ROBERT TOWNSEND: OK. Today we're going to continue the topic of Consumption Smoothing. This sort of sequence began with the optimal allocation of risk bearing and what you would see in the data. We saw the benchmark was doing well in many but not all cases. And there were some examples where coefficients on income were higher and for certain occupations and certain types of income.

Then we took that same benchmark still clinging to it and looked at a whole different kind of data, production data. And we're going to come back to it next time when we talk about labor supply. But while we're on the subject of consumption and consumption smoothing, I thought it would be good to go through some alternative models that are widely used and different data sets and give you a sense of what people do in the literature.

So a lot of this has to do with the permanent income model. It's just kind of the basic sort of thing that people rely on and fall back on constantly or say consumption smoothing over the life cycle. These markets are not complete. They shut down a lot of the risk contingencies allowing households to trade only in risk free bonds or maybe have some other savings account which cannot go below zero or limited credit.

The papers were going to cover, first one, there's one by Kaplan and Violante, which is going to look at the degree of consumption through smoothing from this SIM, which is Standard Incomplete Market model. They're also going to simulate it. So that's using it two purposes there. And we're going to look at smoothing through the lens of that model and in the data against permanent shocks and transitory shocks. This is a language we have not been using so far. But it is standard language in this other literature, and look at how smoothing varies against these various shocks [? varies ?] over the life cycle.

The Violante paper and others draw on this [? Blundell, ?] Pistaferri, and Preston method, which is looking at consumption and income data. And in fact, there is not typically a very long time series in the US to use. The best is a PSID , which used to only have food. So part of this paper is about how to splice together two different databases to create a longer panel and measure the degree of smoothing in the data.

Then we'll move to Deaton and Paxson, which is the same overall topic but actually quite interesting because they're going to be looking more at the variance of consumption in cross sections, which is an implication of the models. But you don't need these panel. You just need a bunch of cross sections for various years. So effectively, if it's representative, you can track people by age. And again, they do that through a permanent income model, a buffer stock model, a model with borrowing constraints, as well as in the background, the full insurance model, and relate it to things in the data.

Then there's this Campbell and Deaton paper, which is about excess smoothness. And I kind of smile every time I see this title because from the standpoint of the permanent income model, you should respond to shocks to permanent income. And they see in the data that consumption is smoother than that. So they call it excess smoothness.

My smiles come from the fact that they haven't considered the full risk-sharing benchmark, which may not fit perfectly well either but may be part of the explanation in the background. And then finally, we have this Krueger and Perri paper, which looks at the US and Italian data. And I hope we get time to say a few words about this today because

AUDIENCE: I'm doing it in the recitation.

ROBERT TOWNSEND: You're doing it in the recitation? OK. Yeah, I originally had this as part of the primary lecture, and then we thought there wasn't enough time in class, which may yet turn out to be true. And then I looked at it again, and I really like it. So I'm glad one way or the other, you're going to see it.

AUDIENCE: [INAUDIBLE] I'll do something else.

ROBERT TOWNSEND: That's fine. Maybe we'll get a lot of questions.

So what's the problem? In the US as I was saying, we need longitudinal data on consumption and income. And on consumption, we need a comprehensive measure. No one's been gathering these kinds of data. So how do people cope with it?

They could use the PS and just look at food. Or they could look exclusively at the Consumer Expenditure Survey, which has a whole bunch of line items. But people are not in that sample for more than three or four quarters. Or you could sort of basically create these synthetic cohorts by sort of extrapolating and merging, which is increasingly a common way to deal with the problem.

Another way to sort of think about this upfront is these models are making a distinction between permanent and transitory shocks, but we only see income. We don't see each of these different kinds of shocks individually. So what to do about that?

You could ignore it and just look at the response to total income change, which is kind of what we've been doing, actually, in the development data sets. You could use proxies for permanent versus transitory shocks. For example, disability might rightly be thought of as a rather permanent shock. Short-term unemployment might be thought of as a transitory shock. So that's another way to try to deal with a problem.

And then there's sort of a public finance tax literature. Different people assume different things about different kinds of taxes. Some tax is assumed to be permanent, others assumed to be transitory. And then you kind of see how households react to that.

So the benchmark model people use for saying, quote, "How much insurance is there actually in the data?" Is this Blundell, Pistaferri, Preston paper, I doubt if we'll get to that today. It is in the appendix, so I have it in the slides, time permitting. But it's basically covariance decomposition of the income and consumption process.

So that's the quote view of the data. Then for the models, you start with the standard incomplete markets model, assuming no access to contingent claims but allowing some kind of insurance through buying and selling bonds. Now let me pause for a second, just in case it's not clear. And this literature uses confusing language, actually.

