9.0.2 Aggregation, risk sharing, and pareto optimality

Our empirical work and application typically uses aggregate or average consumption growth

$$c_t^a = \frac{1}{N}\sum_{j=1}^{N} c_t^j.$$

Let's explore whether that makes any sense. Are we allowed to use $u(c^a)$? Does idiosyncratic risk matter?

- *Identical investors.*

Well, that was easy. If we're all identical then of course $c^i = c^a$ and we're done.

- *Complete market aggregation and Pareto theorem*

In a complete market, there IS an aggregate utility function – the "social welfare function." The planner maximizes

$$\begin{aligned} V(c^a, c^a(s); \{\lambda\}) &= v(c^c) + \beta^a \sum_s \pi(s) v(c^a(s)) = \\ &= \max_{\{c^j, c^j(s)\}} \sum_j \lambda_j \left[u(c^j) + \beta_j \sum_s \pi(s) u[c^j(s)] \right] \\ s.t. \sum_j c_j(s) &= c^a(s);\ \sum_j c_j = c^a \end{aligned}$$

The first order conditions are

$$\begin{aligned} \lambda_j u'(c_j) &= \Lambda \\ \lambda_j \beta_j u'(c^j(s)) &= \Lambda(s) \end{aligned}$$

$$m(s) = \frac{\Lambda(s)}{\Lambda} = \frac{\beta_j u'(c^j(s))}{u'(c_j)}$$

Here $\Lambda(s)\pi(s)$ is the Lagrange multiplier on the $c^a(s)$ constraint. (You can also define constraints without the π and then you end up with contingent claim prices not m. I thought this was prettier.)

These are the same first order conditions as agents in complete markets, so we've proved the Pareto-optimality of the complete markets equilibrium. Also to review, we have perfect risk sharing

$$m_{t+1}^i = m_{t+1}^j.$$

V is the value function – the maximized planner objective as a function of available aggregates. By the envelope theorem – marginal value of a dollar is the same in any use, and the shadow value of the constraint –

$$
\begin{aligned}
v_c(c^a) &= \lambda_j u'(c_j) = \Lambda \\
\beta^a v_c(c^a(s)) &= \lambda_j \beta_j u'(c_j(s)) = \Lambda(s)
\end{aligned}
$$

$$
\frac{\beta^a v_c(c^a(s))}{v_c(c^a)} = m(s)
$$

Thus, *the derivatives of the value function (social welfare function), defined over aggregates, price assets exactly as the derivatives of a utility function.*

- *Issues*

So, what is aggregation theory about? One issue: functional form. Lots of aggregation theorems try to prove that if the individual utility function is of a certain form, and with a certain amount of heterogeneity in preferences or wealth, the social welfare function = utility function over aggregates has a certain form – usually the same.

In fact, I cheated a bit, as I have not even proved that V is separable. In this case, that's not really a problem, as we can just interpret the symbol $v_c(c^a(s)) \equiv \partial V(c^a, c^a(s); \{\lambda\})/\partial c^a(s) \times 1/(\pi(s)\beta^a)$ and price assets with a nonseparable value function. The bigger issue is whether an economy populated by people with power u and a certain amount of heterogeneity produce a value function also with power uility and some sort of average wealth and risk aversion.

I find this less interesting, because we don't really know much about individual preferences, so this is, empirically, not a big constraint. Rather than say "we assume individuals have power utility" we might as well just say "we assume the SWF is of a power form."

Another issue is the impact of wealth heterogeneity. There isn't any here. Yes, the form of the social welfare function can be different for different $\{\lambda\}$. If one class of people are more risk averse then a SWF will be more risk averse if the λ weights them more. But in complete markets, there is no uninsured individual risk. And the λ do not change over time. So really this is again interesting only if we want to build up a SWF from knowledge of individual utility. (We will have models with idiosyncratic risk in a minute.)

