E or C.

4. Yablo (2012), 1013.

The second grade is reached when mathematics is needed to explain how E arises from C at the right level of generality (e.g., a square peg won't fit into a round hole, not because of the specific microphysical configuration of the matter making up the peg and the material surrounding the hole, but because of more general facts about the macrophysical shapes). The third grade is when mathematics is needed to provide the covering law itself, for example: the wax walls that make up honeycombs are arranged in hexagons because hexagons are the most efficient regular polygon that can tile a flat plane (i.e., the perimeter-to-tiled-area ratio is lowest when the tiling figure is a hexagon).[5]

Yablo says that examples of the third grade of involvement might count as evidence that numbers are explanatorily important in a way that is relevant to their existence. In order to determine whether that's so, he says we need to ask further questions, like whether the explanation is existence-implying. There are ways of reading mathematical statements such that they can be true when there are no numbers (they merely describe what numbers would be like if they existed, e.g.), so he says that before we can conclude that numbers exist, we need to say what the existence of numbers adds to the explanatory power of mathematical statements. The proposed existence test seems to be something like this:

• *The object plays a role in the covering laws of our best theories, and those laws would have less explanatory power if the object didn't exist.*

5. Yablo (2012), 1020.

Like Azzouni's test, mathematical objects fail if they don't have any impact on our scientific theories. If our explanations are less good if mathematical objects don't exist, then they (probably) exist. But again, since mathematical objects are non-physical and non-causal, their existence or non-existence doesn't seem like it would affect the explanatory power, truth, accuracy, or any other aspect of our theories about the physical world. Yablo himself began by rejecting Quine's criterion on the grounds that the physical state of the world is orthogonal to whether it contains numbers. Orthogonality in his sense needn't mean that the physical state of the world is totally irrelevant to whether it contains numbers, only that it "neither demands nor precludes" their existence. It could make the existence of numbers more or less likely, but not near 100% or 0%. That's an interesting possibility, but I don't think this test allows for it. Since numbers can't cause anything, it seems unlikely that their existence could add to the explanatory power of any theory. If it can't, then gathering more data about the world won't tell us anything at all about whether numbers exist — they necessarily fail this test, just as they necessarily failed Azzouni's. I think it's less clear whether numbers' lack of causal power means that they fail Yablo's test, because it depends on what we mean by explanatory power. Maybe some notions of it would allow non-causal objects to add explanatory power to theories.

Let's approach this problem from another direction: instead of saying what numbers have to do in order to warrant our belief, let's see what numbers have to do in order to warrant our *disbelief*, and then see if they do that. Occam's razor can be read as a test for the *nonexistence* of numbers. Not everything that passes it exists, but everything that fails it at least has to work hard to get back into our good graces.

IV. Occam's Razor

Very roughly speaking, Occam's razor says that the simplest theory that accords with the data is the best one. To make this specific enough to be useful, we'll need to say what makes one theory simpler than another.

Let's say you're making some observations — you're keeping track of the temperature each day — and after you've done this for a while, you want to predict what the temperature will be tomorrow. If you create a model of your observations so far, or several models, you can use them to predict what will happen next. A model can be thought of as a formula that defines a line on a graph. To find a good model, you plot your past observations on the graph and try to find a line that fits them. To predict your next observation, you just have to continue the line in the same way, according to the formula, and see what it says the temperature will be tomorrow. You might model the temperature in degrees Celcius using a formula like this:

$$t = 20 + 0.5d$$

where d is the number of days since you started making observations. d is a *parameter*, a value that the equation depends on. This formula gives you a straight line, starting at 20 degrees and going up by half a degree per day. You could make it more complicated by adding more parameters, for humidity or air pressure or whatever you want. You could also add *higher-order* parameters, for instance d^2, which would make your line a parabola.