If you thought about a household in isolation or a business, they have ups and downs of income. And it's clear that by saving, you kind of smooth off the peaks so that higher income does not make its way into consumption. And likewise, in the valleys, when income is really low, you can carry forward that savings to that spot or borrow, which makes the valley less deep. So consumption is smooth to a large degree, but not completely against these fluctuations.

What's the difference between that and the full risk-sharing model? The idea there is that you don't have-- not everyone's facing idiosyncratic and transitory shocks at the same time. So as in the Rocky Mountains diagram, peaks and valleys come at different stages. And even after you act as if a household would do everything that it could on its own, you're still left with this asynchronous timing.

So what the risk-sharing model is basically doing is transferring consumption around in the cross-section to get even flatter consumption. And the only thing that would be left then, if you believe the full risk-sharing model, are those shocks that basically somehow get left in aggregate consumption after all that smoothing. The only part of the peaks and valleys that's common across all the households would be left. And everything else would be flattened out.

So this literature, it's true that you can do an enormous amount of smoothing by borrowing and lending. But it is not true that you can do everything that would be possible with greater amounts of insurance. Anyway, so they start with the standard-- note the word *standard*-- it's standard for the certain branches of the macro literature, incomplete markets literature.

You may or may not put on a lifecycle. To be realistic, you'd say, the household faith sees a stochastic probability of dying. And that makes them more and more vulnerable as they get older and older, in some sense, because they can't smooth that out. Or they have to front-load stuff into savings to be able to cope with fluctuations as they get older.

So it does matter whether the horizon is infinite or finite. There are permanent and transitory shocks to earnings while they earn. But if it's the life cycle model as in the US, where there is literally retirement, then earnings go to zero.

You may have some income stream from social security, and these guys will load that in. And you may have some assets you've put in a pension fund, which has kind of an insurance component to it, in the sense that you're getting an income stream from the pension but averaging over the mortality risk in the population. So there's a bit of insurance in the standard pension funds.

Another version has borrowing limits. There's something called the natural borrowing limit or the zero. Zero is obvious. You can't borrow anything at all, or you'll put just a bound, can borrow 10% of your assets or your income.

The natural borrowing limit is something like this. You always have to repay your debt. So you can't borrow more than what you could repay in the worst possible income realization in the future.

And then it matters a lot what you're assuming about the income process. If there's a small chance of a zero income, and your debt is due in that situation, then effectively you can't borrow at all. So it's a rather stringent criterion.

And then they'll take-- you'll see Kaplan and Violante then take the model. And they simulate, hence the SIM, artificial data from the model itself.

There's two reasons for doing that. First of all, this Blundell, Pistaferri, Preston paper is doing something with the data. It's not a direct look at the data. But the summary statistics come from this BPP algorithm. So they want us-- if you want a common ground, which is what's in the data through BPP, then let's look at the data generated by the model through BPP and compare apples to apples.

On the other hand, when you have a model and you simulated it, that is the reality. You know exactly what the economy is like. So they also use this-- pick up distortions, where BPP is kind of giving you a summary statistic of the degree of insurance, which is not actually what's going on in the underlying model. So they use this to look at the biases or what can generate the bias.

The intuition there is pretty straightforward. Zero borrowing or limits to borrowing are like a corner. And you're either constrained or you're not constrained, so that's a non-linearity, whereas BPP algorithm is basically linear. So the distortions get introduced by the more likely you are to be hitting corners, the higher the distortion is.

So the standard incomplete market model generates an insurance coefficient for transitory shocks of 94% in the natural borrowing constraint economy, or 82% in the zero borrowing constraint economy. That's a coefficient for insurance, so that's a good thing.

Be careful you don't get flipped around, because we've been looking at the degree to which consumption fluctuates with income, so you start to think high numbers are bad. But this is 1 minus that. So for them, high numbers means insurance, which means a good thing. Those are numbers, 94, 82, compared to the BPP run on the actual data, which is 95. So that's an excellent fit.

But when they look at the insurance coefficient for permanent shocks, it's 22% or 7%, depending in the model what you assume about credit. But in the actual data, it's 36%. And here already you're seeing this recurrent theme that there's better smoothing against even the permanent shocks in the data than these models allow.

And then if you looked at the life cycle through the lens of the model, you would see better insurance as people get older. But in the data, it's not there. There's no age profile in the data.

So the model-- another way to say this, is generating too much consumption smoothing for the older workers, relative to the data, or too little smoothing for workers in the early stage of their lifecycle, relative to the data. If you're skeptical, you've got to be skeptical about BPP because it's supposed to be, quote, a "fact" that they're finding.

We'll look at some of that data though through the lens of the cross-sectional distributions when we get to Deaton and Paxson. And then you'll kind of see something else. And I haven't thought about how to reconcile that yet.

So what I said about the reliability, when they simulate and use the data from the model, it's fine for the transitory shocks, largely. But it tends to underestimate the true coefficient for the permanent shocks. That's actually making this situation worse because then the BPP measure, which shows good insurance for the permanent shocks, is a lower bound estimate of the actual insurance in the data, so there's an even greater divergence. Yes?

