

[SQUEAKING]

[RUSTLING]

[CLICKING]

HAZHIR OK, let's get started. Welcome to our last class, folks. It's bittersweet. So today, we'll continue talking about your
RAHMANDAD: personal projects. That will be the focus. And I suspect it can be a little bit more open-ended in the sense that this is the last session. What would be most useful for the specific project? Or even beyond that, if you're thinking about a modeling engagement and you want some feedback on it, we can open it up to that type of a conversation.

We don't need to be focused on formulations as we have been in most of the semester. But if there are formulation issues, we can dig into those as well. That's perfectly fine. So that's one part of the class. And I think kind of not the second half, but last third, we can have a little bit more reflections, kind of what you have learned, what are kind of open questions for you in terms of the structure of class we could do differently.

All of those are, I think, useful to have some conversation about. So those are the things I have in mind. But let's see if there is anything else that you folks think might be useful to discuss, cover as we go through it today.

AUDIENCE: One thing I think would be useful is just how to continue our education in terms of getting better from here.

HAZHIR Good point. So we can definitely talk about that. And maybe we can even start with that. We don't need to
RAHMANDAD: prioritize that much. So there is a question of what you want to do with this, that is, how far you want to go on the learning curve. So I think you folks, the folks who have been in this class and have been doing the work now can build models on your own. But it is not easy, and you have experienced that, models that are useful for meaningful applications.

So this is a point where shifting towards actually doing it for real problems that are close to your heart and you care about them makes a lot of sense because it will be hard, but it is not so frustrating that you would give up after a couple of attempts because you see, OK, there is no way I can do it. I think hopefully through this semester, you have reached that point. That's one general path.

For more concentrated, formalized set of exercise type of things, we usually have our PhD students go through all of the challenges in business dynamics, at least the ones that are more challenging, for their preparation for their general exam. So that provides a set of good practice exercises. Many of them are also interesting problems, so you may find the topics are pretty interesting as well.

We also have old general exams that our PhD students have gone through. That's another set of practice exercises. I would say the last couple of assignments that we had in the class, in terms of complexity, we're getting close to what you would have for the general exam questions, perhaps with the exception that general exam questions are a little bit more open-ended.

So I kind of gave very clear set of constructs to use and mechanisms to capture. Those, you need to determine when you're doing a general exam question. So those would be a good next step as well. And if anybody is interested, we are happy to share those resources with anybody who's interested.

There is kind of a whole different area that is connecting models with data, and we haven't really talked about that. There is a good chance the next version of the class that I'll be teaching in the fall, actually, will focus on that. In that space, I think for pragmatic applications, it is very important to bring data into models.

It is both the calibration side, but also thinking about different ways data and models connect with each other, not just for fitting, but also for informing what structures to use and so on. We don't have good, clean text to teach all of that. It's a little bit more hands-on practice. Again, anybody interested, you're welcome to participate in the fall seminar. I'm happy to have visitors.

Beyond that, it is really depending on your use cases. We are, I think, now advanced enough that there is no textbook that I would say, OK, go read that textbook that provides you with the next steps. Looking at good, old models, I mean other people's models and related topics is always very instructional. That's another good area to focus on. Other thoughts?

OK, so let's open it up. Let me see if my Dropbox is happy. Well, I don't want to reboot, so let me get this thing out. And my Dropbox, finally, we have material in this session coming up. So right now we have four model submissions. If anybody else has theirs online, it might be that my Dropbox hasn't yet synced. So that's quite possible.

Oh, my computer decided to restart. It didn't even ask me. So while this is restarting, anybody wants to bring up what they are thinking about, their project, and what they want to do with it next. And even if you haven't submitted something, we can have that conversation. If you haven't submitted something, it's easier to just talk about it.

So I know a few are working on this and related research. So Maria, you are doing kind of research and using this as a little bit of a trying to get something in place. And you're also searching working on it as a research project in the longer term. So we had a look at your model last time. After that conversation, where do you see the project going? And what direction do you feel is most useful and viable?

AUDIENCE: So this was really useful for me. And I learned how to identify the dominant feedback loops driving the behavior of the system. And then I tried to make some sensitivity analysis that I haven't completed yet. But my main problem is with data and how to calibrate it. So maybe I can get more into it.

HAZHIR
RAHMANDAD: Yeah. So definitely, for serious empirical work, you definitely need the connection with data. These days, we have a lot more data. So that's the great news. The challenge still remains in terms of what are the best methods for estimating these models. It's a messy space.

There is no easy template that just follow these steps and you are done. It really depends on complexity of the model, how much process noise we have, which, that is how much of the dynamics are determined by things inside the model and how much are determined by things that are completely outside of the boundary, and you can't observe them that well.

So that makes it harder to estimate these models. But I think if I do the version that is on estimation in the fall, this might be useful for you as well, if you're interested. Good. Maria, what are your thoughts on what's next on this project?

AUDIENCE: I think what's next is, I need to figure out the correct formulations for the existing, and then all the mechanisms I was thinking about. And I need to get the data, documentation.

HAZHIR So you feel a little bit overwhelmed with the complexity of the work that is ahead of you?
RAHMANDAD:

AUDIENCE: No, not overwhelmed. It's just that there are different parts that have their own complexities.

HAZHIR Yes, absolutely. OK, on the modeling side, we have had a couple of back and forths. So you have a little bit of at
RAHMANDAD: least what the underlying stock and flow structure should be. One generally useful idea when connecting models to data is to, first of all, see if you can match the aggregation level of the data that you have in your model. Not always, but sometimes that is a good heuristic.

So if you have data on five different age groups, it might be a good idea to break down your model into five different age groups. Now, if you have continuous age and continuous and individual-level data, for example, it's not worth disaggregating to that level. Sometimes it is, but most often it is not.

But by getting the same level of aggregation in the model, in the kind of construct level to what you have in your data, you are minimizing data loss, basically, and you can capture the measurement process of the data as closely as possible. That's also key. The data generating process should be in your model. So if there are biases in the measurement process, if there are delays in the measurement process, you want to capture those in your model when you are creating variables that correspond to the real-world data.

So always think about, there is a real-world process, and then there is a data-generating process that sits on top of that real-world process that generates this data. And then that data-generating process gives you the observables, gives you the data that you should compare against the simulated values for the same thing. So if you're doing epidemic modeling, there is a data-generating process.

There is a process for the dynamics of the system-- people are getting sick and dying and going to hospitals and so on. And then there is a measurement process, that is, what's the delay for measuring when people got sick and what's the delay for reporting of deaths. So some countries, you look at their data, they had these oscillatory behaviors because they went with weekly cycles in their tracking of the cases.

So if you wanted to match that data correctly, you needed to include these weekly cycles of measurement and, of course, the delays in it to do a decent job. Now, for some applications, that's more important, for some, less, but it generally is pretty important to have that. My wife works in a product management space, and she's dealing with data that is about registrations and sales of coffee making machines.

So there are hundreds of models. There are these numbers. And when you think about the numbers and you're trying to come up with KPIs for how the company is doing, whether the sales is increasing or not, registrations are fraction of sales that gets registered going up or down, you need to actually account for the delays between the time somebody buys something until they register it, because that process takes actually more than a month on average.

And you need to account for the fact that some people return their product and therefore sometimes you may have negative sales. If you don't do these, then the data doesn't make sense. And to do those, actually having a model can help. Having a simple SD model with the stocks and flows without much feedback can be pretty useful. So replicating the data-generating process.

And that's one of the low-hanging fruits for using what you have learned in this course without building a very complex model, just capturing mechanically how the data is being generated that you're using in whatever data analysis you do, and then creating that filter for making sense and inferring the underlying values that you may care about even more. You really care about sales or customer happiness or whatnot, and that is better captured by back tracing from the reported data to the underlying mechanism. Good.

AUDIENCE: So I have a question. What is your advice on how to capture them by mechanisms in terms of people moving between at-risk and not at-risk, and then the air conditioning effect? I'm not sure where to start.

HAZHIR Yes, good point. So basically in Maria's model, there are people who are at risk for heat incidents because of not
RAHMANDAD: having access to air conditioning and people who are at lower risk. Basically, this is largely based on age, if I understand correctly, and to some extent, based on other comorbidities. So that's the risk level.

And then there is whether they have access to AC or not, basically, their house is cooled down or not. That's another dimension. So in the simplest case, you can think about two levels-- low risk and high risk and then AC access, no AC access. So you would have four stocks for people in each of these conditions.

Then you would have transitions either going from left to right, from low risk to high risk, and from top to bottom, from no access to with access. So low risk to high risk, the most basic driver is aging. So if you have a population, you can even think of it as an aging chain with, OK, these are people under 65 and these are people over 65. And you're just modeling the transition between those two if that is your primary risk factor.

Then, if you're thinking, no, actually there is another risk factor, you can capture it in two different ways. One is to think of these population groups, these two population groups, as combining age and the other risk factor. And then the transition would be, how much people are moving based on the age change, but also adjustment up or down based on the other risk factors that these people may have.

AUDIENCE: Also the youngest, the youngest are also part of that.

HAZHIR Very good point. So then, you can do it in two different ways again. You could have it as three stocks for the
RAHMANDAD: youngest, low risk, and the elderly. That's kind of the first and third stock will be higher risk. And the middle one would be the lower risk. I would probably go with that rather than trying to combine the elderly and the very young into a single stock of high risk.

The reason I feel a little bit hesitant about doing the latter is that capturing the flows between low and high risk becomes harder. So I would go with a clean age-based grouping. And you can have actually different risk levels for each age group. That is, 10% of the middle-aged folks may also be at risk, and that's perfectly fine.

So you can calculate the total population at risk using that. Then, you use age to drive the risk profile. And then you will have basically changes in AC in housing as the driver of the other part. So does that help?

AUDIENCE: Yeah. I need to think about it more. And that's on the one hand, in terms of the risk, I need to think also how to formulate. Because the idea in this model is to capture also the fixes that fail effect. So the AC adoption and then how that reduces the risk. And then the excess heat from AC. So on one hand, capturing excess heat in terms of climate change and then excess heat from using aircon. There's a lot of different parts.

HAZHIR Interesting.

RAHMANDAD:

AUDIENCE: But I'm not sure if I want to disaggregate that much the stock slightly.

HAZHIR So it's actually an interesting general question many of you will deal with-- is desegregation helping or hurting in
RAHMANDAD: many ways? So for precision, it usually should help. But for building a model in terms of conceptual model building, disaggregation is not always clear. Early on, it might be actually easier to disaggregate. That is, coming up with good ways of aggregation is not obvious.

You can make cleaner models by disaggregating and then following, for example, the age groups or the categories of AC availability access. If you think different categories of that matters, go for it. So don't be too afraid of disaggregating a little bit more. Because then the formulations come more naturally, because you're not combining multiple processes into a single transition from one stock to another.

You're just capturing one very well-defined mechanism. So aging is very straightforward. You just have a time constant. You're just having first-order movement of people across different stocks. It's much easier to do it that way at least early on. Now, once you are a very experienced modeler, you may feel comfortable combining them and you say, OK, I think this captures the core idea with a lot less complexity, so it enhances communication.