- *An aggregation example*

Here is an example of a simple complete-markets aggregation theorem. Suppose people have power utility, the same risk aversion and discount rate, but different levels of wealth. What does the SWF look like?

$$\begin{aligned}
\lambda_j \beta \pi(s) c^j(s)^{-\gamma} &= \Lambda(s) \\
c^j(s) &= \left[\frac{\Lambda(s)}{\lambda_j \beta \pi(s)}\right]^{-\frac{1}{\gamma}} \\
\frac{1}{N}\sum_{j=1}^{N} c^j(s) &= c^a(s) = \frac{1}{N}\sum_{j=1}^{N}\left[\frac{\Lambda(s)}{\lambda_j \beta \pi(s)}\right]^{-\frac{1}{\gamma}} \\
c^a(s) &= \left(\frac{\Lambda(s)}{\beta \pi(s)}\right)^{-\frac{1}{\gamma}} \frac{1}{N}\sum_{j=1}^{N}\left[\frac{1}{\lambda_j}\right]^{-\frac{1}{\gamma}} \\
\left\{\frac{1}{N}\sum_{j=1}^{N}\left[\frac{1}{\lambda_j}\right]^{-\frac{1}{\gamma}}\right\}^{\gamma} \beta \pi(s) c^a(s)^{-\gamma} &= \Lambda(s)
\end{aligned}$$

Looking at the top and bottom equation, you see we have the same thing – and λ term will cancel in marginal rates of substitution. So, in complete markets we can accommodate differences in wealth. Obviously, we want more, uninsured shocks and differences in risk aversion.

- *An example*

A lot of the problem is solving the Lagrange multipliers. A two-person example lets us look more closely

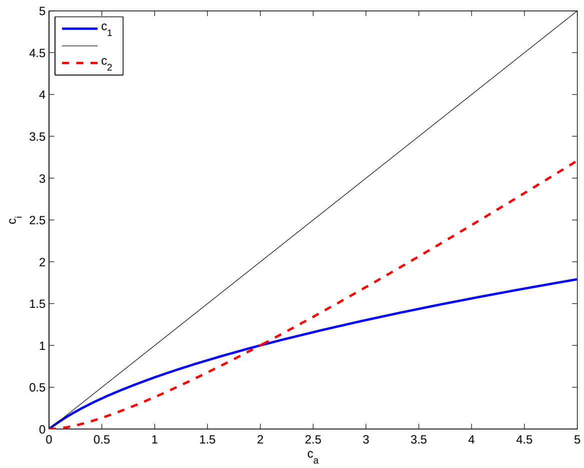
$$\max \lambda_1 c_1^{1-\gamma_1} + \lambda_2 c_2^{1-\gamma_2}$$

$$\begin{aligned}
\lambda_1 c_1^{-\gamma_1} &= \lambda_2 c_2^{-\gamma_2} \\
\lambda_1 c_1^{-\gamma_1} &= \lambda_2 \left(c_a - c_1\right)^{-\gamma_2} \\
\left(\frac{\lambda_1}{\lambda_2}\right)^{-\frac{1}{\gamma_2}} c_1^{\frac{\gamma_1}{\gamma_2}} &= (c_a - c_1) \\
\left(\frac{\lambda_1}{\lambda_2}\right)^{-\frac{1}{\gamma_2}} c_1^{\frac{\gamma_1}{\gamma_2}} + c_1 &= c_a
\end{aligned}$$

Now we have to solve for c_1 This gives us a sharing rule.

Obviously, if $\gamma_1 = \gamma_2$ $\lambda_1 = \lambda_2$ then $c_1 = 1/2 c_a$..For a more interesting example, suppose still $\lambda_1 = \lambda_2$ but $\gamma_1/\gamma_2 = 2$. Then

$$\begin{aligned}
\left(c_1^2 + c_1\right) &= c_a \\
c_1 &= \frac{\sqrt{4c_a + 1} - 1}{2}
\end{aligned}$$



1 is more risk averse, and gets a large share of bad times, in return for a lesser share in good times. Notice this sharing rule is nonlinear.

INCOMPLETE MARKETS

- *Risk sharing in incomplete markets.*

In a complete market, we had $m^i = m^j$, and with power utility perfect correlation of individual and aggregate consumption growth. What about an incomplete market (meaning not all states are traded. We still have "complete" payoff spaces meaning no trading limits, bid ask spreads, short limits, etc.)

$$p = E(m^j x) = E(proj(m^j|X)x) = E\left[\left(proj(m^j|X) + \varepsilon\right) x\right] = E(x^* x)$$

$$\beta^j \left(\frac{c_{t+1}^j}{c_t^j}\right)^{-\gamma} = x_{t+1}^* + \varepsilon_{t+1}^j$$

Recall, there is a unique $x^* \in X$ such that $p = E(x^* x)$ Thus, *the projection of everybody's marginal utility on to traded assets is the same*, or *the mimicking portfolio for everybody's marginal utility is the same.* . People *use markets to share risk as much as possible.* The traded component of all m are all equal. This does not mean the portfolios they hold are equal. (More: See "Risk sharing is better than you think")