AUDIENCE: Can I just make sure I understood? So in the data, we see changes in how I smoothed over the lifecycle. But in the model, we don't see any changes? Or have I got it wrong way around?

ROBERT TOWNSEND: No, it's almost the other way around.

AUDIENCE: Right. So the model we see it as [? profile, ?] but in real life we don't see--

ROBERT TOWNSEND: Yeah.

AUDIENCE: So that seems kind of counterintuitive, I guess, to us. Wouldn't we expect-- maybe I'm thinking of a different method.

ROBERT TOWNSEND: Yeah, the problem is taking-- the model is generating differences by age because of the life cycle. But the data does not show that. So that's why you get this sort of odd-looking statement that looks like it's more criticism of the model. But it's of the model relative to the flat data.

AUDIENCE: Yeah, once we [INAUDIBLE].

ROBERT TOWNSEND: Someone once said, if the data don't fit the model, it must be something wrong with the data. So we're feeling at pressure here. I didn't say that.

How can you get more smoothing to generate less sensitivity of consumption to these permanent shocks? Because again, there's-- it seemed to get turned around so easily. There's more smoothing in the data than in the models. So you want the model to allow more smoothing to catch up to the data.

And one thing you could do is give households a little more information about the future. And then the idea-- it's a smoothing model. So if they know things are going to get better or worse, they can take actions now and adjust consumption. So the overall profile that would result would be smoother if they could anticipate future income, rather than having to react to sudden information.

And a close cousin of that is to move away from some random walk, to put in some serially correlated income process. And that, actually, does a better job in getting the model to match the data. As I said, so far in this class, we haven't talked too, too much about these income processes. But it's really front and center here.

What you assume about an income process, is it IID? Is it autoregressive, for that matter, with high frequency data? What about the seasonality and everything?

That makes you aware of both the great strength and vulnerability of the risk-sharing model, which says it doesn't matter what time it is or what state it is. Just add up total consumption and smooth it out. There's a fixed static rule for allocating risk.

There could be predictable periodic seasonal fluctuations or trends in income. It doesn't matter for the risk-sharing model. But here with incomplete models, it matters much more.

AUDIENCE: You say that two is better than one. If it's AR, it's going to be better than predicting the future. But AR, to a degree, is also the ability to predict the future.

ROBERT TOWNSEND: Yeah, it's very related.

AUDIENCE: So there's something else from the AR that's making it better?

ROBERT TOWNSEND: I guess there are other ways to predict the future. Even if the process itself is not autoregressive, you could get signals about what it's going to be in the future.

AUDIENCE: That's not as good as--

ROBERT TOWNSEND: That's what the claim is, yeah. That's what they find. So here's the model, finally.

There is no aggregate uncertainty. Agents work until retirement, but they don't die then. There's a probability [? z ?] of surviving. This is, I guess, retirement is a really bad thing because if you don't retire, you're going to live with probability 1. So anyway, this is a simplification. So they only start using these actuarial tables after age 60 or 65, whatever they put in for retirement age.

So it's maximizing discounted expected utility, common utility function. We've been doing a lot of that too. Beta is the standard discount rate. [? Z ?] is this-- you only get the utility if you're alive, basically. So that's a standard adjustment to the discount rate.

The income process is a bit complicated. First of all, there is a possible nonstochastic trend, sort of the deterministic part, which could move income along with T. And then you have income, quote, "standard." Here's the decomposition. Income is equal to z, which is this permanent thing, in this case, entirely a random walk, and epsilon, which is the IID transitory shock.

So these papers use different notation. And this paper, z, is permanent, and epsilon is transitory. And eta is another shock. And that's the shock to permanent income, basically.

So in expectation, you expect your income to be what it was. You expect the permanent part of your income to be exactly what it was last period. But this thing will move that around.

And then these have variances-- sigma eta, sigma epsilon. The budget constraint, again, is like this bond, which you may or may not be able to short. So you've got your assets from last period. That's earning interest, plus your income. So this is like your disposable income, which you can spend on consumption and assets, going forward.

Watch the subscript. It'll drive you crazy. It does me. This is dated t plus 1. But it's decided on a t, and consumption is also decided on a t. So in some of these formulas, you're going to see a mysterious interest rate adjustment. And it has to do with this sort of convention of the timing.

This looks the same, except as I said, it's like a pension fund in your assets. So effectively, this gets adjusted. It's not like everyone has their own individual retirement account, which you draw down. There's a pool of money there. Some people continue to live, and some don't.

And so those-- in other words, if you're, quote, "lucky," and you keep living, you could actually draw more from the fund than what you put in there. And this is a very convenient way to kind of adjust.