I can get early results faster for academic journals. That is much more attractive because the reviewers can follow what is happening. You can tell a story more easily with a simple model. So there are a lot of benefits from aggregating and having a simpler model. But early on, it is not easy to do that if you haven't had that much experience.

Finally, the computer is back. So let's see. We can start with the models we have here. So Gerardo, we haven't, I think, looked at your model as much. So let's open that one.

AUDIENCE: Yes. Actually, it's really a work in progress. I tried to start the process of designing the model, but it's not complete. So I haven't had so much time this week.

HAZHIR Sounds good. So actually, maybe you can walk us through what you're trying to do and how you're trying to do it
RAHMANDAD: in the model. So first, what you're trying to do, big picture, and then we can look.

AUDIENCE: OK, big picture, I'm trying to capture the mechanism of how startups have an impact on environmental problem. And with the last problem definition, thanks to the feedback that I received, I modelled this with three sources. One is the startups. There are three types of stocks. One is startup that became firms with scale-up rate.

The green startups now have an impact on the economic capital. So the startups have an average economic capital contribution. And when the startup scales up, this contribution grows a lot. And so we have contribution on the field with the other stuff.

And at the bottom, we have the other stocks related to the energy consumption of the firm and of the startup. And the way that I am trying to think about this problem is really close to the coflow of the building efficiency that we designed in the other class. And so this is really helpful for me to really match these two.

HAZHIR So just as we are looking at this, for example, green startup energy consumption needs to have a coflow outflow
RAHMANDAD: that represents the green startup death rate.

AUDIENCE: Yes, yes, yes.

HAZHIR Yeah, just keep that in mind. OK.

RAHMANDAD:

AUDIENCE: Yes.

HAZHIR Basically, for coflows, this is generally the case. For every flow for the main stock, you need a flow for the coflow
RAHMANDAD: that represents that. And most of the time until you know how to do it differently, the amounts that you take out of those using for those outflows should be the average amount of the material that is available per unit of the stock.

So that is this average energy intensity is what should be coming out. Those two are the basics for coflows. Now, the second one, you could violate under some conditions if you are willing to make some assumptions. But, good.

AUDIENCE: What I'm trying to do is basically replicate the same mechanism for green startup and brown startup in firm. I can do this also with subscripts, I think. But at first, I try to draw these stocks just to better see what is happening in broader terms.

And the concept here is that a green startup is a startup that is defined as a startup that is particularly efficient or is working on technology that is particularly good for the environment. And brown startups are all the other types of startups, for example, that are set up in generative AI or all the other sectors that are not a particular benefit for the environment.

HAZHIR Got it.

RAHMANDAD:

AUDIENCE: And so the idea is to capture the total emission generated by the growth of the two type of startups. And my idea is to try to figure out in which condition the total emissions of the ecosystem would go down.

HAZHIR

Yeah, OK. All of that makes sense. I think the basic structure in terms of stocks and flows is solid. Whether you need the economic capital or not is a little bit unclear to me. Just kind of looking at this stock and flow, it feels to me that right now you would have an average economic capital per firm versus per startup. These two will be different, but that's basically showing their size.

RAHMANDAD:

And the change in that size, that is for some reason the green firms might be growing bigger over time, is not the focus of your analysis. So if that is not the focus of your analysis, you could even treat those as parameters and keep your model simpler until you need to change it, until you decide to incorporate mechanisms that are showing, for example, green firms are becoming bigger over time gradually, or something like that.

But it's OK what you have here. It's just more complexity than you necessarily need. And then I think at the heart of your problem is that these green firms and perhaps green startups, some of them contribute to brown startups' energy consumption. Basically, some of them are in the business of enhancing the energy efficiency of other firms. So these could be consulting operations or other types of technology firms that are producing new technologies to reduce the emissions.

So as such, that is where the two chains get linked together. And that's what a lot of the interesting dynamics come from. So you definitely need to highlight that. And then going with the subscripts, I like that as well. But what you have here is a little bit easier visually. The benefit of going with subscripts is twofold. One is that you can have, then, more than just two categories.

You can have green firms that are not in the business of reducing energy efficiency of others but themselves are much greener at the beginning or they're in low-polluting industries. So you can distinguish based on industries rather than these two big categories. So you can easily extend the model that way. And there are reasons potentially to do that. And then there is the parallelism in formulations that gets enforced much more cleanly when you have subscripts.

So that's a little bit more making sure your model is consistent across the two different sectors and always trying to come up with general formulations that allow for that. One other part of your story that you may need to think how exactly it will work is the energy consumption per unit of the startup or brown firm, so per unit of activity of each of these firms. That energy consumption per unit of activity itself is changing over time, partly because of the green startups creating new technologies.

And so that becomes another endogenous dynamic that your model needs to capture. Just highlighting that because to capture that, you need some concept of what is the current state of energy efficiency for, for example, these brown industries, which it might be mining, it might be cement production, and so on. So they might have pretty high levels of emissions at the beginning, but that high level of emission has been going down over time because of new technologies and new focus on energy consumption. So that's another piece of the endogenous story.

So I would say, simplifying the model by making your economic capital more of a parameter and instead making sure you have those two pieces of the endogeneity in the model would make a pretty nice, solid starting model, which is actually a conceptually more transparent model.

And then depending on what you want to do with this, if you're moving towards combining with empirical data across different industries and so on, then you could disaggregate for different industries rather than just having two. But I wouldn't recommend that at this stage, because I think this is clean and well-defined enough that it's giving you a nice result.

AUDIENCE: Yeah, this equation really I used at the bottom of this model. There is the other part related to how investment in green technology in a certain way related always to the other model, to green building, this is really close. And so here, I use the unit of kilowatt for dollars. And I'm not really sure. In this part, I use this energy intensity of capital, used this. I don't know if it's good.

HAZHIR So I like the minimum feasible energy intensity per capital as a stock. That's, I think, a conceptually useful
RAHMANDAD: variable to have, which can then drive the new energy intensity for the new firms that are being built, both brown startups and green startups. So as such, I like the idea. The unit, you have it as kind of kilowatt per dollar or something. It's a little bit tricky.

I don't like dollar, but I know that most people measure capital in dollars and they are comfortable with it, just because when the exchange rates change the value of dollar, the unit of capital is not changing, energy intensity is not changing. Nothing is changing, but it's just that the dollar value has changed. But it is OK. I think I might have had something that was more dollar equivalent. Based on the prices of a given year, what was the value of that capital or something like that.

AUDIENCE: OK, I see.

HAZHIR But it's not a big deal. I think you're fine with using dollar there as well. Yeah, this makes sense.

RAHMANDAD:

AUDIENCE: And the other part is related to my question. I was thinking about how to really connect the probability definition or I called the limitation effect. This effect of green startup, this is particularly important when it comes with the brown firm in a certain way. And I was thinking, if it's possible to think about it like the retrofit effect in the green buildings, but it doesn't make a lot of sense for me. And so I would struggle just to figure out how to model this part in a good way.

HAZHIR So can you tell, what is it exactly that you're trying to model?

RAHMANDAD:

AUDIENCE: Yeah. I'm trying to model two different types of mechanisms. One is mechanism, for example, the green startup developed a new technology that has less energy intensity to use and to sell this technology. In this particular case, a new firm can, for example, imitate the investment of this startup and start to invest in this technology and to use this technology in a particular business. And so in that case, we have a reduction in the consumption for the brown firm because it's starting to use this technology.

That is caused by the startup, the green startup, in a certain way gain the work on that technology. The other one is related to the same mechanism. It's where the brown firm buy the technology from the startup. That is, from an economic point of view, completely different because there is a benefit from the startup. But from a consumption point of view, it is actually the same.

HAZHIR Yeah. So I would say for that distinction, your model is too aggregate to worry too much about that distinction.

RAHMANDAD: That is in a way, you can think about it, OK, a fraction of it is going to be benefiting the startup, a fraction of it is just benefiting the brownfield company by reducing its energy consumption. And if you assume that fraction remains the same or is not dynamically changing, then you don't need to separate the two.

You just use the average for both the benefit to startup, if you want to endogenize that, and for the reduction in energy consumption. So that would be my suggestion. And in that sense, it would be pretty close to what you saw in the retrofit example that we were working on in terms of formulations.

But then you could have it feedback back into growth rate of startups, green startups, based on the rate of adoption of this thing among the brownfields, because that's what is driving the growth of some of these green startups. And you can capture that explicitly and that's fine.

AUDIENCE: Yeah.

HAZHIR OK.

RAHMANDAD:

AUDIENCE: Thank you, thank you.

HAZHIR Sure. This is good. I think you are moving in a good direction. I think the pieces are pretty sensible. This is going

RAHMANDAD: to be challenging still, but it is a challenge, I think, you can manage.

AUDIENCE: Thank you. Yeah, the other thing that I'm trying to do is try to endogenize the death rate and the birth rate of startups. So in this, I think, for the moment it's really too much complex. But I think it's a good part because this opens to some entrepreneurial mechanism that are really important.

HAZHIR Absolutely. So I think this is very important. Have a hierarchy of requirements that you will add in future, but

RAHMANDAD: right now don't worry about them. So exogenous birth rate even death rate is fine for now. But later on, you can incorporate that. And we kind of went through a few of these mechanisms that you can add increment by increment. In a way, maybe your model is already a little bit bigger than what it should be for starting the simulation.

Because you want to have a simulating model as soon as possible so that you can understand what you currently have with all the assumptions that go into it, and then add another layer. When you are not a very experienced modeler, it is risky to have a model that may have two, three pages of structures before you simulate it. Because then once you simulate it, it becomes a unwieldy mess that you don't know how to bring it back under control. So start with something that is under control.

It could be as simple as, OK, I just have a green startup and green firm stock and flow structure and that's it. And then I add the energy consumption per unit. And then I add the other two. And they are disconnected from each other. So each iteration, I make sure it is working, the model simulates well. I understand what it does.

Then add the next ones. So once you add the brownfield, but they are not connected, they are just really the same thing, that's easy. Then make the first connection. That's the impact of greenfield on the rate of improvement in the brownfield or something that you think might be the first priority. And then just add in these iterations. The iterative modeling process, and we highlighted that at the beginning of the semester, is really your friend.

Without that, it very quickly becomes very frustrating. So if you have a big model that has a lot of moving parts that you don't understand, you can spend a whole week full-time trying to deal with it and make sense of it, and it just baffles you and defeats you and it becomes very frustrating. So I see a lot of nodding heads. So that shows you this course has done what it has to do. Yeah.

AUDIENCE: Can ask a question? I think it can also open a discussion here related to in which case a model can really make a theoretical contribution to a field of research.

HAZHIR Yes, very good question. So in this case, first of all, you always go back to what is the current state of
RAHMANDAD: understanding in the field. So if the current state of understanding has no models and has no interaction of greenfield firms and brownfield firms, then as soon as you build these quantitatively and make a connection between the two and show that connection starts to matter and there is some interesting non-trivial dynamic, you are making a contribution.

Now, depending on the audience, how big that contribution is, and how much you need to continue to make it empirically grounded or not, to add five different mechanisms and look at the boundary conditions for the impact of each or not, those are a little bit more field specific and also journal specific. So if you go with a very top journal in an economics type of audience, you need to do a lot more analysis than some of the engineering journals, for example.