Lots of lines might fit the data pretty well, and they don't all make the same prediction about tomorrow's temperature. How do you know which one(s) to use? There are many different criteria that scientists use when choosing a model — the Bayesian Information Criterion, the Akaike Information Criterion, etc. One main difference between them is the penalty they add for extra complexity, which in this context means extra parameters or higher-order parameters. Complexity might not seem like such a bad thing — you're trying to fit the data as best you can, and your data points might jump all over the place, which would mean you'd need a really complicated formula to match them exactly. The reason complexity isn't always good is that the data you've gathered so far isn't perfect. Whatever pattern the temperature is actually following, your thermometer hasn't measured it exactly, and random fluctuations in the weather have distorted it a little. The pattern you're trying to capture with your model is the "signal", and the random fluctuations and inaccuracies are the "noise". You want to filter out as much of the noise as possible and model the pure signal. If your model doesn't have many parameters in it, it won't be able to vary too much (the simplest model is a horizontal line, with no parameters), so it won't fit the signal very well. If it's too complex, it will have too much freedom to vary (it will be more wiggly), so it will fit the noise too well. So if we find the simplest model that fits the data well enough, we've probably got the best model (given the data we've collected so far).

Criteria that are applicable to these kinds of models aren't straightforwardly applicable to theories with and without numbers, though. Those theories aren't complicated curves that we can use to predict all our future observations, and numbers aren't some extra parameter or exponent. When platonists say their theory describes or explains mathematical practice more accurately, they don't mean it literally fits some set of data points better. We're also left with a lot of wiggle room for saying what counts as a good enough fit, and how heavliy to penalize extra complexity.

If we want to get more specific about what complexity might mean, there's always Solomonoff induction, which in an ideal world (in which the halting problem was solvable) would be an algorithm we could actually run in order to evaluate hypotheses, but is sadly uncomputable. Still, it's interesting to see what this idealized kind of induction would require, so let's look at it anyway.

Here is how it works. Solomonoff induction gives us a way of deciding which of our available hypotheses is most likely given our evidence. We can think of competing hypotheses as computer programs that generate predictions about what evidence we'll receive. Uncomputable hypotheses aren't very useful for making predictions, so we'll only consider the computable ones. But we don't want to restrict ourselves too much, so we'll consider *all* computable hypotheses. Our goal is to guess which program is the "real" one, which one is producing our evidence.

We can think of these programs as being written in a certain programming language (one that doesn't make it easier to express any particular thing, so that more complicated or random outputs that intuitively contain more information require a more complex program to output them), and that we translate each one into a sequence of ones and zeros. There are many equivalent ways to express each program; for simplicity, we'll imagine these equivalent ways as tacking extra "junk" bits onto the end of the shortest possible way of expressing the program. (So if the shortest way to express a certain program is as the string '1001', all the 6-bit equivalents of that program are: 100100, 100101, 100110, and 100111.) We're going to use Bayes' Theorem to figure out which of these programs is most likely given a piece of evidence E. Bayes' Theorem says that

$$P(\text{program} \mid E) = P(E \mid \text{program})P(\text{program}) / P(E)$$

where P(program) is the *prior probability* (the probability before we observed E) that that program is the correct one, and P(E | program) is the probability that we would observe E, given that that program is correct. P(E), the prior probability of the evidence, is the same as the sum of $P(E \mid \text{program}_n)P(\text{program}_n)$ for all *n* of our programs, so once we know how to calculate P(E | program)P(program) for any one of them, we'll be able to figure out how likely each one is given E.

Finding P(E | program) is pretty easy: we run the program and see what it outputs. If it matches our observations up until we observed E, then we look at what it predicts will happen next. If it predicts E, then P(E | program) is 1, otherwise it's 0.[6] Now we need to assign prior probabilities to each program. Let's look at two of them: A,

6. Each program is deterministic. Given a certain input, its output is determined. But we can still get probabilistic predictions using Solomonoff induction, because our evidence needn't eliminate all but one possible program. What we usually end up with is a bunch of programs that make different predictions, weighted by their probabilities. If, say, we end up with two possible programs, one of which is 4 times more likely to be correct than the other, and the more likely one predicts E and the other one predicts not-E, then we can give E a 0.8 probability.

which is four bits long, and B, which is six bits long. Let's say A is '1001' and B is '111100'. Now, how likely is it that the universe is running each of these programs? We want to make as few assumptions as possible, so we'll consider all the strings of 1's and 0's to be equally likely. But each program is encoded by many different strings, so it might well turn out that not all *programs* are equally likely. To make things easier, we'll compare the probabilities of A and B, rather than putting a number on each one. Let's say the longest program the universe could be using is 1 million bits long. Of all the million-bit programs, four times as many are equivalents of A than are equivalents of B (because A has two extra spaces for junk bits, and there are four ways to fill in those spaces). That is, there are four times as many programs the universe could be running that would make it agree with the predictions of program A. So antecedently, A is four times more likely than B.