Oh, and if you're retired you get your social security benefits. They're going to put some numbers in here from the life cycle literature and other sources, using various US data sets. And I'm not going to dwell on the numbers.

But this is a classic picture of the life cycle. So here are the earnings going up. You guys are down here somewhere, on average. You can expect to do better eventually.

And then you get to be 60, and it doesn't go to zero. That's if you were a US citizen, that's where you start drawing. Actually, you could just be an immigrant. You start drawing social security benefits.

With this profile, you try to keep consumption steady. It's not completely flat. It just looks flat relative to everything else on this diagram. It's actually increasing somewhat. But the goal of the life cycle smoothing is to try to have more or less steady consumption.

Again, this is-- there are two lines here. There's a natural borrowing constraint and zero borrowing constraint. It doesn't make too much different for consumption.

I guess here the zero borrowing constraint looks more constraining. So when you're young, expecting on average higher future income, you might want to borrow against that and repay the debt later. But you hit one or the other of these constraints, so you can't do that too much.

Of course, the gorilla in the room is this wealth profile. And that's where all the action is, basically. In fact, you're basically over the life cycle, accumulating and then deaccumulating wealth. That's what's keeping consumption steady.

And remember, when income goes to virtually zero, apart from social security, you've got to live off the interest ir your-- you have social security, plus the adjusted actuarial interest off the pension fund. And that's typically a lot smaller flow than was your previous income.

So you need a ton of assets to be able to survive into old age. This is drawn out, looks like to 90 something, 95.

That could have been shown first, but that's kind of a summary of the life cycle model and what Kaplan and Violante are doing, and they're generating these smoothing statistics from date by date with shocks. But this is the overall pattern you would see.

Let's go to Deaton and Paxson for a very different look at it through the lens of these cross sections. The starting point, again, is that with a permanent income hypothesis, consumption of each person should basically follow a random walk. Again, the idea is you should basically be eating at the level of your permanent income. And when you're permanent income is shock, then your consumption is shock.

But the implication that consumption follows a random walk means, basically, you'll see the equations for it. Date by date by date, you get hit with positive or negative shocks and their IID. So basically, some people have their ups and other people have their downs.

So you start getting this fanning out. And then you go-- then you take the cohort that is still all alike. Go one more period. They have ups and downs. Those guys fan out a bit. So the cross-sectional distribution just keeps getting fatter and fatter with bigger and bigger tails.

So all of the slides, basically, are going to be about this increasing consumption inequality in the cross-section. They measure it by the standard deviation of the log of consumption, age group by age group.

But again, if you thought about full risk-sharing, you get this dramatic contrast. At least for the cases where people have uniform risk aversion and so on, then we know that what pins down the level of your consumption, on average, is going to be your Pareto weight. And you kind of have those fixed at t equals 0. That's going to fix the intercepts.

And your consumption might fluctuate with aggregate shocks around that, but it's not going to move with age or anything else. So basically, the cross-sectional dispersion of consumption does not increase in the full risk-sharing model. But you will see that in the data, depending on the country, it does happen.

You have to be a little mindful of going back and forth between this cross-sectional age-dependent inequality, and inequality for the economy as a whole, because demographics can shift. Even if you give me someone's age, and I'll tell you sort of what the cross-sectional dispersion is in the population, when you have demographics going on, you may have an increasing number of young people or old people, depending on the economy.

So you're basically taking an age-weighted average as the demographics shift, and that will shift inequality. So an issue in this literature is how much of aggregate inequality can be explained by these kinds of life cycle smoothing considerations, and how much by baby booms versus fertility drops and all of that.

So this is kind of nice because they have the US, Britain, and cross sections, which are three countries that have a quite ample amount of cross-sectional consumption, for that matter, also, income data. They're going to use 47 annual surveys spread out over the three countries and look at the implications of these models for the kind of cross-sectional inequality and see what it takes to explain it.

So how does this work exactly? So we want to follow cohorts over time. We cannot follow an individual person over time, because it's not panel data. But you make the assumption that, for example, if you see a bunch of 31-year-olds in 1976 in Taiwan, you're getting a good summary statistic of the cross-sectional dispersion of their consumption in 1976.

And then you go to 1977, and they are now 32. Actually, "they" isn't quite correct, because it's a different set of people. But again, you hope you have enough of them and that the basic assumptions are correct, that you can take these kinds of statistics as representative of what you would have seen if you could have been able to continue to track the same people.

Some people actually think this is better than panel data, in the sense that you don't have to worry about people dropping out, and the data becomes unrepresentative. And there is some truth to that.

There are issues about whether the data are coming from households or individuals. You have to make some adjustments for household composition. And the surveys being used here are the Personal Income Distribution Survey in Taiwan, the Consumer Expenditure Survey in the US-- that's the same one we've referred to before-- and this Family Expenditure Survey in Great Britain.