- *Pareto optimality in incomplete markets*

This competitive equilibrium is constrained Pareto efficient. Here is an example. Suppose the planner can only allocate across consumers with a state-contingent pattern given by $x(s)$. The planner chooses weights w_j that each person holds in the asset, with a constraint that the planner can only reallocate the endowment in each state:

$$\begin{aligned}
\max_{\{w_{ji}\}} \sum_j \lambda_j \left[u(c^j) + \beta_j \sum_s \pi(s) u[c^j(s)] \right] \text{ s.t. } c_j(s) &= e_j(s) + w_j x(s); \sum c_j = e^a \\
\text{total transfer} = 0 \qquad & \sum_j w_j = 0; \sum_j c_j(s) = e^a(s); \\
\frac{\partial}{\partial w_j} \quad \lambda_j \beta_j \sum_s \pi(s) u'(c^j(s)) x(s) &= \lambda_j E\left[\beta_j u'(c^j_{t+1}) x_{t+1}\right] = \Lambda_x \\
\frac{\partial}{\partial c_j} &: \lambda_j u'(c^j_t) = \Lambda
\end{aligned}$$

$$\text{"}p\text{"} = \frac{\Lambda_x}{\Lambda} = E\left[\beta_j \frac{u'(c^j_{t+1})}{u'(c^j_t)} x_{t+1}\right]$$

Again, the first order conditions of the constrained planning problem are the same as the competitive equilibrium condition, that $p = E(mx)$ for the assets x.

- *Incomplete markets, aggregating m vs. c, nonlinearity*

For *each person* we have

$$p_t = E(m^i_{t+1} x_{t+1}) = E\left(\beta \frac{u'(c^i_{t+1})}{u'(c^i_t)} x_{t+1}\right).$$

We therefore can always "aggregate" by averaging *marginal utilities* $\frac{1}{N}\sum_i \times$..

$$p_t = E\left[\left(\frac{1}{N}\sum_i m^i_{t+1}\right) x_{t+1}\right] = E\left(\left[\frac{1}{N}\sum_i \beta \frac{u'(c^i_{t+1})}{u'(c^i_t)}\right] x_{t+1}\right).$$

or, we can average over marginal utilities.

$$p_t \left[\frac{1}{N}\sum_i u'(c^i_t)\right] = E\left(\left[\frac{1}{N}\sum_i \beta u'(c^i_{t+1})\right] x_{t+1}\right).$$

Of course in complete markets, all m are the same, this is trivial. Why bother, you say, why not use the individual data and get N more moment conditions? I think the answer is measurement error. It's also a useful concept theoretically.

We cannot in general aggregate by averaging *consumption*

$$p_t \neq E\left(\beta \frac{u'(\frac{1}{N}\sum_i c^i_{t+1})}{u'(\frac{1}{N}\sum_i c^i_t)} x_{t+1}\right) = E\left(\beta \frac{u'(c^a_{t+1})}{u'(c^a_t)} R_{t+1}\right)$$

Brav, Constantinides and Geczy (2002) look in micro data and aggregate m not c

Note what's crucial here – *nonlinearities in marginal utility* are what causes the problem. If marginal utility were linear, we could go back and forth from consumption to marginal utility and aggregation would work fine. Let's pursue that idea

- *Quadratic utility*

With quadratic utility linear marginal utility, aggregation works nicely. in complete or incomplete markets. (Lars Hansen "calculating asset prices in three example economies" is a great paper on this. Also my "a mean variance benchmark" paper.).

Here is an example. People have quadratic utility, with stochastic individual-specific bliss points c_t^{*j}. (Since $\gamma = -cu''/u' = c/(c^* - c)$ different bliss points means different risk version) Markets are incomplete with all the idiosyncratic risk you could want.

$$p_t\left(c_t^{*j} - c_t^j\right) = E_t\left[\beta\left(c_{t+1}^{*j} - c_{t+1}^j\right)x_{t+1}\right]$$

Apply $\frac{1}{N}\sum_{j=1}^N$, and denote by a the average versions,

$$p_t\left(c_t^{*a} - c_t^a\right) = E_t\left[\beta\left(c_{t+1}^{*a} - c_{t+1}^a\right)x_{t+1}\right]$$

So, asset prices are exactly the same as those generated by a representative consumer with quadratic utility (same functional form). This is a great example that shows how all the problems of aggregation stem from nonlinear marginal utility.