It will be useful later to have considered one more example here: compare A's prior probability to that of A*, which is the same as A but with two specific bits on the end. Let's say A is '1001' and A* is '100100'. Again we get that A* is four times less likely than A, but in this case it's a little easier to see the reason why: A* is just one *specific version* of A. A* makes all the same predictions as A, because it *is* A with two specific junk bits on the end. If we think A* is the right theory, we think that not only will we observe certain phenomena, but that those phenomena are produced in a certain kind of way — moreover, in a way that doesn't have any effect on our observations (because A and A* make exactly the same predictions).

Once we have our priors, we gather evidence and conditionalize on it, using Bayes' Theorem, to calculate the posterior probabilities for each program. Occam's razor as traditionally understood balances two things: *simplicity* and **fit**. **Ceteris paribus** *the simpler theory is probably correct*; don't *multiply entities* **unnecessarily**. Solomonoff induction does that too. Simpler hypotheses get higher initial credences, and hypotheses that fit the data best get higher credences after conditionalization. But Solomonoff induction also says something about what it takes for one theory to be simpler than another: the simpler theory puts fewer constraints on the universe. (Technically, the simpler theory has a lower Kolmogorov complexity, which means it contains less information, in a certain sense of information. But let's stick with the intuitive gloss for now.)

A couple of examples should make this clearer. Let's say we're trying to come up with a theory about a certain machine. The machine is a black box, so we can't just open it up and see how it works. It dispenses marbles at regular intervals, and that's the only evidence we can get. What we want is a theory that will predict the colors of the marbles the machine will dispense. Here are some theories we might want to compare:

(Simple) The machine will only dispense blue marbles.

(Complex) The machine will dispense a blue marble, then a red one, and alternate between them.

Simple gets a higher initial credence, because it takes a shorter program to run it — its predictions can be described more compactly than Complex's predictions. But then we make these observations about the colors of the marbles: blue, red, blue, red. Now

Simple gets an extremely low credence (zero, or nearly zero). But our credence in Complex is increased, because Complex fits the data.

Now let's compare Complex to this:

(Complex*) The machine will dispense a blue marble, then a red one, and alternate between them, and it does this via a mechanism involving gears.

Complex and Complex* will always stand or fall together in the conditionalizing-on-the-evidence stage of our investigation. Any evidence for Complex is also evidence for Complex*, and vice versa, because we can't observe the inside of the machine. We can't ever get any evidence that bears on whether the machine has gears, and the two theories make exactly the same predictions about the only evidence we can get, which is the colors of the marbles. *But* Complex and Complex* don't get the same initial credences. Complex* is more complex than Complex, because Complex* makes an additional claim about what the machine is like. It's less likely that the machine is the way Complex* says it is than the way Complex says it is. Since Complex* had a lower credence initially, and it can never be better supported by the evidence than Complex is, Complex will always be the more likely theory of the two. We'll never have a good reason to prefer Complex*.

So what does all this mean for abstract objects like numbers? If theories that are committed to numbers make the same predictions about the physical world as theories that aren't, then, like Complex and Complex*, the physical evidence we get will never discriminate between them.[7] So if numbers fail Occam's razor, it has to be because

---

7. If number theories and non-number theories make *different* predictions about our observations, then

number theories have a lower prior probability than non-number theories. The assumption that there are numbers has to put an extra constraint on the universe. A computer simulation of a universe with numbers has to be generated by a more complicated program than a computer simulation of a universe without numbers. This isn't quite the way Occam's razor is usually put. The lesson that's often drawn from Occam's razor is that commitment to *entities* (or kinds of entities) is bad, unless the evidence demands it.