So it takes a bit of staring to get used to this. So we start out with people who are age 15 in 1976. And they kind of start here in calendar time, 1982 or something like that. Again, we don't track the same people. But we can look what happens.

This paper is not the most recent thing in the world. It's just a really nice one. So they basically stopped using data after 1990. So we go back to the 1976 survey, all the way up to the 1990 survey, with repeated cross sections and looking what happens to inequality in this cohort, as they age.

These guys were 15. These guys were 20 in 1976. So they start older, and they're going to end a bit older. And then you keep going here to the aged 55 people in 1976 and so on.

These are pretty flat, then these profiles start to pick up. You kind of in your mind sort of can imagine. And I'll show you what happens when we splice all this together.

AUDIENCE: It seems like a lot of the action is brought about by the-- towards those middle-aged cohorts. Those who start middle-aged did in '76, like, from 40 to 55.

ROBERT TOWNSEND: Let me just--

AUDIENCE: The rest is kind of flat.

ROBERT TOWNSEND: Let me just jump because I-- those are the three countries. Here's what's going to-- when you use all the surveys over all the years and just focus on the age-- yeah, so you see this. You're right. You see this is pretty flat, and then it picks up around age whatever, 45 or 50 or something. And these are the results from the other countries.

AUDIENCE: So is there any explanation for why that's the case?

ROBERT TOWNSEND: Yeah, we're going to get into that. You'll be sorry you asked.

AUDIENCE: And only in Taiwan.

ROBERT TOWNSEND: Huh?

AUDIENCE: And only in Taiwan. The others seem to be--

ROBERT TOWNSEND: OK, so this is kind of flat, and then it kind of comes up, and then it flattens out again. So Taiwan is flattening out, although, unfortunately, the cohorts are getting thin the older they are. So you get some sort of shaky data because you can't average quite so well.

This one flattens out. Great Britain kind of flattens out. The US, I don't know, wasn't very steeply climbing to begin with. It's not so obvious. This kind of looks concave. This looks concave but almost linear. This actually has almost an inflection point and then comes in. So the question is concave, convex, et cetera, and that's where we'll do a little modeling.

By the way, you saw a little bit about demographics when we did China. One of those explanations of the savings rates, et cetera, wasn't so much about state-owned enterprise as it was the different demographics, and young savers in China, and sort of debt-loaded or less saving-inclined households in the US, and so different demographics.

Just to remind you, it's not the first time that we've thought about demographics. So it's possible to put that-- there's trivial demographics in some of the macro models, in the sense that they start with a lifecycle model or two-period overlapping generations model. But we haven't actually focused on much of this in the sort of micro part of the class.

So this is what I said in words. They claim Britain is slightly convex. I didn't see that just now. And again, don't misread those things to be measures of the total increases or decreases in inequality, because the demographic structure is moving around by age over those 20 or so years.

Let's start with the basics. Here is basically Hall's permanent income implication for consumption. Consumption should essentially be a random walk. And it's only moving around because of the innovations or the shock to permanent income. I think that was [? JPE, ?] 1970. The dates here, I think it was in the [? JPE. ?]

There are assumptions. This isn't always true. He assumed quadratic utility, additive preferences, discount rates equal to interest rates, and, I think, another life cycle.

This is supposed to be an innovation, so for every person at least, there's no covariance between lagged consumption and this innovation, because it wasn't supposed to be predictable. That's the same conversation, just how much, if anything, you see about the future. Here, zero is seen.

So I described this in words. But basically, if it were stable like this on average, then consumption dispersion shouldn't be changing. But you're just layering on repeatedly, shock after shock after shock. So with that zero covariance assumption, the variances are increasing by the variance of that [? u ?] t shock.

And basically, if you started plotting these distributions, the one at t is going to have fatter tails than the one at t minus 1. So t minus 1, [? quotient, ?] first order stochastically dominates.

You have to be a little bit careful throughout all of this stuff. You think about things in the point of view of a person tracking a person over time. But that statement about c t and [? u t ?] is actually a statement about the cross section, as well as the intertemporal dimension.

This model typically didn't have taste shifters or heterogeneous preferences. You can put those things in. Depending on how you put them in, it does make a difference. Arguably, if we hadn't done the demographics right, but they claim-- Deaton and Paxson claim for Taiwan, then you kind of need preference shocks to help with the observed dispersion.

Here is the basic sort of budget equation again. And I'm afraid each one of these papers kind of has a different convention about the timing. So it's really kind of hard to keep tracking and track it.

But basically, you have sources and uses of money. The sources are previous assets plus interest. The uses are-- source also is income, and the use is consumption. And then what you don't spend accumulates on to the next period.

If you do do the life cycle with a terminal date, for sure, say, age 99 or whatever you want, then you can actually get a formula closed form for consumption, under these other maintained assumptions. And basically, it's kind of saying, consumption ought to equal something about the return stream through the interest rate on assets, and something about the forward-looking part, which is expected future income shocks.