The scales are pretty different. It's almost an order of magnitude different amount of work that goes into a typical publication in one versus another. So that is not general. But it is always the starting point is the front line of what is already done. Now, because the types of tools you are bringing into the picture are usually a little bit different-- if you're building a system dynamics model and going into an audience that has not been using models in the past, there is another both positive and negative element to it.

The positive side is that just bringing these tools may provide them with novelty, with something that they may appreciate as, oh, this is interesting. Let me understand it. So that's good. The negative side is, you actually need to do your homework. It's not negative, but it is more work. You need to do your homework to understand how your framework and your modeling work is already explored in the other language, in the other way of thinking, and the other way of doing research in that field.

You really need to understand that and be able to speak about your model in that language. You shouldn't go and talk in system dynamics lingo and talk about feedback loop dominance and so on if you haven't yet understood how they think about this phenomenon using their own vocabulary, using their own tools and methods. That is really the most important part of making a contribution.

Because ultimately, if you are the host in some other field's journal and in some other field's research, conversation-- if you're a guest, sorry, not the host. If you're a guest, you need to follow the host's language. You can't just come and say, oh, I have this model. I think it is great. I don't care to read your journals and understand how you think about this, but you should come and read my paper and tell me how great it is.

That's not how the world works. You actually have to completely reverse it. You need to go read what they have done, how they think about the connection between the green firms and brown firms, then write your paper in the language of that field. This is what we know from all the other people who have worked on the relationship between green firms and brown firms, but one mechanism has been missing from this or three mechanisms have been missing from that conversation which could be important.

Now, I'm going to show why this could be important and under what conditions it is important. And to show that, because it requires looking at the interactions between these things, we actually need tools that might not have been used in this field before. That's how you bring modeling into the picture. So that's one of the things.

Like, you ask John Sterman, he would say, he gets pissed off whenever he sees a title of an article in our conference that says "X, Y, Z-- A System Dynamics Approach." Because that is usually trying to sell the work based on using system dynamics. Using system dynamics per se is not a good or bad thing. It is a question of, does it add value?

And if it does add value for that audience compared to what already they are doing, you need to show that, not tell. Show, not tell is kind of the idea. It's not a sales pitch. It is something that people read and at the end say, oh, this is really interesting. I'm now understanding my own field a little bit better. I'm understanding my own phenomenon a little bit better. OK, let's switch to--

AUDIENCE: Thank you.

HAZHIR Sure. No, I think this is going in a good direction. Good work. So we have anybody else has submitted but is not
RAHMANDAD: seeing their thing here? Let me see. Oh, my Dropbox is finally up to date. OK. So let's this time start with Kathy's model. So you have a little bit of an overall of the model. Good. So you have the low frequency versus high frequency readers. And there is basically, right now an exogenous shift in these numbers. So this is not an endogenous story.

AUDIENCE: Yes.

HAZHIR OK.

RAHMANDAD:

AUDIENCE: Yeah. So the shift is not endogenous. The things that are endogenous are just the social influence impacting the desired reading rates and desired buying rates separately. And then there's an intrinsically-- so I separated out what drives book buying and book reading and social value and intrinsic value. So those are the two.

HAZHIR OK. So let's look at those. So there is book buying and then desired fraction of books. And it's a fraction of books
RAHMANDAD: that are read. That's good. OK. And then you have the social influence coming from reading level.

AUDIENCE: So this is how many books out of max capacity people are reading. So if everyone's reading at their max capacity, then it should be one. And if everyone's reading at half of their max capacity, so the capacity is kind of normalized between the high frequency and low frequency folks right now.

HAZHIR Got it. So there is a capacity utilization. So let's see what this behavior-wise looks like. So if the reading level
RAHMANDAD: goes up right now to-- what is the unit here? Dimensionless? What is the average?

AUDIENCE: It's the percent of capacity.

HAZHIR This is average reading level as a percentage of your capacity to read. And your capacity to read, where is it
RAHMANDAD: defined so that I can make sense of it?

AUDIENCE: It's to the left and right.

HAZHIR Yeah, I got it. Max HFR reading capacity in this case is 10 books per month per person, for example.
RAHMANDAD:

AUDIENCE: Yeah, and then two below.

HAZHIR Good. So then when average reading level goes up, the social influence goes up and then comes down. So what's
RAHMANDAD: the story here with the formulation?

AUDIENCE: So it's formulated with this exponential growth based on the reading level increases it. The k parameter I just set it to 1. Or sorry, if capacity utilization is 1, then that balances it. Then, that's it.

HAZHIR I see. So there is indicated desired fraction of books. No, that's coming from it.
RAHMANDAD:

AUDIENCE: I think, yeah, there's a bit of an overshoot as reading level spikes.

HAZHIR So reading level is going up.
RAHMANDAD:

AUDIENCE: Yeah, very quickly.

HAZHIR Yes. I'm just trying to understand why social influence goes down here. It's a formulation question than
RAHMANDAD: something being wrong.

AUDIENCE: It could be wrong.

HAZHIR So for me, this formulation is already too complex. That is, separate the components out of it. So if you have
RAHMANDAD: social influence strength on HFR, that's fine. 1 minus capacity utilization, if that is the effect of saturation, if that is the idea, then put an effect of saturation out. Because otherwise it's just hard.

It leads to a little bit of confusion for people looking at it. And then the last piece, so this one also has 1 minus e to the power of this thing. It's OK. But that can be a separate effect if you want to have it as a separate effect. So then this thing is this effect and that effect multiplied by each other.

You could keep one of them inside it, but don't keep both of them inside. So then that leads to more buying of books. So this social influence, conceptually, what is it? Is it about reading or is it about buying?

AUDIENCE: The social influence right now is coming from reading.

HAZHIR It's coming from reading, but is it about pushing people towards what?
RAHMANDAD:

AUDIENCE: Oh, pushing people towards buying.

HAZHIR OK.
RAHMANDAD:

AUDIENCE: Yeah.

HAZHIR So you see others are reading books.
RAHMANDAD:

AUDIENCE: Yeah, and then you want to buy the book.

HAZHIR And then you want to buy books. But are you seeing people read the books or are you seeing them talk about it
RAHMANDAD: on social media? That was kind of your starting point.

AUDIENCE: Yeah. So I think I'm assuming that the influence is coming from people who are able to speak about the book. And thus, they've read the book. But that doesn't mean that the buyer will then read the book.

HAZHIR OK. So I think the process right now you have is that, actually, it seems like people go buy it and that average
RAHMANDAD: reading level is determining the social pressure, whereas the story you were telling us was more about, oh, there are a lot of people who kind don't really read the book. They just go and do it, look at a summary of it, and then bring it and use it in their social media to boast that this is a good book or a bad book and we know about it or something.

AUDIENCE: Yeah. And then I started thinking about the influence isn't necessarily coming from the aspirational book buyers. It could be formulated another way. This is just the way that I've done it on this draft.

HAZHIR So it's OK. It's just kind of the story is now a little bit has shifted. So that's why I'm asking whether you wanted to
RAHMANDAD: make that shift consciously.

AUDIENCE: Yeah. Because I was thinking about the people who are influencing probably have read it or at least can talk about it coherently enough.

HAZHIR Yeah. So the thing is, the average reading level is fraction of your maximum that you're willing to read or you're
RAHMANDAD: actually reading for these high versus low.

AUDIENCE: Yeah, and it's combined right now.

HAZHIR Oh, it is also combined. So that adds a little bit even more. Let's see. So this is high frequency readers multiplied
RAHMANDAD: by maximum and low frequency readers. So this should be divided by high frequency readers plus low frequency readers. Now, I suspect you have put them as fraction of population from the beginning. So that doesn't play a role. So high frequency readers plus low frequency readers add up to 1.

AUDIENCE: Well, that's the denominator.

HAZHIR Oh, there is a denominator?

RAHMANDAD:

AUDIENCE: Yes, so that on top is the numerator. And then it is divided.

HAZHIR Oh, yes, yes, yes, yes. You're right, sorry. I was just looking at this second line and thinking that's the

RAHMANDAD: formulation. You have this first numerator. OK, yeah. Yes, and these are the maximum.

AUDIENCE: Yeah.

HAZHIR OK, that makes sense. So this is kind of a social norm, for what fraction of your total you want to be reading. If

RAHMANDAD: that is the social norm idea, then that becomes a very good anchor for this, what fraction you are going to be reading. So then when I go to the equation where it uses this, which is here, I want see it as an anchor.

Whereas, right now, you have it in this complex formulation that is $1 - e^{-k}$ multiplied by average reading level. So this is doing a nonlinear transformation on an anchor that, in my view, should have been just the main thing. I don't understand why I need this nonlinear transformation. Now, what was your thinking when you built that?

AUDIENCE: I think I was just overcomplicating it.

HAZHIR Yeah, but there was something in the back of your mind that you decided to put this complicated equation. So

RAHMANDAD: what was that?

AUDIENCE: I wanted to have the steepness of the social influence to attenuate how quickly we get somewhere. And then I was kind of just rooting around for formulations that have this non-linear shape, because I was thinking that it might approach it in that regard.

HAZHIR Got it. So I think really at the heart of the problem here is that you built an input that was really an anchor, but

RAHMANDAD: you were talking about it and thinking about it in terms of social influence, because you were framing this as social influence. So if you're doing social influence, you probably needed some other mechanism. Because social influence is about people basically talking about the fact that they are reading books.

So you needed to be counting how many books are being read or some version of that, and then multiply that by what fraction of people are willing to put it on their social media, and each person who puts it on social media, how many people will read it, and then when they read it, what fraction of those people get interested in going and buying a book.

AUDIENCE: Yeah.

HAZHIR So that becomes a social influence pathway. But you didn't build that infrastructure to go into it. What you built is

RAHMANDAD: a little bit of a different way that says, oh, people look at their friends and see how much they are reading books, and if they are reading a lot of books, they anchor to that level. And then therefore, based on that, they should decide how many books they want to buy.

AUDIENCE: What's the conceptual difference between those, like psychologically?

HAZHIR Yes, so the cover stories are a bit different. So the cover story of the first one is, I see folks posting about books
RAHMANDAD: and that triggers me to go and buy books. The second one says, I see my friends are spending time reading books and they are spending two hours a day reading books, so maybe I should be spending two hours a day reading books.

So one is about how I think about allocation of my time, the second one. The first one is, what triggers me to go and buy books? So if you go with the first one, I would be worried about buying too many books, because that might easily happen. It doesn't have a bound on that. The time-based one has boundaries of time that you can enforce but doesn't have the exponential growth built into it.

It doesn't have the potential for exponential growth that easily. Because, yeah, the norms are spreading, but usually you're kind of in the middle, whatever is happening. So that's a source of inconsistency that then will make the rest of the model internally a little bit inconsistent on that.

AUDIENCE: That's helpful. So I just need to decide whether it's about anchoring versus--

HAZHIR Yes. And whichever you pick, you can potentially include a little bit of the other one as well. So if I'm anchoring to
RAHMANDAD: what fraction of the time my friends are spending on reading books, then I might get tempted to buy more books or less books, depending on how many times I see folks on social media talking about books. That's perfectly fine. I might do that.