- *Continuous time*

Another way to make utility linear is in continuous time, where the local linearity is all that matters. Then we *can* aggregate. (Grossman and Shiller (1982)). Again, there are N agents, with idiosyncratic risks and incomplete markets, and varying risk aversion. γ can be the local derivative of an arbitrary utility function, this does not have to be power utility. The basic asset pricing equation for agent i.

$$E_t\left(dR_t\right) - r_t^f dt = \gamma_i E_t\left(dR_t \frac{dC_t^i}{C_t^i}\right)$$

You can see again the partial risk sharing result:

$$\gamma_i \frac{dC_t^i}{C_t^i} = \frac{d\Lambda_t^*}{\Lambda_t^*} + \sigma_i dz_t^i \; E(dR_t, dz_t^i) = 0 \text{ for all } R, i$$

The tradeable component of marginal utility is exactly the same for everybody . Now, aggregation. The aggregate consumption is

$$\begin{aligned} C_t^a &= \sum_{i=1}^N C_t^i;\; C_t^a = \sum_{i=1}^N dC_t^i \\ \frac{dC_t^a}{C^a} &= \sum_{i=1}^N \left(\frac{C_t^i}{C_t^a}\right)\frac{dC_t^i}{C_t^i}. \end{aligned}$$

Thus, take a clever weighted sum of the individual first order conditions. Apply this to both sides

$$\frac{\sum_{i=1}^{N}\frac{1}{\gamma_i}\frac{C_t^i}{C_t^a}\times(\cdot)}{\sum_{i=1}^{N}\frac{1}{\gamma_i}\frac{C_t^i}{C_t^a}}$$

The result

$$E_t\left(dR_t\right)-r_t^f dt=\frac{\sum_{i=1}^{N}\frac{1}{\gamma_i}\frac{C_t^i}{C_t^a}\times\gamma_i E_t\left(dR_t\frac{dC_t^i}{C_t^i}\right)}{\sum_{i=1}^{N}\frac{1}{\gamma_i}\frac{C_t^i}{C_t^a}}$$

$$E_t\left(dR_t\right)-r_t^f dt=\frac{1}{\sum_{i=1}^{N}\frac{1}{\gamma_i}\frac{C_t^i}{C_t^a}}E_t\left(dR_t\frac{dC_t^a}{C_t^a}\right)$$

Or, cleaning it all up

$$\begin{aligned}E_t\left(dR_t\right)-r_t^f dt &= \gamma_a E_t\left(dR_t\frac{dC_t^a}{C_t^a}\right);\\ \frac{1}{\gamma_a} &= \sum_{i=1}^{N}\frac{1}{\gamma_i}\frac{C_t^i}{C_t^a}\end{aligned}$$

Take a step back and admire. Despite incomplete markets, arbitrary preference heterogeneity, arbitrary income heterogeneity, etc. etc., we obtain a representative investor based on aggregate consumption, whose "risk aversion" is a consumption-weighted average of the risk aversions of the individual agents.

The distribution *does* matter here – as less risk averse people get more wealthy, the representative investor becomes less risk averse. *Shocks* to the distribution matter. Though the perfect-market lambda weights were constant over time, if less risk averse people get more wealth, the aggregate becomes less risk averse.

And vice-versa. There is a simple mechanisms here that can generate the holy Grail, a time-varying risk premium. The high beta, low γ rich lose more in a downturn, shifting the aggregate risk aversion coefficient to higher values. Voilá.

- *Constantinides and Duffie*

The canonical example that of how to get heterogeneity to matter. (See treatment in "financial markets and the real economy") You can see how hard it is going to be to get individual shocks to matter to asset prices. Individual shocks are, by definition orthogonal to anything common to all, like asset returns. $m^j = x^* + \varepsilon^j$ and $E(\varepsilon^j x) = 0$. So how do we get ε^j to matter? Answer: we give *consumption* shocks, and then *nonlinear* marginal utility.