But traditional formulations of Occam's razor don't say exactly what it takes for one theory to be simpler than another. Solomonoff induction does. It says that finding a good theory isn't about finding one that adequately accounts for your observations and is committed to as few things as possible. It's about finding the *most likely* theory (or combination of theories) that adequately accounts for your observations. Sometimes the most likely theory is the one that's committed to fewer entities. Let's say I agree with the predictions that the best scientific theories make about the world, but I also think there's a gremlin under my bed. He hides whenever you look for him, doesn't produce any infrared radiation, and in fact has no effect at all on anything. My "theory" makes all the same predictions about our observations as normal theories do, but it's a lot less plausible than it would be if it didn't include the gremlin, because adding the gremlin adds information to my theory — I have to describe the gremlin, and adding that description makes my theory longer. It's less likely that the universe matches the

our evidence will discriminate between them, so both their simplicity and their fit will be important. But for now I'll consider simplicity alone.

predictions of the non-gremlin theory *and* makes that gremlin exist. The gremlin theory is just the non-gremlin theory with a specific set of junk added onto the end.

But sometimes theories with more commitments aren't any harder for the world to satisfy. Let's say you have a theory about heredity, and your theory contains parents, grandparents, great-grandparents, etc.. I come along and say, "My theory is much better! I get all the same results as you, and all I'm committed to is parents." You say, "I can easily prove your theory wrong: here's my grandmother, so grandparents exist." I say "Sure, she exists, but she's not your grandparent. She's a parent of one of your parents." At this point there's no use arguing which of our theories is better according to Occam's razor. It's no harder for the world to satisfy your grandparent theory than my parent theory. Once parents are around, so are grandparents. I can do without a certain word, so in a sense, my theory quantifies over fewer (kinds of) entities. But our theories are equally likely initially, and since they make the same predictions, no amount of evidence will ever tell between them. They're equivalent, so neither has a simplicity advantage.[8]

My gremlin theory, on the other hand, fares worse than your non-gremlin theory because although both of them make the same predictions, the gremlin theory is a specific way of generating those predictions. My theory makes a demand of the world — that it include an unobservable gremlin — that your theory doesn't. My theory

8. If this seems too easy, because we've agreed that 'grandparent' and 'parent of a parent' are definitionally equivalent, then substitute this example: your theory includes rabbits, and mine only includes (proper) rabbit parts, and I'm unwilling to define 'rabbit' in terms of a certain collection of rabbit parts. Still it seems like our theories both describe the same world. It's no harder to simulate yours than mine.

is just one specific version of yours, as A* was a specific version of A and Complex* was a specific version of Complex. *That*, not its commitment to an extra entity per se, is what makes mine much less likely than yours.

The question now is whether numbers are like the gremlin or not. Does adding numbers to a theory add more information? As with the other simplicity criteria, it's hard to see exactly how to apply this to theories with numbers and theories without numbers. Numbers have Kolmogorov complexities, but a theory that uses numbers and a theory that doesn't might well have exactly the same Kolmogorov complexity overall. In fact, they might be the exact same theory. Literally writing out a platonist theory and a nominalist theory in ones and zeros and calculating their prior probabilities isn't feasible. But we can get some useful general ideas from Solomonoff induction and other ways of comparing scientific theories for simplicity. Are platonists really adding extra unobservable entities to the world for no gain in predictive accuracy, or are platonists and nominalists using the same theory, but describing it in different terms? Are numbers like gremlins or grandparents? What exactly does it mean for numbers to exist, and is it antecedently more likely that they exist or don't exist? These are the kinds of questions that Occam's razor really tells us to worry about.

The idea that there might be something wrong with applying the traditional formulation of Occam's razor to numbers is not entirely new among philosophers. Burgess (1998) concludes that what really governs scientific practice is not "don't multiply entities beyond necessity", but something like "don't multiply causes beyond necessity" or "don't posit causes beyond the ones we already know about unless the ones we know about are inadequate to account for the data"[9]. This gets some support from Solomonoff induction. If hypothesis A includes more causal entities than hypothesis B, then A is probably more complicated than B, so its posterior probability will end up lower than B's unless B is less likely given the evidence. But if A doesn't include any more causal entities than B, then it might be that neither is more complicated than the other. Again, though, the key question is not which theory posits more causal entities, but which one is antecedently less likely. Those questions might often get answered the same way, which may be why Burgess thought (rightly) that the causal version of Occam's razor was more plausible than the way Occam's razor is usually phrased.