The other thing that is a little bit conceptually unclear to me here, you have these kind of low frequency versus high frequency readers. It feels like this is a built-in difference between two groups. It's like you have a group that is genetically even. It's very hard coded. I'm not saying you're saying that. But it's hard coded.

And then when they decide how much of their time they are spending on reading books, they go and look at the people who are not high frequency readers, and they look at them and say, OK, this person is doing two hours per day, per week, or whatever is the number.

And that two hours is 100% of what they would do given their internal interests or capacity or whatever. So because it is 100%, now, I should do 100% of my 10 hour per week. That's, again, I feel a little bit uncomfortable with making that connection, that strong assumption across the groups.

AUDIENCE: So I should keep average reading level?

HAZHIR So it's a little bit tricky. How you would need to deal with it is not obvious to me. Because my first-order response
RAHMANDAD: is, even if people are inherently very different in their propensity to read books, when they look at the other person, they wouldn't know if you're from the high propensity or low propensity by default.

And they wouldn't adjust for that even if they knew that, oh, yeah, this friend of mine reads a lot of books, and this other friend doesn't read a lot of books. They wouldn't adjust for it as, OK, this is an internal capability, so I should discount it that way.

So I may just go with just look at the average of the hours, not fraction of time, and use that hour as the driver. Now, you may have some inherent interest in books versus those who are not interested. And that becomes another anchor that is my underlying anchor is, I enjoy reading books 10 hours per week and this other person enjoys reading books 30 minutes per week. And that is a separate anchor.

And then you have some kind of averaging between these anchors perhaps, or something like that. So in both of these cases, this is really about thinking about these mechanisms. That is, you see how we are bringing out what the equation is telling about the underlying assumptions. You need to think about those assumptions and make sure you feel those are good assumptions, those are defensible assumptions. That is what is good modeling. Yeah.

AUDIENCE: OK.

HAZHIR OK, very good. Any questions?

RAHMANDAD:

AUDIENCE: Yeah, I think I wrote a little blurb on the upper left, proposed next steps. But I think I am open to feedback about whether I should--

HAZHIR Yeah. So it's really, I think this was more of a problem for this course. I don't think this is strong enough a

RAHMANDAD: problem for a research project, so I wouldn't be spending too much time on this at this point.

This is good you got the learning experience you wanted out of it, but I would probably not spend much more time on that. It is a good idea to try to still improve it based on the feedback, so one more round is perfectly fine. But beyond that, I wouldn't spend time on this. OK. Let's see, let's look at Jason's.

AUDIENCE: I'm back to one flight.

HAZHIR OK.

RAHMANDAD:

AUDIENCE: Yeah. So we do have two separate categories of shoppers that are used in the industry to identify different booking patterns. Future shoppers are trying to book earlier before the flight business comes closer to departure and is willing to pay more.

HAZHIR Yeah.

RAHMANDAD:

AUDIENCE: And as a result, the revenue manager's job is to project some fraction of the seats on the aircraft for that demand that they're expecting to come later. So I built a model that kind of runs forward with that setup. There's no feedback yet on the forecast. That's the direction I want to take this. It's based on the realized demand. But we need to see the next step.

HAZHIR OK. So let's dig in. So you have the leisure shopping. You just have some exogenous arrival of people per day

RAHMANDAD: who are interested in buying that flight.

And then you have the business shoppers going through another much closer to the flight date. Pretty sharp demand going up and coming down here. Perhaps sharper than it is in reality, but that's all right. You can change those easily. Then, this leads to total shopping. And then you have a fulfillment fraction.

AUDIENCE: The idea is the business shopper is flexible, so they're willing to buy down if there's a cheaper pair available. So they'll purchase the cheapest thing on the shelf that's available when they go to book. So you can't distinguish between business and leisure when they materialize.

HAZHIR So for now, let's leave the left-hand side, the demand on one side. Let's go and see what's happening with the

RAHMANDAD: bookings. So you have low-fare and high-fare seats that are remaining and then those that are booked.

AUDIENCE: Yes.

HAZHIR And then there could be movement between low fare versus high fare based on what the airline thinks is the

RAHMANDAD: demand.

AUDIENCE: And they can reallocate.

HAZHIR They can reallocate. OK, good. So for now, let's leave out reallocation as well and let's focus on the booking rate

RAHMANDAD: for low fare. So booking rate for low fare comes from leisure and business. And low-fare bookings are some fulfillment fraction for low fare and low-fare fulfillment fraction is?

AUDIENCE: Just shipping from an inventory. Since its capacity constrained things, there's 100 low-fare seats remaining, we compare that to the desired--

HAZHIR So this is kind of a max low-fare booking rate that is driving it. So that is really the main driver unless you don't

RAHMANDAD: have enough shoppers.

AUDIENCE: Yeah, because if you have seats left, everyone's going to drop in the--

HAZHIR Yeah, so this is really not even incorporating price yet. So low fare is just a fixed price, whatever it is.

RAHMANDAD:

AUDIENCE: Right. But we do have a separate category, shoppers willing to pay more. And if in fact low fare is sold out, then they'll drop in that high fare.

HAZHIR And this is kind of a minimum booking time that drives the max low-fare booking rate. And that gives the

RAHMANDAD: fulfillment fraction. Fulfillment fraction is applying to both of the low-fare and business.

AUDIENCE: Proportionally allocated, yes. So I could redo this.

HAZHIR No, that part is fine. So now that gets the bookings and then the remainder of the high-fare?

RAHMANDAD:

AUDIENCE: Goes to draw from the high-fare pool. If their demand wasn't satisfied from the low-fare stock, they'll look at the high-fare stock.

HAZHIR And then fulfillment ratio in this case goes boom, boom, from 0 to 1, 1 to 0.

RAHMANDAD:

AUDIENCE: One fare's sold out.

HAZHIR It's pretty fast shifting here.

RAHMANDAD:

AUDIENCE: For now.

HAZHIR For now. OK. So then let me go back. So you have the low-fare booking, residual business shopping. This is those

RAHMANDAD: who are not satisfied because they didn't get it.

AUDIENCE: They would have bought a low fare, but there were none left. So now they have no choice.

HAZHIR Yeah. So this is why you have this residual folks buying the high fare. That brings down the high fare and that

RAHMANDAD: takes care of the rest. OK, so I think the basic is making sense. So if you have only two price categories, this is the way to do it. If you have n price categories, then you have a cascade.

AUDIENCE: Yes.

HAZHIR So basically the same idea, but you start from the top, you allocate it among anybody who is interested in that

RAHMANDAD: price range. So you may have multiple demand groups, basically. So you have multiple demand groups and multiple fare groups.

All the demand groups are not interested in all the fares. So you start by allocating it, the first, the cheapest among all of them, then the next cheapest among those who are interested. And usually, you can make the two groups the same number just to keep things.

AUDIENCE: Separate passenger groups as fare groups.

HAZHIR As fare groups, so that you basically take out folks in the second passenger group. The second passenger group is

RAHMANDAD: not going to buy anything other than the cheapest. That's the idea.

AUDIENCE: Right.

HAZHIR And then the cheapest price group.

RAHMANDAD:

AUDIENCE: Right.

HAZHIR Yes. And so on, so that would become-- and then if you do it that way and you do it that robustly, then this works

RAHMANDAD: well. But what you are missing is still is the fact that people don't completely go away. That is, that lowest passenger group may stick around, may come back occasionally.

So you would have basically these two stocks of people who are still buying and people who are already satisfied. So the total demand goes into those who are buying or could at the beginning go directly into the demand. And then when they are not satisfied, they come back to this waiting stock.

And with the time constant, they come back into trying to buy. So that would be the structure. And then the management of fares based on these protection rates is the rest. It's how many of the seats do you allocate to this versus that. And that question is, what is the revenue managers in the system?

AUDIENCE: Right. And that decision is a dual process. There's a kind of output from a forecasting, from an optimization procedure, that says, OK, given the value of these fares and the expected demand that's coming, this is the right number to hold back such that we're indifferent from taking a low fare now to taking a high fare a week from now. But there's also behavioral element too, where an individual can adjust that up or down based on their feelings. And then those behavioral decisions get baked into future forecasts.

HAZHIR Yeah. Yep. OK. So right now the adjustment. you have here is, you have a remaining high fare.

RAHMANDAD:

AUDIENCE: Right now, it's just whatever my expected bookings are. And then I put in a multiplier in case I want to buffer it. If I feel really that a high fare is a lot more than the low fare, then it's probably worth my time to bump up and protect even more seats. This is a very heuristic rule. In reality, they do an expected marginal seat revenue thing.

HAZHIR So basically between the gap between high fare booked versus what you expected to book at this time? Or high

RAHMANDAD: fare remaining versus what you expected to be remaining at this time?

AUDIENCE: Yeah, right now, the forecast is static. So I'm saying I expect 35 units of popular demand.

HAZHIR Well, what does I expect 35 units of high-fare demand mean? A total integral under the curve of that?

RAHMANDAD:

AUDIENCE: Total for this flight. And then every time I get one of those bookings, I say my demand remaining is 35 minus however many I have right now.

HAZHIR OK. And how do I decide that is good or bad at any point in time?

RAHMANDAD:

AUDIENCE: I need more. So in this single flight model, you can't unless you have a comparable booking curve.

HAZHIR Basically, you need a booking curve comparison.

RAHMANDAD:

AUDIENCE: Yeah, so that's why last time I had multiple flights.

HAZHIR Yeah. Well, you can have a booking curve comparison here as well. That's not a problem.

RAHMANDAD:

AUDIENCE: It would be exogenous.

HAZHIR Yeah, it would be exogenous. But the thing is that right now the time doesn't play a role. And that is a little bit

RAHMANDAD: confusing. I expect if I have sold 5 high-fare seats out of the 35, by a day before the flight, I should be really worried. But I have sold 5 high-first seats a year before that, I would be very happy.

AUDIENCE: So if I want an exogenous booking curve, how would I do that technically?

HAZHIR Any exogenous input can do that for you or any of the methods you have here for creating these exogenous,

RAHMANDAD: anything works. I don't know. Yeah, it is a cumulative curve. So basically it is cumulative number of high fares that you have sold in the past as a function of time before thing.

So that is your reference curve. And you're always comparing how many I have booked so far versus that reference curve. If you are above that, you have booked more, you would shift more high-fare seats because demand is good.

AUDIENCE: So I can use this model to generate reference curves and then feed those reference curves?

HAZHIR You could, but you could also think a little bit harder about it and come up with what is the distribution that you
RAHMANDAD: are using for your leisure shopping and for your business shopping that is your high-fare and low-fare demand. So think of them as distributions, think of their cumulative distribution formulation, and use that.

And then create small deviations in the realized demand so that they are initially consistent. So it could even be the same distribution, but on the left-hand side, you can have it as stochastic. That is, the realized demand is coming from the same underlying distribution, but it is a stochastic. So you use a random number.

