Structural Transformation and Economic Growth (STEG)

STEG CEPR

STEG Virtual Course on "Key Concepts in Macro Development"

Friday 9 April 2021

Lecture 9: Heterogeneous agents models and methods

Ben Moll (LSE and CEPR)

Presentation Slides

Video

Q: So you take time t to the limit 0, so not partial derivative of value v with respect to time t? I expect to see value v with respect to time t in HJB. *no partial derivative of value v with respect to time t

> A: I believe this comes from stationarity because the time horizon is infinite.
>
> The HJB equation is exactly analogous to an asset price equation. The return on the asset is \rho*V, which is the current profits or dividend, r(x,y) + expected capital gain v'(x)*f(x,\alpa).

Q: What is the boundary condition associated with the HJB?

> A: See equation (10) here. http://benjaminmoll.com/HACT/ for some intuition. Basically, the boundary condition says that if you're AT the lower boundary, your savings will always be positive so you are pushed away from the boundary.

Q: This may be more appropriate for the end, but in thinking about bridging macro to micro, how can one think about parameter identification in these models? Do we still have the same fundamental problems that some parameters are not identified (i.e. the discount factor)?

> A: It depends how you want to think about identification. If you are thinking about it in a structural sense, these models are typically well-identified, even overidentified, if you have data on the wealth distribution, investment rates, etc. for example. You have many moments to target that will discipline the discount rate, risk aversion/intertemporal elasticity, etc. On the other hand, this is structural identification using observational data. Experimental methods of identification can also be used. My work (Kaboski and Townsend, Econometrica, 2011) has an example of this.
>
> A: That's very useful, thanks for that reply! Although in the context of observational data, I am a bit confused about the differences, if there are any, between this type of moment matching you are describing and typical GMM estimation in structural micro.

A: The difference is twofold… 1) GMM typically uses the statistical properties (variance of errors) to weight the various moments a la Hansen. Calibration uses more subjective weights, relative importance of different moments, to choose or weight moments, 2) GMM leads to standard errors… instead calibration usually uses sensitivity analysis to quantitative implications.

Q: I understand from the consumption-wealth plot that high income guys consume more, but the opposite may be the case when MPC is used. Any clarifications on this? I was just kind of just wondering if it is always the case that high income individuals tend to consume more. This is important especially because we are dealing with consumption-wealth relationship in a heterogenous agent environment

A: If I understand correctly, the answer is no... Because consumption depends on wealth as well as income, the people with high incomes and low wealth may end up consuming less than people with low incomes and high wealth. The heterogeneity in wealth is precisely what will make this relationship more complicated. Is that addressing your question?

A: I was referring to the wealth-consumption behaviour. We know from a standard Keynesian consumption model that poor people tend to have a higher marginal propensity to consume (MPC). This can be the implication of continuous time models

A: So the point here is that ‘rich' and 'poor' is now complicated by the need to distinguish between wealth and income. In this model, high-income people and low-income people don't necessarily stay that way. High-income people may become low-income with some probability, and vice-versa.

A: With financial frictions, poor people may be liquidity constrained… they are poor because of a low-income shock, there consumption is therefore low, which gives them a high propensity to consume.

A: Absolutely, thanks. I guess I was sort of using income and wealth analogously.

A: And one of the nice aspects of these models is that we can distinguish between these two things... That also means that we can consider a case where (for instance) the poor differ in terms of their assets, or their education, or some other variable that may matter. The richness of the models is very helpful for thinking about the role of various frictions in developing country contexts. A credit market failure, for instance, may matter for those with little wealth but maybe not for those who have substantial assets. So these models give us a lot more scope to approach real-world data.

Q: Very basic question, but in this model the law of motion (using language from previous slides) is savings conditional on assets and income realization?

A: Right, it’s the evolution of the state variable (here assets). And that depends on savings (that are endogenous to both wealth and income).

Q: Wait more specifically, it is the derivative of assets with respect to time conditional on assets and income. Is that right or am I misunderstanding?

A: Yes, if by “conditional on” you mean that it is “a function of”.

A: Right, so given your assets (i.e. wealth) today, savings decisions and the random evolution of income, what is your stock of assets tomorrow. And the Kolmogorov Forward equation just tracks this law of motion of assets for each individual. So the KF equation tracks the distribution of assets over time.

A: So I think it took me until now to realize why Joe emphasized "as a function of". Because the whole point (or one major point) is to get the derivative to be a function of only income and wealth and not a function of time. Thanks for being patient with me both of you :)

Q: How do we solve a problem with an unbounded state space?

A: One way is to transform the problem into a bounded state space problem. So for example, as an analogy, we transform the BGP of an unbounded neoclassical growth model into a steady state problem in normalized variables.

Q: In discrete time optimization, there is a max operator which usually is the computationally intensive piece of the code. In the finite-differenced method here, there is no such thing. Is there any intuition why we don’t need a max operator and still end up at an optimal solution to the household problem? Thank you!

A: I'm not sure if this answers your question but you implicitly do have a max operator because for consumption c, we substitute its optimal value from the FOC. In particular, see slide 33 of Ben's slides: c = (u’)^-1(v’).