

DARK Configuration Proposal

May 31, 2023

Christopher D. Carroll

This document aims to outline the modifications we would need added to the configuration of models in Dolo yaml files in order for those config files to have the information necessary for them to be used for solving the kinds of models we currently solve in HARK.

1 The Problem - in Bellman Form

The usual analysis of dynamic stochastic programming problems packs a great many events (intertemporal choice, stochastic shocks, intertemporal returns, income growth, and more) into a small number of steps and variables. For the detailed analysis here, we will be careful to disarticulate everything that happens in the problem explicitly into separate steps so that each element can be scrutinized and understood in isolation.

We are interested in the behavior a consumer who begins stage t with a certain amount of ‘capital’ \mathbf{k}_t , which is immediately rewarded by a return factor R_t with the proceeds deposited in a bank account balance:

$$\mathbf{b}_t = \mathbf{k}_t R_t. \quad (1)$$

Simultaneously with the realization of the capital return, the consumer also receives noncapital income \mathbf{y}_t , which is determined by multiplying the consumer’s ‘permanent income’ \mathbf{p}_t by a transitory shock $\boldsymbol{\theta}_t$:

$$\mathbf{y}_t = \mathbf{p}_t \boldsymbol{\theta}_t \quad (2)$$

whose whose expectation is 1 (that is, before realization of the transitory shock, the consumer’s expectation is that actual income will on average be equal to permanent income \mathbf{p}_t).

The combination of bank balances \mathbf{b} and income \mathbf{y} define’s the consumer’s ‘market resources’ (sometimes called ‘cash-on-hand’, following ?):

$$\mathbf{m}_t = \mathbf{b}_t + \mathbf{y}_t, \quad (3)$$

which are available to be spent on consumption \mathbf{c}_t .

The consumer’s goal is to maximize discounted utility from consumption over the rest of a lifetime whose last stage is date T :

$$v_t(\mathbf{p}_t, \mathbf{m}_t) = \max_{\{\mathbf{c}\}_t^T} u(\mathbf{c}_t) + \mathbb{E}_t \left[\sum_{n=1}^{T-t} \beta^n \mathcal{L}_t^{t+n} \hat{\beta}_t^{t+n} u(\mathbf{c}_{t+n}) \right] \quad (4)$$

$$\begin{aligned}
\mathbf{a}_t &= \mathbf{m}_t - \mathbf{c}_t \\
\mathbf{p}_{t+1} &= \mathcal{G}_{t+1} \mathbf{p}_t \Psi_{t+1} \\
\mathbf{y}_{t+1} &= \mathbf{p}_{t+1} \boldsymbol{\theta}_{t+1} \\
\mathbf{m}_{t+1} &= R \mathbf{a}_t + \mathbf{y}_{t+1}
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}_t^{t+n} &: \text{probability to } \mathcal{L}\text{ive until age } t+n \text{ given alive at age } t \\
\hat{\beta}_t^{t+n} &: \text{age-varying discount factor between ages } t \text{ and } t+n \\
\Psi_t &: \text{mean-one shock to permanent income} \\
\sqsupset &: \text{time-invariant 'pure' discount factor}
\end{aligned}$$

Transitory and permanent shocks are distributed as follows:

$$\begin{aligned}
\Xi_s &= \begin{cases} 0 & \text{with probability } \wp > 0 \\ \boldsymbol{\theta}_s / \wp & \text{with probability } (1 - \wp), \text{ where } \log \boldsymbol{\theta}_s \sim \mathcal{N}(-\sigma_{\boldsymbol{\theta}}^2/2, \sigma_{\boldsymbol{\theta}}^2) \end{cases} \quad (5) \\
\log \Psi_s &\sim \mathcal{N}(-\sigma_{\Psi}^2/2, \sigma_{\Psi}^2)
\end{aligned}$$

where \wp is the probability of unemployment (and unemployment shocks are turned off after retirement). The problem can be normalized by the level of permanent income \mathbf{p} , yielding the resulting Bellman representation:

$$\begin{aligned}
v_t(m_t) &= \max_{c_t} \quad u(c_t) + \sqsupset \mathcal{L}_{t+1} \hat{\beta}_{t+1} \mathbb{E}_t[(\Psi_{t+1} \mathcal{G}_{t+1})^{1-\rho} v_{t+1}(m_{t+1})] \\
&\text{s.t.} \\
a_t &= m_t - c_t \\
m_{t+1} &= a_t \underbrace{\left(\frac{R}{\Psi_{t+1} \mathcal{G}_{t+1}} \right)}_{\equiv \mathcal{R}_{t+1}} + \boldsymbol{\theta}_{t+1}
\end{aligned}$$

Although our focus so far has been on the consumer's consumption problem once m is known, for constructing the computational solution it will be useful to distinguish a sequence of three moves (we use the word 'moves' to capture the fact that we are not really thinking of these as occurring at different moments in time – only that we are putting the things that happen within a given moment into an ordered sequence). The first move captures calculations that need to be performed before the shocks are realized, the middle move is after shocks have been realized but before the consumption decision has been made (this corresponds to the timing of v in the treatment above), and the final move captures the situation *after* the consumption decision has been made.

But even this breakdown of the problem is insufficient for the general case. As ? shows, some problems with multiple state or multiple control variables (or both) can be solved much more efficiently if the solution machinery lays the problem out as a sequence of single-control-variable choices. For example, one might have a consumption stage that results in some stock of assets, and then a portfolio choice stage that determines the allocation of those assets between investments in the risky and the safe asset.