So this kind of helps you see how future things are already-- if this is going to be very high, for example, then this is already going up. So it's kind of smooth against-- or low, vise versa. It's tricky with a finite horizon and a cap t because there is a coefficient premultiplying this, but it's for that reason, basically.

Obviously, the closer you are to the end to cap t, the less smoothing you can accomplish. And that's inherited into this formula. Yep?

AUDIENCE: The shocks here are IID all the time?

ROBERT TOWNSEND: Yes. Well, the shock to permanent income is IID. But permanent income is an autoregressive process.

The standard incomplete markets model had an infinite horizon. And it had a discount rate equal to 1, or basically beta times 1 plus r equals 1. In the finite horizon case, beta, that thing premultiplying consumption, turns out if you stare at it, is concave. We're already beginning to anticipate these formulas and decreasing in t.

And if you go back to those two equations, you'll get kind of a simplified version, that beta t times the change in consumption is equal to this, quote, "consumption innovation." And here's a formula for the consumption innovation, which is, again, the difference in expectation.

So you're forecasting the future all the time. But you're also getting these shocks. So as of t and t minus 1, you were forecasting y t at t plus k. Sorry, y of t plus k, not either to either, to t or t minus 1. It could be 10 years out.

And you're constantly adjusting your expectations, based on the current information you have from the income process. And you're summing up over all those future income shocks and discounting backwards to get these, quote, "innovations."

So then imagine starting this out at some date t equals 0, and then t equals 1, and t equals 2, and so on. And you could just write out, literally, consumption in each one of those dates. And of course, when you go from one date to another, you get hit with this shock.

And you're premultiplying by the beta, so you end up with this, an explicit formula for consumption equal to its initial given level. Plus, the beta moves on the right-hand side, this sort of beta discounted. But it's not beta in the preference function. This is that beta thing that had to do with little t, big T, and all of that, that took care of the lifecycle considerations.

So finally, we see again the familiar theme that the variance at t is related to the initial cross-sectional variance at zero, plus the sum of variance terms. So the issue-- so roughly speaking, it looks like, yeah, it ought to increase-- unless you retire and you have no more things then forecast, then this piece goes away.

But before that happens, you're looking at future and making expectations of future income. So you just keep adding on. So that's why inequality is increasing in the data, through the lens of these models. And then the issue is, is it linear or concave or convex? That has to do with this beta and also has to do with what we can infer about sigma, about the consumption innovations.

So in response to your question, if we assume this innovation is entirely white noise, normalized by the interest rate, then sigma square eta at t is a constant that doesn't depend on t. And this is the resulting expression with the sigma square eta pulled out front. And then the convexity is entirely due to this beta thing, which if you start taking some derivatives and so on, it's a bit messy. You can convince yourself that that thing is convex.

Here's an orthogonal example. I guess, [? Jan, ?] I jumped the gun a bit because this actually allows something a bit different, which is this autoregressive thing. This is IID, and then this is autoregressive, so this income process is persistent.

It's a bit odd because this was the innovation that had to do with the innovation to-- the right-hand side had to do with all those expectation differences. So it's like, let's just assume it looks something like this. This is more explicit about the income process. And you can derive the formula for the change in consumption, where this beta r thing is now this mess.

And it turns out, this will be concave if you can sign this expression. And it may also be decreasing in age if this is true, and so actually, except for some days of theta, both things--

So two of those three countries, the profile was, say, not concave. It could have been linear, which is a weak case of concavity, or even convex. And so basically, then you could rule this out because these are-- this kind of persistence would give you concavity. But you don't see that.

But then we have the sort of paradox that something like these white noise processes must be closer to the truth if we're going to get convex profiles. But that's an assumption about the income process. And in particular, one way to get it is to say that there's a large stationary component to income. So you're not getting those innovations in your forecasts. And we don't see that in the data.

Income profiles are not stationary. There's heavy lifecycle components. So nothing's perfect, I guess. Every model is a benchmark. Almost every model is going to deliver something that might fit well. Otherwise, we're probably not reading it in a journal. Or it may actually also generate things that are not consistent with the data, and if you're a good scholar, you're reporting that too.

What about the permanent income and the dispersion of income by age? So again, disposable income takes on this form, which is sort of asset return stream, plus income. And then you use it for consumption and savings, adjusting for where you are in the life cycle.

Or you could take the difference. Savings is the difference between income and this beta-adjusted consumption. And some of these models you can get a closed form expression for savings, namely the so-called Campbell's "rainy day," which is really kind of cool, which is you save enough at t to be able to cover the discounted shortfalls in sort of these earnings innovations. So you want to smooth those out. And you save just enough to do that in expectation.

Anyway, so we're rewriting disposable income equal to the savings component. We already had that, less the savings-- plus the savings component. Sorry, I get confused because savings was negative. It used to be here, and now it's out here.