Here is the brilliantly simple example. Individual consumption is generated from $N(0,1)$ idiosyncratic shocks $\eta_{i,t+1}$ .First the great shock distribute sees C_{t+1}, then variable y_{t+1}, and then distributes shocks $\eta_{i,t+1}$)

$$\ln\left(\frac{C_{t+1}^i}{C_t^i}\right)=\ln\left(\frac{C_{t+1}}{C_t}\right)+\eta_{i,t+1}y_{t+1}-\frac{1}{2}y_{t+1}^2.$$

You can see by inspection that y_{t+1} is the cross-sectional variance of individual log consumption growth.

$$\sigma\left[\ln\left(\frac{C^i_{t+1}}{C^i_t}\right)|\frac{C_{t+1}}{C_t};y_{t+1}\right]=y_{t+1}$$

Aggregate consumption really is the sum of individual consumption – the $-\frac{1}{2}y^2_{t+1}$ term is there exactly for this reason:

$$E\left(\frac{C^i_{t+1}}{C^i_t}\middle|\frac{C_{t+1}}{C_t}\right)=\frac{C_{t+1}}{C_t}E\left(e^{\eta_{i,t+1}y_{t+1}-\frac{1}{2}y^2_{t+1}}\right)=\frac{C_{t+1}}{C_t}.$$

Now, the result:

$$1=E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}e^{\frac{\gamma(\gamma+1)}{2}y^2_{t+1}}R_{t+1}\right].$$

How do we get there? Start with the individual's first order conditions,

$$\begin{aligned}
1&=E_t\left[\beta\left(\frac{C^i_{t+1}}{C^i_t}\right)^{-\gamma}R_{t+1}\right]\\
&=E_t\left\{\beta E\left[\left(\frac{C^i_{t+1}}{C^i_t}\right)^{-\gamma}\middle|\frac{C_{t+1}}{C_t}\right]R_{t+1}\right\}\\
&=E_t\left\{\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}E\left[\left(\frac{C^i_{t+1}/C_{t+1}}{C^i_t/C_t}\right)^{-\gamma}\middle|\frac{C_{t+1}}{C_t}\right]R_{t+1}\right\}\\
&=E_t\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}e^{-\gamma\left(\eta_{i,t+1}y_{t+1}-\frac{1}{2}y^2_{t+1}\right)}R_{t+1}\right]\\
&=E_t\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}e^{\frac{1}{2}\gamma y^2_{t+1}+\frac{1}{2}\gamma^2y^2_{t+1}}R_{t+1}\right]\\
&=E_t\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}e^{\frac{\gamma(\gamma+1)}{2}y^2_{t+1}}R_{t+1}\right].
\end{aligned}$$

Back to the result. This is one of many (a general class really) of extensions of the consumption based model. It's of the form

$$1=E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}Z_{t+1}R_{t+1}\right].$$

and thus generates a two-factor model

$$E_t(R^e_{t+1})=\beta_{\Delta c}\lambda_{\Delta c}+\beta_Z\lambda_Z$$

To make it exciting, variation in β_Z (like hml) generates the big premium.

This paper is a construction proof: There *exists* a specification of the time-varying cross-sectional variance of shocks y_{t+1} that can rationalize any asset pricing anomaly. Neat, eh? And plausible. In bad times, it is certainly plausible that the cross-sectional dispersion goes up. I forgot who said "the great depression wasn't that bad if you had a job." And 75% did.

Now, to critiques. Like everything else in finance, it's quantitative. In micro data, does the cross-sectional dispersion of consumption go up *enough* to justify the equity premium? The value premium? Momentum? Preliminary work isn't that encouraging. But let's not give up too fast.

Second critique. *It all falls apart in continuous time.* It's building off the nonlinearity, and that nonlinearity disappears in continuous time. You need discrete trading, or you need to introduce jumps to avoid ito's lemma. Do we really trust the *nonlinearity* of marginal utility that much?

In sum... to get heterogeneity to matter, need to drive a wedge between *consumption* heterogeneity and *marginal utility* heterogeneity. That needs *nonlinear marginal utility*, incomplete markets, and jumps/discrete time.

- *Back to the Sharing rule*

In my little two person example, we ran into problems aggregating because the sharing rule was nonlinear

$$\left(\frac{\lambda_1}{\lambda_2}\right)^{-\frac{1}{\gamma_2}}\left(c_1^{\frac{\gamma_1}{\gamma_2}}+c_1\right)=c_a$$

That stopped us from cleanly connecting $\Lambda(s)$ to $c_a(s)$. In continuous time-ito's lemma though, we only look at local changes,

$$\left(\frac{\lambda_1}{\lambda_2}\right)^{-\frac{1}{\gamma_2}}\left(\frac{\gamma_1}{\gamma_2}c_1^{\frac{\gamma_1}{\gamma_2}-1}+1\right)dc_1+(\text{ito term})=dc_a$$

This is locally linear – and we build up to the nonlinear sharing rule because the derivatives change over time.