We therefore want to allow for the problem in each period to have as many ‘stages’ as may be convenient for solving any particular model – say, a labor supply stage followed by a consumption stage followed by a portfolio choice stage.

We need to define notation for these three moves within a stage. We will use s_- as a marker for the move before shocks have been realized, s_\sim to indicate the move in which the choice is made, and s^+ as the indicator for the situation once the choice has been made.

$$v_{s_-}(k_s) = \mathbb{E}_{s_-}[v_{s_\sim}(\overbrace{k_s \mathcal{R}_s + \theta_s}^{=m_s})] \quad (6)$$

$$v_{s_\sim}(m_s) = u(c_s(m_s)) + v_{s^+}(a_s) \quad (7)$$

$$v_{s^+}(a_s) = \beta v_{s_-}(\underbrace{k_{s+1}}_{=a_s}) \quad (8)$$

If we turn off the mortality shocks ($\mathcal{L} = 1$) and family-size shocks, and redefine the pure time preference factor as $\beta = \beth$, the problem has first order condition

$$u'(c) = \beta R v'_{s^+}(m - c). \quad (9)$$

2 Creating An Instance of a Model

The first thing to do in building a model would probably be to create an empty model framework, perhaps with a command like:

where any information that was globally true for the model (like, maybe, its name, some links to external resources describing it, etc) would be encapsulated in ‘modelsetup.yaml’).

The next thing to do would be to define at least one ‘period’ in the model, where a period describes a set of things that are envisioned to happen in a single time period: where any information that was globally true for the period would be defined.

We want to define a ‘period’ as a standalone environment that must consist of at least one stage, each of which may have several moves, but in which all moves and stages share a common set of resources, which include parameterizations, symbol definitions, calibrations, etc. These common resources would be defined in a standalone yaml file – say, `bufferstockperiod.yaml` which would get digested as the first thing to be done as the construction of the period commences. The file might look something like this:

Next, one would construct at least one ‘stage,’ which consists of a sequence of moves that solve a problem. The task of defining the solution then be to define what happens serially backwards as each move gets added.

So, for example, for the buffer stock consumption problem we would have a file that defines the consumption optimization move (say, `coptstep.yaml`) of a problem, which would have an `equations` block that might look something like this:

where `vEndOfStge.deriv['a']` is a placeholder for whatever the correct notation would be to retrieve the derivative of `vEndOfStge` with respect to a and the assumption is that the equation defined as `arbitrage` presents an expression whose value should be zero at the optimum. The arbitrage equation therefore is equivalent to a rewritten version of our FOC (??):

$$u^c(c) - v_{s+}^a(m - c) = 0 \quad (10)$$

The result of processing this move should be the construction of the optimized value of c for each of the points in the state space defined as the input state space for each of the values of m constructed previously in the definition step, and construction of the `vMidOfStge` value function for the same states. (For the moment, we are not using EGM, because it is not clear how the syntax for EGM would work for dolo).

The next (earlier) move would compute the expectations that would contain the information necessary to construct the beginning-of-period value function. That file would need to define an operation that does not exist in dolo (as we currently understand it). That operation is basically to calculate the expectations of the value function (and other interesting objects) at the gridpoints already defined for the state variables.

The `equations` block of `vexp.yaml` might look something like this: where the expecto-rate operation would be defined to construct `vBegOfStge` by calculating the expectations at the grid of points for k defined by `cstagesetup.yaml`.

This would complete the definition of the solution of this stage, and if the problem were one in which there is only a consumption choice the foregoing pseudo-code would have constructed not only the solution for the consumption stage but the solution for the period as a whole (because the only stage is the consumption stage).

If we want to move back to construct the problem that takes the now-constructed period as its successor (that is, we want to continue the backward iteration to a preceding period), it is necessary to define a mapping between the end-of-period state variable(s) in period $t - 1$ and the beginning-of-period state variable(s) in period t . This `transit` operation belongs to neither period t nor to period $t - 1$ – it is the ‘connective tissue’ between them.

In the case where the previous period’s problem is another consumption problem identical to the one just defined, the `transit.yaml` file might look something like this: and possibly it would need to have further explicit information about things like how to map whatever grid exists for k in the current period to the corresponding grid for a in the prior period, and any information needed to map between the value functions.

With these files in place, the way to solve a period of the model might be something like this:

The payoff to all this work is that it should allow us to add components modularly. Suppose, for example, that we had the code to solve a portfolio choice problem of choosing the optimal share of risky versus safe assets; assuming that the risky return is realized at the same moment as the shocks to labor income (so that we can potentially allow the asset market return to be correlated with the labor supply decision – say, periods with bad stock market performance are also periods when unemployment is

higher), that decision problem should be insertable into the structure above between the `transit` and the `expectorate` move:

(This would also require specification of the stochastic process for the rate of return; following the scheme sketched above, this would presumably be located in the `cstagesetup.yaml` file, though possibly we might decide that it makes more sense to describe the stochastic shocks in the `expectorate.yaml` file that calculates the expectations over those shocks).

3 Maybe there should be two kinds of transitions

Arguably, the transition between periods is such a different phenomenon than within-period transitions that the best way to deal with it might be to declare that each ‘period’ is defined by a set of within-period steps, and a single transition equation that connects the period to its predecessor. In that case, the pseudocode for defining the solution to a period might look like this:

followed by .