Savings is not negative, but these innovations are negative because you save to cover shortfalls. And then you get implications here for if one thing is a random walk, and the other thing is a random walk, then the sum of two things that are random walks is also a random walk.

So basically, savings ought to be stationary in that sense. And it ought to be dispersing at the rate that consumption is dispersing. And they don't see that in the data either. And in particular, earnings is dispersing, which is a point that was made on the other slide as well.

And then finally, I guess this is like the fourth version of something that starts to look like the same thing. Most of those analytic expressions were assuming something like quadratic utility. That's how you managed to get this really tight reduced form analytic expression. If you don't do that, if you have constant relative risk aversion, for example, or a more general utility function, than the Euler equation basically is going to equate the marginal utility today to something like marginal utility tomorrow.

We might say, well, where's the expectation operator? Instead, the shock is put over here on the right-hand side. So this is like an Euler equation. If this were 0, and the interest rate was equal to this delta discount rate, then basically the marginal utility of consumption, if we could deduce it, would be not going anywhere. It would be the same over time.

But again, with r equal to delta, we kick on this extra orthogonal shock, which makes the marginal utility of consumption dispersing in the future, relative today. And you can even have r not equal to delta, as long as delta is greater than r. The same logic applies because you've kind of like amplified this thing, carrying it through to tomorrow, along with the u shock.

So that's another way to get at the variance of consumption increasing, depending on what you assume about the utility function. Some functions, though, have a force that go the other way. In particular, if you're really, really into it, you get into third derivatives. But when you're really sort of cautious, that lambda is convex.

So you're taking a second derivative of already the marginal utility. That's why I commented about third derivatives of utility functions. And that, if you start staring at this thing, can actually push you back against this dispersion of consumption.

So I'm having a sense that this is going on and on. But the point is to take each of these different forces and then tell stories about what must be true in Taiwan, versus Britain, versus the US, to try to rationalize what we're seeing in the data, and similarly, here.

So let's just say three words about excess smoothness. To repeat, the lifecycle model implies that shocks to permanent income should be fully incorporated into consumption, while innovations to the transitory part are not. Basically, all the income-- all the transitory income fluctuations should not be appearing in consumption, which actually is something we've talked about with regards to the full risk-sharing model, except there, we didn't make this distinction between transitory and permanent.

But when they go and look at the data, they say, hey, consumption is too smooth. It doesn't react to innovations in the permanent component. And other people have found something similar. And we're inching forward to models where we're going to need to modify the models. We're going to have to introduce something else, like private information, and more on that in a second. Yup?

AUDIENCE: Why do all these assume the market is incomplete? So maybe just because the market is relatively complete, so these [INAUDIBLE]?

ROBERT TOWNSEND: So that's my gut reaction. Actually, so this Pavoni paper puts in sort of some unobserved savings and other things. And it actually comes close to the data. So it isn't all the way toward full insurance, but it is more full insurance than what these models imply. So it's trying to reconcile the puzzle.

I'd like to think I'm quite agnostic. In my rural [? Thai ?] data, certainly, when you include other variables and not just consumption, we see rejections of full risk-sharing. I'm not determined to always find full risk-sharing in the data. But likewise, I would hope these guys aren't clinging to the permanent income model as the only game in town, because it's clearly also suffering from its own sort of anomalies, relative to the data.

And there's no reason to think that it has to be this same model for every village or every country, and actually, Deaton and Paxson's stuff, looking at the cross sections, kind of saying that it's different. It could be different in different countries if you're going to try to reconcile the observed movements.

But anyway, you guys need to know about this literature. You need to have a view of the different consumption smoothing literature. I guess I should also say sort of, it looks as though it's a macro-ish, US-ish literature. But that's not true, certainly, for Paxson and Deaton, who are avid development economists and looking at that data from developing countries.

And likewise, the initial risk-sharing stuff, that wasn't peculiar to development. That was being tested by Cochrane and Mace and so on in US data. So there's really never been this idea that somehow macro is in one room, using one subset of data, and micro is another room, using another subset. That would be rather silly. So we really need to know how all these models work.

So this is, again, Campbell's saving for a rainy day. It's repeated because it's a different paper. They look at-- Campbell and Deaton now are looking at US data and saying the labor income is described by this autoregressive process with a positive serial correlation. And what that means is innovations are, quote unquote, "more than permanent," not just random walk-ish, but basically--

Now you get even into more trouble because if consumption should reflect the innovations to permanent income, and what you see in innovations is more than permanent, that means consumption should respond even more. And it's not, in the data, responding that much. It's much less variable.

So either they're getting this fact wrong. Or, again, maybe the households have more information somehow. Or they have stupid expectations or whatever.

These guys were not able to resolve the puzzle. If you want to see the equations, there they are. I'm not sure at this point that it adds all that much more to the discussion. And this is what I was saying about Attanasio and Pavoni.