AUDIENCE: That makes sense. I have the cumulative curves just coming out of the--

HAZHIR Yeah, from the potential.
RAHMANDAD:

AUDIENCE: And then I could build the stochastic. So the stochastic part would come off of the rate and then I just accumulate the stochastic rate. And then feed that as a reference curve.

HAZHIR You can't feed it back as a reference curve, because that thing is what is realized this time. You need the
RAHMANDAD: reference to be in place from beforehand. That's why I'm saying you need to solve that to have an analytical curve, basically. And a probability distribution is usually the best. One that has an analytical cumulative curve will give you what you need.

AUDIENCE: OK.

HAZHIR Good. OK. So I think this is good progress. This is in the right direction. The concepts are working. The part that is
RAHMANDAD: really not yet there is this feedback that is changing the target PL, protected level. It is doing something right now, but what it is doing is not as consistent with the things changing over time.

AUDIENCE: Yeah, it's not dynamic.

HAZHIR It's not consistent with how they would probably think in this context.
RAHMANDAD:

AUDIENCE: Yeah.

HAZHIR OK, good. Anything else that you want to discuss on the model?
RAHMANDAD:

AUDIENCE: I think before I put in more fare pockets, because that's more just technical.

HAZHIR Yeah, that's technical.
RAHMANDAD:

AUDIENCE: Probably the PL feedback and think clearly about the process for adjusting that protection level by comparing the prior year booking. And then one day, if there's a--

HAZHIR So you could also think of this as happening over multiple cycles. So the first time you start with the distribution
RAHMANDAD: that I talked about, the next time you update the distribution based on the realized sales. And this is based on realized bookings, not the demand really. And then the next time you update it further and further. Yeah, so you can do that.

AUDIENCE: That's the direction.

HAZHIR Cool, OK. Great. So again you have research that is related to these topics. So I don't know how much this model
RAHMANDAD: will be useful for those research projects. Don't spend too much time unless it is really at the heart of a project you're already working on. Again, this is useful practice for modeling in Vensim setup, but this is not the easiest way of building models for this problem.

That's why there is a whole kind of literature on revenue management that people build models of different sorts. Actually, maybe we should take a break. I think we're at 10:30. So let's take a break. When we come back, we'll look at the remaining couple of models and then have our general conversation.

OK, let's get back. So we have Ross and Hassan's models we'll start looking at. I have model parameters. These are the analysis that we discussed last time. So just as a reminder for our thinking, we had basically two types of customers, one that were more price sensitive and one that was less price sensitive, and one that was kind of driven by exclusivity and the other that looked at prestige as, OK, if the exclusive group have it, maybe I should have it.

That's the basic idea. And then the firm maximizing profitability was the key idea. And we started with quantity and price decisions for the firm. And with that analysis, the profit maximizing option here comes down at low levels of quantity and high levels of price.

AUDIENCE: Yeah.

HAZHIR That's pretty interesting. And actually, the shape of the curvature of this is also not trivial. So you can increase
RAHMANDAD: quantity supplied and bring down relative price. So there is kind of a whole different range of profitable practices.

Of course, high price and very high quantity is not good for you. And it's interesting that for all of these more quantities, not necessarily good, it of course, depends on that elasticity assumptions that go into your model.

AUDIENCE: It's also about the accessibility.

HAZHIR Can you speak up?

RAHMANDAD:

AUDIENCE: Oh, I think it's also about the accessibility of the good that is observed, like on the street, for example, that affects prestige. I have comparative slides.

HAZHIR OK. So let's look at it. So here you have the prestige as well. And the higher prestige happens, of course, with
RAHMANDAD: lower quantity supplied and with higher levels of price.

AUDIENCE: And then the one picked is prestige at 0. It's just ordinary.

HAZHIR OK, so this is really just to take out the effect of prestige. Good analysis, so that's a nice way of just saying, OK,
RAHMANDAD: without prestige effects, what would happen? And the short answer is, either you pay at high prices with low quantities or high quantities and low prices.

AUDIENCE: I was trying to look at the optimal strategy, which is the black dot. I think the contrast between them was the thing that I tried to see.

HAZHIR Yes. So you were saying the maximum profit happens here versus there. So yeah, I think that shift is also

RAHMANDAD: interesting. And in fact, you could get this as a function of different parameters. Is that what you're doing here?

AUDIENCE: Yeah. So now I'm changing what percentage of the demand is price sensitive. So A is like everyone's price sensitive. And then it drops down to 90% being price sensitive.

HAZHIR So the first one is 1. Next one is 95%. And as you increase the fraction of price sensitive people-- oh, decrease it,

RAHMANDAD: you go up. But the shift is pretty rapid.

AUDIENCE: Yeah. That's why I narrowed it down to 90%.

HAZHIR Yeah. But do you have any intuition why that shift is so rapid?

RAHMANDAD:

AUDIENCE: I'm not so sure about that.

HAZHIR So some of it is, of course, the flat payoff landscape. That is, you can have fairly similar profitability for a big

RAHMANDAD: band. So that is part of the story. In fact, here you can see it very clearly. And therefore, small shifts can shift the optimum quite a bit.

The other part, the more interesting question is whether the price-sensitive customers, the existence of this small group of prestige focused can create a really tipping dynamic. That is, it activates the prestige dynamics. And you have a little bit of it here, maybe. It activates the prestige dynamics in ways that then brings price-sensitive customers and makes the optimal price much higher.

AUDIENCE: Can I ask a question? The profitability, is it percentage or is it the amount of profit?

AUDIENCE: It's the amount of accumulated profit. Yeah.

AUDIENCE: So it is with size, too?

AUDIENCE: Yeah, so it's not comparable across the different panels.

AUDIENCE: Sure, but for one panel it is?

AUDIENCE: Yeah.

AUDIENCE: OK.

HAZHIR Good. By the way, you have generated these writing the code with the equations separately outside of Vensim?

RAHMANDAD:

AUDIENCE: Yeah, I used Python.

HAZHIR That's fine. So just for technical alternatives that you have at your disposal, this is a simple model. You could just

RAHMANDAD: write the code easily online. There is a PySD package that takes Vensim models and turns them into Python code, so that takes the translation work out of your hair.

Now, that doesn't work perfectly for every function in Vensim. Some of the more complex functions, it doesn't have the translation, but it works for most of them. And the code it generates is not optimized in any way. So if you want to do some code optimization, you may need to think about that. So that's number one.

Number two, you have actually built-in Vensim functionality to do all these types of analysis, sensitivity analysis, and generate the data and export the data. You can automate the generation and exporting of the data by connecting Vensim with Python. So there are two different ways to connect Vensim and Python. One is through DLL, which is dynamic libraries.

And that basically allows Python to have full control over Vensim, sends commands that Vensim runs and then gets the results out of it and processes the data. And the other is basically using Vensim script language. So Vensim has a scripting language for generating code to run Vensim and get data and so on. And you can basically tell Python to write that code and then send that code to run Vensim.

Both of those work OK. If you have a more complex model, then translation of Vensim to another language or running the simulations in another language could become the bottleneck, then these solutions are more attractive. So just for what you have done, this is good. I think this is the right choice. You always want to do fancy graphing outside of Vensim anyway.

AUDIENCE: The last one is try to see accessibility affecting prestige. So the first is when the sensitivity to accessibility is 0, so you don't really care how much you see it on the street, for example.

HAZHIR So sensitivity to accessibility is 0 means you don't care if others have it or not?

RAHMANDAD:

AUDIENCE: Yeah, it doesn't affect the prestige. You have more wiggle room in terms of that.

HAZHIR Yes. So what that means then in practice, it brings the quantity a little bit further to the right because now it

RAHMANDAD: doesn't erode the performance. Makes sense.

AUDIENCE: What effects prestige besides accessibility? I'm just curious.

AUDIENCE: For this simulation, it's just accessibility, which is quantity. And then the price, relative price.

AUDIENCE: So when is accessibility is 0, it's just price.

AUDIENCE: Yeah.

AUDIENCE: OK. So that's why it's parallel?

AUDIENCE: Yeah.

HAZHIR Yeah. So this time prestige only increases with price. Does the prestige increase with price or with fraction of the

RAHMANDAD: customers who are of one type versus the other?

AUDIENCE: Oh, it's going to just increase the price. But the response to the prestige is going to have different impacts.

HAZHIR Got it. And then sensitivity to accessibility when you bring it up, you see that shift. It also tightens the range of

RAHMANDAD: profitable strategies, basically, because of lack. Very interesting. Very good. This is pretty nice analysis.

Good job. It's kind of clear. And I think all of the mechanisms make sense, but they are not trivial. So you can see that the model is adding some value, but it is also the right level of complexity. So this is the right choice.

I think, again, perhaps this is not something you want to spend too much time on it. But if you wanted to do more in this direction, some dynamic policy of what would be building a new brand, what would be the pathway to building something profitable and of decent size and impact is an interesting question.

AUDIENCE: I have a question. The last time, you said you can't linear in state variable kind of price policy in Vensim. And I was wondering how you do that?

HAZHIR So the way for that would be, you can say, OK, my quantity that I'm going to produce and my price, either of
RAHMANDAD: them, you pick that. You say it is a function of the state of the system. The state of the system is time and the number of customers I have in each of these two groups. So those are the three variables. The simplest functional form would be a linear one.

Not sure if it's, in this case, the best choice, but it's an OK starting point. I always start with that and then change that. So it would be a constant plus alpha multiplied by the number of customers in category 1 plus beta customers in category 2 plus gamma time. So that becomes your quantity supplied, similarly for price.

Then you do an optimization on those eight parameters that is maximizing cumulative profits over the time of the simulation by changing those eight parameters. And it will find the right coefficients that maximizes your profitability. And those coefficients will generate a dynamic response in terms of the quantity and price over time.

Now those are linear response strategies. They are not necessarily great because if there is any non-linearity that it is better to get a lot of good reputation, but then at the very end, bring down my price or do something that expands my market or whatever, that is not going to be feasible because it doesn't change it easily. So then you can have a second-order factor come into it or something like that.

That's one way to do it. Another way to do it is to say, oh, my price over time is going to be dynamic. And so I define a piecewise linear function that says my price is changing. I put 5, 6 points over time and let those determine my dynamics of price over time and similarly for quantity over time.

And then I just optimize the coordinates, let's say the y values, of those 5, 6 points and then do a linear approximation between those 5, 6 points to generate my price values. That's another one. You can do that inside Vensim as well or elsewhere. That concept is similar. That is an open control policy, open in the sense that it is not getting feedback from the state of the system.

But for a deterministic model like this, it is fine. You can go with an open response policy. It gives the same result as a closed response. That is, your policy is a function of the state variable. It gives similar insights. And in fact, it would be a more nuanced answer when you do it with this over time linear, piecewise linear, because it can be pretty nonlinear.

In fact, you can make it even more flexible. So one way to do the piecewise linear is to set fixed points over time, so time 10, 20, 30, 40, whatever, and get the y variables for each of those and optimize over what is the y value and then do the piecewise linear. The other is to make the x also flexible. So then, if there is a sharp change, it will bring a few x's in that place and do the sharp change.