To see why some non-causal entities don't merit our ontological commitment, recall the gremlin. The gremlin is non-causal (at least in some sense), since it has no observable effects, so the gremlin theory doesn't posit any more causal entities than the non-gremlin theory does. But in saying there's a gremlin, we *are* making an extra claim about what the world is like, and (partly *because* the gremlin is non-causal), we can never be warranted in making that extra claim. Grandparents, on the other hand, are causal entities, although they don't cause anything that parents of parents don't cause, so the grandparent theory doesn't posit any more causal entities than the parent theory does. But, unlike the gremlin theory, the grandparent theory doesn't make any demands of the world that the parent theory doesn't make.

9. Burgess (1998), 210.

So, where have we ended up? The criterion of non-existence that we have so far is something like this:

- *Among theories that accord equally well with the data, the one that makes the fewest demands of the world is the most likely to be true.*

And now of course we have to say exactly what we mean by "fewest demands". Solomonoff induction gives a precise criterion based on information theory, but that's tricky to use in practice, so we'll have to come up with a rule-of-thumb version that's informative enough to tell us whether numbers exist. In fact, we needn't even think that Solomonoff induction points us in the direction of the right answer to "what makes one theory simpler than another"? But we have to answer that question somehow.

V. Conclusion

If we want to figure out whether something exists, we have to come up with an existence criterion and defend it, and explain what answer it gives. If we can't, then maybe there's something wrong with the question we're asking. I've looked at several criteria, and none of them seem quite right for numbers. But thinking about Occam's razor might give us a strategy for solving this problem. There are versions of Occam's razor that don't penalize theories for their commitment to additional entities, which means that platonism and nominalism might be equally simple. That all depends on which simplicity criterion is the right one.

Moreover, maybe Occam's razor as used by scientists isn't quite right for philosophical purposes. Deciding which hypothesis to favor is partly a matter of deciding whether you prefer one that gives you lots of accurate predictions or one that's more likely to be true. Scientists tend to care more about predictive accuracy; when they're assigning priors to all the possible "programs", they don't really care whether the world actually is in some sense a program, or whether our experiences are generated by anything that could be thought of as a program. They're choosing a good model to use, a good way to picture the world as being. Whether that's "really" how the world works under the hood is less important. Solomonoff induction tells you which model of the way the world works is more likely to generate lots of accurate predictions (remember that we originally restricted ourselves to computable hypotheses because they're the only ones that we could use to predict things, not because the universe has to obey computable laws). But maybe philosophers care more about truth, about the way the world is under the hood, and maybe this should lead us to adopt a slightly different criterion for simplicity than scientists would use, and maybe the philosophical version would give us an answer to questions like "Are there numbers?".

Or we could come up with another kind of test, one that has no ties at all to observation. If the nominalist and platonist theories are different, and even if the platonist really is adding extra unobservable entities, maybe that's made up for, not by giving us more accurate predictions about our observations, but by giving us a better handle on something else — like more accurate predictions about what ideas will end up being fruitful, or more insight into the fundamental structure of reality.

The more pessimistic conclusion is: it's hard to find a test for existence that isn't connected somehow to observables. If the existence of numbers depends partly on the state of the physical world, then it's contingent, and that doesn't seem right. But if it has nothing to do with physical things, then maybe there is no good test for the existence of numbers, and in that case maybe it doesn't even make sense to ask whether they exist.

Works Cited:

1. Azzouni, Jody. "Taking the Easy Road Out of Dodge." *Mind* 121.484 (2012): 951-965.
2. Burgess, John. "Occam's razor and scientific method." *The Philosophy of Mathematics Today*. Clarendon Press, Oxford (1998): 195-214.
3. Gauch, Hugh G. *Scientific Method in Brief*. New York: Cambridge UP, 2012.
4. Rathmanner, Samuel, and Marcus Hutter. "A philosophical treatise of universal induction." *Entropy* 13.6 (2011): 1076-1136.
5. Yablo, Stephen. "Explanation, Extrapolation, and Existence." *Mind* 121.484 (2012): 1007-1029.
6. "An Intuitive Explanation of Solomonoff Induction". *Less Wrong*. 11 July 2012. Web. Accessed 21 June 2014.