So the plan is-- there's a little more coming today. Don't worry. Or you might say, oh, darn. But the plan in the ordering of the lectures is to do labor and wage variation and smoothing, adding labor supply to a consumption model, and talk about elasticities and sensitivities. That will be done with both the full risk-sharing model, as well as a version of these incomplete market models. So we'll see what survives, depending on what you assume about the market structure.

And then we're going to go to have a whole lecture on models that are much more explicit about moral hazard, unobserved income, and so on. And this Pavoni paper is one of them in that literature. We're not going to cover it in class.

So actually, the front part of the lecture was about this BPP. It's probably a bit mysterious what it was, so we can review that for a second. The idea is that you postulate that log of earnings takes on this process where x is a vector of shocks, a is some vector of coefficients, the shocks are IID, the x's are, and they have variances.

This allows a whole bunch of stuff, not only just random walks, but autoregressive integrated moving average processes on income. This is what you see. And then this general formulation allows a bunch of different specifications, which they're going to try to estimate.

And somehow what they want to back out is this insurance coefficient. So here's, again, how it's easy to get turned around. This is how a sort of, like, quote unquote, a regression coefficient of how consumption is moving with this particular shock, inferred somehow. But 1 minus that is then the degree of insurance because when consumption doesn't move with that shock, this is zero. And so insurance is like, perfect, measured at 1.

And you could put t's on these things if you sort the data by age, which came up earlier. So we don't really see all those permanent versus transitory shocks. You kind of have to infer them, somehow.

And this is a bit like using some assumed function on the income process, and then filtering the data through that function and looking at the covariance with income change. But they don't just do three alone, because this would be a statement about income alone. They jointly estimate this function, g, essentially, jointly with consumption.

So for example, if you assume income took on this classic form that we were actually tracking through the lectures, then income would be the sum of a random walk with its own innovation and variance and epsilon, which would be this IID transitory shock. If you believe this to be the structure, and there are many, many other candidates, then you take a first difference. And by construction then, what would be left would be this innovation in the permanent part and a time difference in the transitory part, which is still transitory.

So then the consumption model tells you about-- puts restrictions on how the change in consumption should be responding, basically taking advantage of the time delays. As you go back like, two, three periods, there is nothing that happened that far back that's influencing anything at all today.

So those past data kind of become predetermined variables, almost as if they were instruments. And so that kind of mysterious notation about g on y, this is a version of it, postulating a model and then going back far enough time, so you would see, basically, zero covariance.

So here it is, actually. This g is just basically the time difference in income at t plus 1. And given the other assumptions they make, you back out some of the key things that you want. For example, how on earth are we going to know about the variance of the transitory shocks if we never see transitory shocks?

Basically, that turns out to be the covariance of two things we do see, the time difference of consumption at income at t and the time difference of income at t plus 1. Hopefully, you're seeing the spirit of this, if not following every line of the algebra.

So that's where we get that object. And how would you get the variance of the permanent income part, the innovation to permanent income? Remember, that thing is persisting over time. So it's a bit more complicated.

It turns out to be this daunting object, which is to-- and everything here you see. That's point number one. It's the covariance of the time difference of income at t, against the sum of the changes in income at t minus 1, t and t plus 1. So there, again, you see this sort of time structure at work. The spirit of it is go back far enough in time so everything is predetermined.

And this is the covariance of consumption with that permanent shock. So this is a key object. Remember, that insurance formula is, how much is consumption moving with the innovation to permanent income? It all seems quite mysterious, but here is an explicit formula for how they get it.

So clearly, it's a linear model, using a lot of these variance covariance formulas, given the assumed structure. They don't have to-- their starting point could have been something else. As I said, you had these ARIMA processes, but it is some structure. When you specify the order of the moving average part and the order of the autoregressive part, you get restrictions on the data. So hopefully that helps resolve some of the mysteries about what this BPP algorithm and what they do to the data to measure these insurance against idiosyncratic and permanent shocks.

So I'm going to leave for [? Whit ?] to do the version of smoothing, a bit in the Italian data. But I will just say, by way of motivation, it's again looking at various models, although they're not exactly nested, and looking at the responses of consumption to innovations, but also of wealth to innovations.

So we saw in my data sort of how wealth in various lectures is moving around in the cross section and moving around over time. We talked about responses to shocks and whether they're using savings accounts and so on. So these guys in their own way are doing something similar.

And the paper backs out the movements in consumption and wealth that are predicted from certain kinds of innovations, but not just between this year and next year, but this year and two years from now, all the way up to six years, or even longer. So you kind of get the sort of time profiles of responsiveness.

And very much in the spirit of what we're talking about today, yet again, a bit different, those response patterns are very different, depending on whether you're talking about the permanent income model or the life cycle buffer stock type model. That's all for today.