And then in other parts, it will be more distant because there is not much changing. That's also doable. Again, these are things with a little bit of manipulation of subscripts you can do. Did I mention that I think my 2011 paper in *System Dynamics Review* on openness and pricing, optimizing openness and pricing, has kind of the infrastructure in Vensim to do that in a differential game setting, if it is worth it.

This is all kind of manipulation of Vensim. There is nothing conceptually complex. But that's half an hour of time it can save for you, perhaps. So those are a few alternative ways to play with that. I would say all of that is good practice to get some skills and familiarity with a different space that in system dynamics we normally don't talk about a lot.

But it's pretty important. People who care about designing policies and strategies are always also bringing these ideas into the picture. And it's a useful starting point into that space. Great. Any other thoughts on this? OK. So let's go to Hassan's model. So which one should I open?

AUDIENCE: The second version. So maybe another view, the first one.

HAZHIR OK. So we have low smoking history, population with no nodules; high smoking history, population without
RAHMANDAD: nodules. So you have low and high smoke levels and you have the nodule progression from no nodules, non-detectable, detectable, symptomatic.

AUDIENCE: And then the last four are treatments.

HAZHIR In symptomatic treatment. TX is treatment? So detectable would not be treated, correct?

RAHMANDAD:

AUDIENCE: It will be detected if they're screened, but won't have symptoms.

HAZHIR Yes, but could they be treated or not? So then you would want detectable with treatment as well.

RAHMANDAD:

AUDIENCE: Yeah, so these are who are screened sent to treatment. The two on the right are those who are symptomatic sent to treatment. The two on the bottom right are not screened.

HAZHIR Got it. OK, perfect. But why, so if you have low smoking history population with detectable nodules, they get a

RAHMANDAD: screened and they get treated?

AUDIENCE: So they won't extreme because they're not eligible for screening. Only the highest, so the eligibility for screening comes from high smoking status.

HAZHIR I see.

RAHMANDAD:

AUDIENCE: And I'm assuming that because there's a concept called accidental screening, which means if someone's in a car accident, they'll get screened, I'm assuming that's very small.

HAZHIR Got it. OK. With that explanation, it all makes sense. And basically, the rate of screening is determining how

RAHMANDAD: many people you catch in this stage. And then, that helps. And then the recovery rate here from treatment?

AUDIENCE: So I'm not very comfortable with how I formulated that. So we have in data, we know how big a percentage of people in each one of these four types of treatment will survive or will recover, essentially. But the times are different for recovering or passing. I wasn't very comfortable.

HAZHIR Yeah. So I think one thing that is probably important to have is you need death rates from symptomatic nodule in

RAHMANDAD: both of these stocks before they get treated as well. Not everybody who is symptomatic gets treated unless you're in a country and set up that you really have--

AUDIENCE: Yes, I'm assuming that there's infinite capacity for screening and treatment at this point. But you're right. And I'm assuming that there is no other cause of death but lung cancer in this case, which is a very big assumption.

HAZHIR Yeah.

RAHMANDAD:

AUDIENCE: But I see your point. I think another thing that I can definitely add is not just passing from other causes, but post 80 years, people are not eligible for screening, so they will exit the whole chain at some point in time.

HAZHIR Yeah. So I think the exit rates need to be there. Otherwise, it creates these kind of weird dynamics. Right now,

RAHMANDAD: everything is moving towards right and bottom. And in the steady state, that doesn't make much sense. OK. Good. I think the basic story makes sense. If you want us to dig into any of these, we can. So you have a screening participation. This is as a fraction of people.

AUDIENCE: This is exogenous at this point. I'm hoping that I can make that endogenous based on false negatives, false positives, perception of risk and benefits. That's the long-term plan.

HAZHIR Yes. But the unit is fraction? It's fraction of these folks who are in time to uptake screening, decide to be

RAHMANDAD: screened? OK. This is when they get screened. And then there is?

AUDIENCE: That rate depends on the growth rate of nodules based on the comparison between the thresholds.

HAZHIR Yes. So this is the symptomatic nodule size threshold. So the terms here are a little bit confusing for me-- 1

RAHMANDAD: minus is screening participation. So this is actually tricky. Why do you have 1 minus screening participation in the movement of people from medium size to large-sized nodules or whatever it is?

AUDIENCE: I'm assuming that if they participate in screening, they'll enter treatment. So that will stop the growth of the nodule. And they will not be in that pipeline anymore. They'll move for treatments.

HAZHIR Yeah. But if everybody is participating in the screening, so if your screening participation is 1, what does that

RAHMANDAD: mean? It means right now that the residence time here that is under the time to uptake screening, they will move down to the screened group. But during that time, their nodules don't stop growing. You see what I mean? So during that time, their nodules are actually growing. So the rate would not shut off.

AUDIENCE: No.

HAZHIR This rate would not shut off.

RAHMANDAD:

AUDIENCE: But now it does.

HAZHIR It does. So this is actually pretty fundamental, basic. It is a very common confusion that the mechanism for

RAHMANDAD: moving to the right here is about growth of nodules. It has nothing to do with the screening. So if you increase the screening or decrease the screening, that process does not shut off.

What does change is that more of the people go down and therefore they are not exposed to this movement to the right. But those who are exposed to moving to the right, you should let them move to the right with their baseline rate. So there shouldn't be any impact of a screening on this flow variable. This is pretty important.

AUDIENCE: That's a great point. Follow-up question on that. Will the survival rate or recovery rate of treatments work the same way?

HAZHIR Pretty much so.

RAHMANDAD:

AUDIENCE: Because right now I'm doing that 1 minus 4 for people who are passing. But I'm not very comfortable with that.

HAZHIR So which equation?

RAHMANDAD:

AUDIENCE: Any of those treatment stocks have two outputs.

HAZHIR Yes. So you have the passing and you have the recovering.

RAHMANDAD:

AUDIENCE: Yes.

HAZHIR So right now because you don't have aging or getting out of these stocks through other channels, it is a little bit

RAHMANDAD: confusing. So basically, you are letting people, you're saying once you are starting in this journey, you go all the way either to death or to recovery from treatment, and you have no other choices.

That part makes it harder to say what exactly this outflow is. Is it aging out of the group? Is it fully recovering from cancer? So if you think of it as recovering from cancer, the way you are formulating it here is you say--

AUDIENCE: Let's say recovery proportion times the population divided by the time that it takes to recover.

HAZHIR Yes. So in this case, what I would do is not use two different time constants.

RAHMANDAD:

AUDIENCE: What if you know that the time constants are different, actually takes different amounts of time for, let's say, if someone who goes through treatment or someone who passes?

HAZHIR Yes. So basically the underlying mechanism there is that people are going through different levels of

RAHMANDAD: complications. So you actually have multiple steps. And the further you go to more complicated cases, the death rate increases. So you can have recovery from early stages of complications that the drug helps, you survive, and you're not dying.

So let's think about it without this model. You have three different levels of complications-- low, medium, high. Low complication, there is kind of a rate of outflow that you are completely recovering. You have that fixed rate coming down. And there is no death from low complication. You can only go to medium complications.

Similarly, from medium, the last one you go to it, you die. You don't have recovery, for example. Or you could have a recovery, but it would be very small. So if you don't have recovery from the last one, you just have recovery from the first two, so that is the disaggregation that captures the idea you are trying to capture.

Now, if you're aggregating all of this, you basically have a fraction that is coming out from this end or a fraction that is coming out from the other end. And all you need for that is two degrees of freedom, basically. The way you can implement these two degrees of freedom could be a fraction dying versus the rest that are not dying, which is what you have, and an average time constant for how long is the average residence time in this total three-stock thing.

OK, so you use an average time constant and a fraction dying. Introducing a third parameter doesn't give you anything. It just adds one degree of freedom that is not needed and will confuse you in interpreting or estimating the model.

So that's the way I think about it and why I said you don't need these two parameters, you just need one. It just doesn't give you anything, basically. Yeah. Good. So this all makes sense. Let's see what else you have here. Passing thresholds.

AUDIENCE: So the idea is the decision maker can change that eligibility threshold, make it more inclusive or exclusive.

HAZHIR OK. So that will then change?

RAHMANDAD:

AUDIENCE: Change the rates that people go from the top to the bottom. That's something I haven't done, but there's a nodule development rate and a nodule growth rate. That needs to depend on that threshold.

HAZHIR Why does it depend on the threshold?

RAHMANDAD:

AUDIENCE: If I make the threshold more inclusive, then the average smoking history of the highest smoking history will be lower, and the average smoking history the lowest smoking history--

HAZHIR Oh, that is tricky.

RAHMANDAD:

AUDIENCE: Right now, it's not dependent on that.

HAZHIR Yeah, but you can't actually change the history completely here. Because once you're in the kind of a low

RAHMANDAD: smoking history, you can think of these as some threshold, but that's a fixed threshold. Changing that threshold for whether you are in the top group or the bottom group over time is not going to work here, because those are the definitions of people in these stocks.

This is not a coflow. Once you are in the upper one, you are saying you have whatever, some cumulative threshold of smoking in the past. But that is the definition. You can't meet simulation, come and change that definition.

So what are your choices? One choice is to have more categories so that you have 100, 200, 300, whatever it is, the unit of it. And then if you're changing the threshold, you're going to go from one of these to the upper one or lower one and allow them to be tested.

AUDIENCE: Let's say you have three layers. And then the second layer can be either including screening or not?

HAZHIR Yes, so that's one way. Another way is to say, I want to keep just these two to keep things simple. But given the
RAHMANDAD: changing threshold, I would have a larger fraction of the top row or bottom row included into the eligible group.

AUDIENCE: Right now we're doing 0%, 100%.

HAZHIR 0% and 100%. Now, making that not 0% and 100% becomes tricky again, because as soon as you make that
RAHMANDAD: flexible, now you have taken some people with more or less history into a different category, and you are changing the average history for that group. So it requires more complex formulations because now it is kind of a coflow, that is, you're changing the average of the thing in the stock.

Short answer is, I don't recommend doing it that way. There are ways to do it. And so my model of COVID, if you had time to go into the details of it, you would find that it has an equity that is, think of it as number of symptoms that somebody may have. And that number is changing whether you go into hospital or not, whether you die or not.

So people who have more equity are more likely to go to hospital, more likely to be detected, and therefore whether you are detected or not-- so all of these are a function of equity. So I need to actually move people around with their equity levels changing dynamically. So there are ways to do it. It is tough. It becomes complex.

So it's a nice challenge to do it and try to do it. But also it is not easy to get it right. So those are the two options.

AUDIENCE: My goal at this point is just to turn this into a reasonable model that I can test changes of inclusivity that can be discrete changes in three layers.

HAZHIR Yes, exactly. So with that in mind, I would say go with, first of all, subscribing things so that you can have
RAHMANDAD: multiple layers without worrying about, oh, I have to recreate this whole structure.

AUDIENCE: Subscripted lower layers?

HAZHIR Yeah, at least the movement up to bottom. But I think nodules, you can also subscript as well if you want to. And
RAHMANDAD: then use those discrete thresholds, movement of thresholds based on those discrete categories. That would be cleaner for what your purposes are. And it's easier to explain to outsiders, much easier to explain as well.

AUDIENCE: Can I ask a clarifying question? What's the proposed fix for the nodule growth still happening even after you moved into the screening for a solution?

AUDIENCE: What are you talking about?

AUDIENCE: What Hazhir pointed out.

AUDIENCE: I think the proposed fix was the formulation for the growth. I'll take the 1 minus the screen participation out. Independent of whether or not they're screening, the participating in the screening, they will move to the right.

AUDIENCE: But how can they move to the right if they've moved into the screening?

AUDIENCE: Those of them who are moving down to treatment, they will. So whatever's left--

AUDIENCE: Oh, you'll have another. You'll have a flow coming--

HAZHIR
RAHMANDAD: These people are coming. So for this group, you have people from high smoking history, detectable nodules to screened.

AUDIENCE: Yeah.

HAZHIR
RAHMANDAD: Once they are screened, they are not moving to the right. They are not getting worse. Or at least that's the assumption right now. You could easily have the assumption that they do get worse, but much slower rate because they are under treatment. Some people would not respond to treatment, so I think it's pretty reasonable to include the possibility of moving to the right but not with the same rate.

AUDIENCE: OK. I think I misunderstood. I thought you wanted to account for growth while they're waiting for screening.

HAZHIR
RAHMANDAD: Yes, it is being accounted for. But I was saying that this screening participation should not go into this rate. Because the movement to the right has nothing to do with screening. The movement to the right has only to do with the growth of the tumor. As long as you are in this group that you have not yet been detected, you are exposed to the risk of growth in tumor.

It has nothing to do with screening at that point in time. The only conditions that we have a fraction moving this way versus that way, our conditions that you can imagine it to be at the end of your stay, you have two different pathways. You die or you survive. Or you go to the upper grade or you go and retake your class or whatever.

This is, you have a bifurcation of pathways at the end of a decision choice possibility. Those are the places that we will have two parameters going into both of them, one for the time constant, how long you stay, the other for what fraction go right or left. This is not one of those.

OK, good. So I think you have the pieces and I think you understand what you need to do to get it ready for now exposing it to data or whatever you want to do with it. In terms of data availability, which of these stocks are in the typical datasets you might be looking at?

AUDIENCE: Data availability, so technically, I have access to anything NCI has because we're part of the NCI Modeling Consortium, but I haven't seen much time series data. They have a lot of these large-scale cross-sectional datasets, which I'm not sure how I can bring in here. But a lot of the parameters that I've left here, I have values for them at the national level or state level, stratified values for them. So a lot of them I don't necessarily need to calibrate.

HAZHIR
RAHMANDAD: Yeah. So I suspect you would have some time series like death at the end of these. Those, you usually have time series. Screening and finding new cases, maybe you have. I don't know.

AUDIENCE: I think we do.

HAZHIR Yeah, that would be my guess, at least nationally you may have.

RAHMANDAD:

AUDIENCE: Capacity constraints on that we're not looking in here, so that may be different.

HAZHIR Yes. So the capacity can be built into the time to uptake screening. Basically, you are combining two things. One

RAHMANDAD: is the demand side, how often people want to go and screen. The other is how much availability of screening is. You just turn both of them into a time constant, and that time to uptake screening can be state specific, can change over time.

So in a regression mindset, you can say there is a time effect and there is a location fixed effect. You combine those two to get a time to uptake screening that can vary over time and across the states. And that does a decent job. So in that sense, you would have a way of doing that nuanced analysis you may need on that. Good.

AUDIENCE: I will need to endogenize the participation at some point. And if you don't mind, I'll reach back to you again with whatever formulation I come up with. I have the luxury of working with the clinical team who may know or have an insight of what's going on in there.

HAZHIR Yeah. So participation rate right now, what is it reflecting? Is it the people deciding to get a screened? Or is it the
RAHMANDAD: authorities that set the level?

AUDIENCE: No, it's the percentage of those who are eligible that show up to screen.

HAZHIR Yes. So that part is largely related to this time to uptake screening, the way I think about it. In fact, if you don't
RAHMANDAD: have this screening participation feeding back into growth of nodules, these two parameters become superfluous. You don't need both of them.

You just need one degree of freedom for each of these. Now, you may have different parameters for those with larger symptomatic nodules versus detectable nodules. So it may or may not, but you could separate two different parameters for those.

But it also depends on how much resolution you have on being able empirically to say what fraction of people are in the detectable versus symptomatic phase. If you have that, then it makes sense to separate them and have these two parameters. If not, you could even combine those two groups.

AUDIENCE: Can I add that those are not the only two stocks who are going through screening? The other two stocks to their left will experience screening, too. It's just their screening will be negative, the results. So they won't move their stocks. So there will be a number of--

HAZHIR Negative cases?

RAHMANDAD:

AUDIENCE: --false positives included, essentially, or eligibility false positive. So if eligibility is 100%--

HAZHIR Yes, yes, yes, yes. So from your total tests, those who are smoking but don't have detectable anything. So in

RAHMANDAD: terms of fraction of tests that get positive results, you need to capture that with whatever screening criteria is. So that's absolutely the case.

And this kind of goes back to the issue that we talked about earlier today on capturing the data generation process. When you go into the field and you ask yourself, OK, what do we have data on? We may have time series data on how many screening has been done in the US over the last 40 years.

So in your model, what does that represent? It represents all the people in all of these stocks who are getting screened. These are high smoking history people. They would get a screening, assuming some age requirement and so on. So how do I capture that? I definitely need to account for these guys as well, even though they may not get positive results.

And then I may have something about the number of people with detected cancer at any point in time. That is the sum of people here and here and here. All of those need to be counted. And even here, those four, so that's the data generating process and so on. So that's the idea. Great.

AUDIENCE: Thank you.

HAZHIR Sure. Anybody else on the models for today? I think we covered everything we had here. So maybe if not, we can
RAHMANDAD: go to our last conversations, just a little bit of reflections.

So this is just open conversation. I want you to reflect on what you have learned in the semester, what might be tips that are useful for others. I think, hopefully, this course will be online soon and others can also benefit from your thoughts and reflections. Also, what we didn't cover that could be helpful as you are building your skills and capabilities, what you feel is still not there and you need to build that.

And also any feedback on the class structure, how we could do things differently. I don't know when would be the next time we offer this version of the seminar, but hopefully every two or three years, we will have another offering related to this. So that would be useful for the next time. So this is open floor. Any thoughts are very welcome.

AUDIENCE: Can I just mentioned that, for me, there's been what I've learned and what has changed in my habits as a result of this.

HAZHIR OK, good.

RAHMANDAD:

AUDIENCE: There's a lot of stuff that we learned, theoretically, which I may come back to some of the notes and then revisit whenever I need them. But as habits that changed, I wasn't as careful as I am right now about units. I wasn't as careful as I am right now about different versions, smaller versions of the model. These two specifically stand out as now, whatever I'm doing with modeling, I will stick to these two.

HAZHIR That's actually a very good way of thinking about it, the habits that might have changed in how you tackle it. So I
RAHMANDAD: think that's a nice reflection maybe to add to the list of questions. Good. Other thoughts? Did you find others kind of similar changes in habits, other dimensions that were more salient for you?

AUDIENCE: For me, it was the change of pace or maybe getting used to having to build the model quickly. I think that was reduced for me. Maybe, otherwise it would have taken me much longer than a week, let's say, to sit down.

HAZHIR I see, I see, I see. So a little bit of a push to actually get started and build models was sometimes helpful for
RAHMANDAD: some folks. I remember when I was learning modeling, I took System Dynamics now almost 30 years ago when I was an undergrad in Iran, our teacher, our professor, was a graduate of the program here, actually. And he pushed us to build maybe 20 models over two semesters of introductory course and a follow-up course.

And I feel I learned a lot of the modeling that I learned because of that being under pressure and trying to build many models. And all of those headaches that you have experienced this semester, that was where I kind of saw, OK, I have to struggle. And it's hard to learn these skills without going through some of that struggles. Now, the less struggle, the better. But you just can't completely avoid those. You're not learning if you're not hitting some dead alleys and wrong passes.

AUDIENCE: Especially at the beginning, we all were failing at the same core conceptual things. So taking the time to maybe do the first model together, like we could all come up with our own, and then together think about those subtle issues with modeling. Or at the very least, see an example of how it would be correctly done.

It's like, OK, well, we all failed in doing it. Then, it's like, if none of us actually figured it out, then it's like more helpful. I don't know if anyone actually went back and sat down and thought about how to do it correctly, but assuming you don't, then you lose that.

And then the other thing being, this might just be me, because I guess I probably haven't taken as many analytic courses as you guys, but I didn't really know how to represent formulations that were very, very basic, like adding or multiplication. And maybe covering a little bit more like, what do all those, smooth delays, when are the right times to use all of those functions.

As someone who's never actually taken a class, I understand how to do that. It might have been helpful, so moving forward, I wouldn't be like trying to do some weird variable modeling when I could just use smooth variable.

HAZHIR Yeah, that's a very good point. And I think kind of a little bit of a challenge for a class like this is exactly not
RAHMANDAD: everybody comes with the same background of what is covered in their current level of knowledge. And in a way, I kind of assumed everybody knows the basic molecules, that they have a lot of these components. But we have never taught those even in our 873 class. So you're absolutely right, I think.

The other point about going through a model that is well built is also a very good one. I tried for the first half of the class to create a version of a complete model, but after the class. And then the later, I haven't had the chance to do that. So maybe I should try to spend some time and build those and have that in the resources that come with it.

AUDIENCE: I thought that was useful. I did go back through, particularly for some of the specifics. I remember we had a very long discussion when we did the bigger SIR model about the vaccinated, indicated protection situation. I definitely looked back at that at different times, because I think that was a key concept that we used for a lot of them, in terms of creating stocks at the right time.

And I thought that was very useful. And it felt more tangible to be able to look at something like that than the molecules website we covered briefly in the beginning, not necessarily because it was like in a context we used. But the SD molecules website, I almost didn't know what to go look for sometimes.

HAZHIR Yeah. And for that, there is a way to systematically learn the molecules. And that would be basically go for each molecule, go read the equations, then build it yourself, then make a few examples of it. What are the use cases where I can use this structure? But that's a big investment.

If you do it for even the 10 bigger ones, because some of them are almost trivial, but there are about 10 that are big enough that you need to actually go through that process. And then you need to simulate it. It's not just the examples, but understand.

Each of those may take two or three hours at least. So it's a full week of work to just go through the molecule systematically. I would recommend that for somebody who wants to build models regularly, but maybe it's not the most efficient way for our class. So that's, I think, pretty fair.

AUDIENCE: I'd like also, maybe to have a look at the dynamics a little bit more. Because I think we looked a lot at the structures.

HAZHIR The structures, but we didn't do the dynamics. You're absolutely right. I think that's probably one thing that in the overall balance, we should have moved a little bit more towards looking at the dynamics. Because I think formulations need to be there. But even once you have the formulations, there is a whole process of making sure the model is in the right parameter space, the part of the parameter space you're targeting with your baseline.

If it is very off, even if all the formulations are correct in the sense that the final model will not change at all, still, you will have weird behaviors that you can't make sense of and you may get disappointed. And so just shifting the model parameters in the right range is a skill. But also during that process, you usually detect a lot of formulation issues that, OK, this formulation I taught makes sense, but it actually doesn't. And we didn't go through that source of feedback.

So I think that's definitely important. And for our PhD students who are taking general exams, they have gone through that suffering more because going through your example general exams, you will learn that it is not just coming up with good formulations, it is also getting the model to behave in the space that you want it to behave and replicate the basic reference modes that you're trying to capture.

AUDIENCE: The time that we did do that with the predator-prey model was very helpful in terms of figuring out how to expand the model.

HAZHIR Yeah, yeah. And that has actually two important components. And I think breaking that down into detecting what is generating the behavior, we have gone through that two or three times in the class, fewer times than we should have. But I think it gives you a feel for it with the causal tracing, come up with a hypothesis, testing that hypothesis. That's really needed.

If you don't do that, you can't shift your model towards the right direction. This is very important. I think the most common failure mode, once you have even a decent model in terms of formulations, is that you want to shift it in the right space of behavior, but you don't know how to do it, so you end up just ad hoc parameter changes. And that is groping in the dark.

So think about even the simplest models we have, they have a dozen parameters. So you're walking in a dozen dimensional space. If you do a random walk in dozen dimensional space, it is hopeless. There is nothing you can do that makes a meaningful change. It's a completely hopeless process, so you definitely need to narrow it down. What is generating the current behavior?

Based on that understanding, what are one or two parameters that really matter because they are shifting the behavior away from what I wanted to reflect? And let me focus on those two parameters or two equations, shift those, see the outcome. You need to do this hypothesis generation and testing or it will be very, very hard. Good. Other thoughts?

AUDIENCE: I took 873 and 879, so I think it would be more helpful if there was some more--

HAZHIR Could you speak up a little bit?

RAHMANDAD:

AUDIENCE: Oh, I took 873 and 879. And it'd be nice if there's another course, like an intermediate level, where they do all dry structure and like molecules thing. And then this would be more like once you have some of structure, how do you use it like downstream.

HAZHIR Yeah. That's fair. So we used to have a course that was called 875, which was actually that exact intermediate
RAHMANDAD: course. In fact, we had 876 as well. It's just because the structure of courses here at Sloan have moved towards whatever more MBA students take them survive and whatever fewer MBA students take die, we haven't been running those versions. And basically 879 has become the course where we iterate through different variants of modeling.

But we really don't have an easy bridge between the two that is kind of that. And you're right on that. Now, the versions that are kind of doing a little bit of this is basically project-based versions, that rather than doing 10 different models during the semester, you're doing one model, but you get a lot more feedback on that so that you build that model, or maybe two, over the course of the semester.

But most of the students don't get the opportunity to take two of those, except for our PhD students in the C group. Most others don't have the time or interest in taking two different versions of it at the same time. So when I was a student 25 years ago, 875 and 876 existed, and they were kind of distinct courses. And one was based on modeling with projects with real-world clients.

And it was a kind of a lab-based type of a setting. You would work with the client to help solve a problem using system dynamics. And then you would have an analysis course that was more about understanding model behavior and basically all of the things that we discussed on how to understand. That was a whole course. There is enough complexity there that you can do a full course on it.

AUDIENCE: I think this is a good way of getting a lot of those skills through a different program every week. The thing that I become uncertain of is just what don't I know that should be basic, you know?

HAZHIR Yeah. It's a good question. So I think what you don't know that you should know is when you start modeling right
RAHMANDAD: now, what will be the bottleneck for you? And my guess is from looking at the projects, some of it is that we didn't spend that much time on conceptualization, that is, starting from a messy problem to a well-defined problem definition, which we tried to do with the personal problems.

I think we did some of that. But each of you got one data point there for your own project. You got some feedback. You tried to clarify it. If you had received 10 of those, those steps of the process are, I think, pretty important. They are even sometimes more important and more complex than the formulation parts. But they are also harder to pin down into an algorithmic way of thinking about it.

So for that, I think even looking at what we have done with other students, that is, look at the first version of the problem definition. What is your critique? What is the critique that I offered? What was the last thing that they got to? That is helpful. You need feedback. You need to put your stake in and say, this is what I think they should do, and then get an alternative view of it, how I critiqued it, how that finalized.

That is one way to get that type of feedback. But we didn't have time in the class to do it. So that would be, I think, my first place. I think on the formulation stuff, if you have a well-defined problem, actually, I was pretty happy with, by week 9, what you folks were doing in terms of the modeling of well-defined problems, well-defined in the sense that the structures are relatively well known from the problem definition and the feedbacks are relatively clear.

You started to build models that were close. Comparing SIR example, that we had in the first week and the growth of a professional field we had in week 9, the gap was big. It was a much simpler. SIR is much simpler model compared to the one we had. And you guys were doing much better with the second one. So I think I could see that learning was happening pretty well, maybe better than I expected, given where, on average, the classes started.

And SIR, I think we had a few quite a few problems in everybody's submissions. But if you start with a new problem, the conceptualization to get it to a well-defined problem will remain still complex. We didn't do anything with data, and that is a whole big additional thing that one needs to bring into modeling for real-world applications. And we don't have time to do that in one iteration.

AUDIENCE: One of the things I thought was helpful too, honestly, about the past few weeks, is that I feel like going through everybody's models with different problems is really helpful in terms of getting into the process of, oh, it actually is totally possible to look at an existing model that I do not understand at all yet, and come to understand both what the actual problem is and also the formulation surrounding it.

HAZHIR Yeah, good. I'm glad to hear that, because I don't know how easy to follow it is for you when you are seeing one
RAHMANDAD: of these new problems. So it is always a risk that we go too fast. And then half of the class is looking at this like, what's happening? I don't know what this model is about and why.

And I have had this challenge of balancing how quickly we want to do things. So I kind go very fast into feedback, boom, boom, boom, boom, versus trying to get everybody in the class connect with that feedback and compare that with what they think about it. So it has been a challenge for me to balance that.

And I'm actually curious what everybody's response to that, what we did. Was it on the too fast side? Or was it OK, speed-wise? I don't think it was too slow. And not only for this last three weeks, I think even before that, there was always this balancing act of trying to go through as many models as possible and give as many pieces of feedback, but not overwhelm folks or "I didn't figure out what was being said" type of response.

AUDIENCE: Sometimes it was confusing, but I think usually the way you started out was you spent like 30 or 40 minutes on the first model, that was like slow enough to catch up. And then by the time you were on model 5 with the same problems and terms, it was not as bad as when it started out.

HAZHIR OK. That's also a good reminder for me to just use that first model to just settle into seeing the big picture. And I
RAHMANDAD: think by week three or four, we started looking at the structures of the models first, so that kind of big picture is relatively clear on how the stocks are cut.

I think Jay Forrester always used to say, cutting the stocks, picking the right stocks is half of the modeling. It's really the hard part of the modeling. So if you do that well, the rest is much easier to get into the right place. And I think that is a very good observation generally.

AUDIENCE: I'll ask question that is on my mind, which is the class was good for figuring out how to represent certain things, but what to represent is still kind of an open question.

HAZHIR Yes.
RAHMANDAD:

AUDIENCE: And there was an obvious transition in the problem statements you gave us, where weeks one through four or five, it's pretty clear what we had to represent. And then we all practiced how to implement that and learn the power of the molecules and linking them up, and the expressiveness you can get from just connecting those correctly. But then the latter half was more open ended. And thinking through what is important to represent to capture a particular problems is the essence of that.

HAZHIR Yeah. No, absolutely. And I think that conceptualization part is what we didn't cover much. And I feel actually,
RAHMANDAD: this is the right way going through the experience. That is, that conceptualization is the harder problem.

You need to have enough experience with the components that give you affordances for capturing different things to later come and say, OK, now I have a new problem. How do I cut it so that I can use those affordances, those capabilities to capture it?

AUDIENCE: The risk you run, though, is you overindex on your avoidance capability and it shortcuts any clear thinking about what's actually going on here.

HAZHIR Yes, that's true. I think I hope Nelson's session gave you a little bit of a focus on the other way of, OK, you really
RAHMANDAD: need to ask the hard question of, why am I modeling in the first place? And if I'm modeling it in the first place, what is really critical and what else is out? And we didn't really discuss that, but that's an important part of that conceptualization question.

AUDIENCE: The most helpful question to me that he really asked us was, what is the hypothesis you have that this is a dynamic problem, as well? Specifically, what feedbacks kind of happened.

HAZHIR Yes. And that question is very, very important in getting concrete contributions that academics may appreciate.
RAHMANDAD: It might be secondary importance for a lot of simple operational problems. Like if I'm thinking about these sales of coffee makers thing that I was talking about, there is really nothing that complex dynamically.

I just need to make sense of some data that I have and see what are the trends. And to do that, I need actually a small chunk of a dynamic model. There is nothing really interesting dynamically there. But if I don't do the accounting correctly, I wouldn't make sense of my data. That's a different type of a use case.

So for some of you, those use cases may be even more important. In real-world problems, 80%, 90% of it is these kind of a basic, low-grade, small dynamic issues that you actually need to handle one way or another. Now, sometimes just stock and flow thinking is all you need.

And then you will write your code in your Excel spreadsheet or in Python to transform the data in a way that captures those small dynamic factors. And that's very legitimate. You wouldn't be publishing papers with those. But it is still something useful. And therefore, I didn't discount those ways of doing it in this class, partly because I think it's a useful skill in the real-world, practical applications.

AUDIENCE: I find that to be my issue. And I think that that's what I was struggling with the final project is determining or, in general, actually, is all these problems that I think of, it's not lending to a really interesting dynamic model. And then when it does, maybe there's a skills mismatch or I don't know if the right data or something. I don't know for this course, but brainstorming or how to have a taste for what kinds of problems are really the good ones.

HAZHIR
RAHMANDAD: It's a very important, very hard question. I think I fully agree with you on its importance. And some of it really comes from, so I see two ways of getting closer to it. One is understanding the domain you are contributing to. You really need to know what is known, what is unknown, what is assumed, so what are the mechanisms that people account when they think about that problem.

You have to have those mechanisms as your starting point. That's starting from where the current literature is. And then add an additional mechanism that you find relevant, important, interesting from usually qualitative data, from some insight that has not been yet incorporated into the literature. And then having enough-- so this is the second path-- enough intuition from building models that you would say, I think that mechanism is going to have a meaningful impact.

It will change the dynamics in a way. That is my hunch. And those hunches are uncalibrated at the beginning, so you have no clue. And as you build models, those hunches become more reliable. So a very experienced person would come to a new field, look at the literature, would say, OK, these are the structures, I can see them as the stocks and flows, how they work.

But then there is this mechanism that I think is important, nobody has seen, that makes a good paper. That's how you bring those two pieces together, typically. Some of it is the homework on the domain you are contributing to. Some of it is the homework in your system dynamics skill development so that you can detect those interesting mechanisms.

That's as much as I can put a structure around that. I don't think there is much more you can say in general about it. Yeah. OK, I think it's a good time to stop. We have a speaker coming for a seminar, so feel free to stick around for our talk. And thanks a lot. It was a great semester with you. Thank you.

[APPLAUSE]