Stackelberg Games of Water Extraction

Alain Jean-Marie
Université Côte d'Azur, Inria
LIRMM, University of Montpellier/CNRS, France

Mabel Tidball
INRA, Lameta, Montpellier, France

Fernando Ordóñez, Victor Bucarey López
Universidad de Chile, Santiago, Chile

Realized as part of the SticAmSud project DyGaMe

11th ISDG Workshop, Warsaw, July 2017

Ideas of the paper

We analyze a simple problem of water extraction.

- Myopic agents harvesting a common water resource
- A regulator has some control on the cost function
- The regulator takes into account rainfall and strategic interaction between agents

Ideas of the paper (ctd.)

- Groundwater extraction: marginal cost depends on the level of the aquifer
- In general, resources with accessibility problems: cost depends on scarcity
- The main ingredient: make the cost depend on the projected evolution of the resource: before or after the extraction or rainfall
- The goal: deduce its economical and environmental consequences
- The method: a Stackelberg Game
  linear-quadratic static game between agents,
  different kinds of optimization for the regulator.

The model of Provencher and Burt

We consider the extraction of groundwater by K players.
Dynamics of groundwater:

$$G_{t+1} = G_t + R - \sum_{k=1}^{K} u_t^k, \quad G_0, \quad \text{given.}$$

We suppose R is a constant.

Instantaneous profit:

$$\pi^i(u_t^i, G_t) = F_i(u_t^i) - C_i(G_t) \; \times \; u_t^i.$$

Marginal extraction cost ($C_i(.)$): depends on the current level of the groundwater.

The extended model

Introduce the more general instantaneous profit function:

$$\pi^i(u_t) = F_i(u_t^i) - C_i(G_t + mR - n\sum_j u_t^j)\, u_t^i$$

where $n, m \in [0, 1]$.

The extreme cases:

- $n = 0, m = 0$ (the standard case): cost based on current resource
- $n = 1, m = 1$: cost based on the state of the resource in the following period.

When $n \neq 0$ the profit function of player i depends on the action of the other player: strategic interaction not just through the dynamics.

Angles of analysis

In a previous work, we developped the non-cooperative, Nash-feedback solution for agents.

Here, we develop the supervised setting

- myopic followers, static Stackelberg leader
- myopic followers, dynamic Stackelberg leader

$\rightarrow$ an exercise in sensitivity analysis of LQ dynamic games, and non-LQ, non-concave optimal control problems.
Mostly work in progress...

Contents

The case of myopic agents

Assume agents play Nash with their instantaneous profit:

$$\max_{u^i} \left\{ F_i(u_t^i) - C_i(G_t + mR - n \sum_{k=1}^{K} u_t^k) \; u_t^i \right\}.$$

For convenience, we continue with the particular linear-quadratic functional form:

$$F_i(u) = a_i u - \frac{b_i}{2} u^2, \quad C_i(x) = z_i - c_i x > 0.$$

Myopic agent reaction

In the symmetric case we find:

$$u(G) = \underbrace{\frac{c}{b+(K+1)cn}}_{\alpha} G + \underbrace{\frac{a-z+cmR}{b+(K+1)cn}}_{\gamma}.$$

The value function of each player is:

$$\pi(G_0) = \frac{(cG_0+a-z+cmR)^2}{(b+(K+1)cn)^2} \frac{b+2cn}{2}.$$

Myopic stock dynamics

The stock dynamics is:

$$\begin{aligned} G_{t+1} &= G_t \;+\; R \;-\; K(\alpha G_t + \gamma) \\ &= (1 - K\alpha)^t G_0 + (R - K\gamma)\frac{1-(1-K\alpha)^t}{K\alpha} \;. \end{aligned}$$

The asymptotic stock is:

$$G_\infty \;=\; \frac{R - K\gamma}{K\alpha} \;=\; \frac{Rb}{Kc} + \frac{K+1}{K}\, Rn - Rm - \frac{a-z}{c} \;.$$

Contents

Supervised setting

A regulator is in charge of selecting the cost model, by choosing the cost parameters n and m.
The choice is announced to players, who play Nash.
Followers are myopic:

$$u_t^i = \frac{c}{b + (K+1)cn} G_t + \frac{a - z + cmR}{b + (K+1)cn}$$

The supervisor optimizes her own criterion, taking this reaction into account.

Optimal static choice

The supervisor gets to choose n and m once and for all.

Optimal pricing problem

The supervisor's problem is:

$$\max_{(n,m)\in[0,1]^2} \left\{ \sum_{t=0}^{\infty} \beta_L^t \sum_i \pi^i(u_t, G_t) \right\}$$

with the myopic agent reactions and stock dynamics

$$G_{t+1} = G_t + R - \sum_i u_t^i \ .$$

Optimal static choice (ctd.)

Thanks to the explicit solution for the dynamics:

$$V_{nm}(G_0) = \frac{b+2cn}{2}\left[\frac{\alpha^2(G_0-G_\infty)^2}{1-\beta(1-K\alpha)^2} + \frac{2R}{K}\,\frac{\alpha(G_0-G_\infty)}{1-\beta(1-K\alpha)} + \frac{R^2}{K^2(1-\beta)}\right]$$

with

$$G_\infty = \frac{R-K\gamma}{K\alpha}\,.$$

$$\alpha = \frac{c}{(K+1)cn+b} \qquad \gamma = \frac{mRc+a-z}{(K+1)cn+b}\,.$$

Optimal static choice (ctd.)

$\rightarrow$ optimization wrt (n, m) for each G_0.

Results:

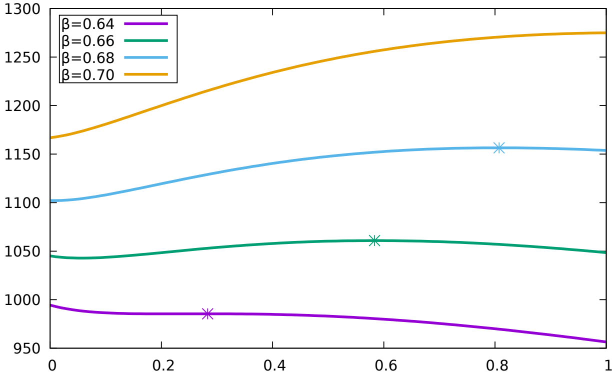
- $m^* = 1$ is always optimal if $K\alpha < 1$ (monotonous traj.)
- the optimal n is:

$$n^* = \begin{cases} 0 & \text{if } \beta_L \leq \underline{\beta}(G_0) \\ 1 & \text{if } \beta_L \geq \overline{\beta}(G_0) \\ n^*(\beta_L) & \text{otw.} \end{cases}$$

but $n^*(\beta_L)$ is the root of a 4^{th} degree polynomial.

Optimal static choice (ctd.)

A situation where $n^*(\beta_L) \notin \{0,1\}$:



Stackelberg with myopic followers

Followers being myopic, they apply the Nash controls. If symmetric:

$$u^i(G) = \frac{c}{b + (K+1)cn}\, G + \frac{a - z + cmR}{b + (K+1)cn}.$$

The benevolent regulator maximizes the total discounted profit on their behalf:

$$\max_{\{n_t, m_t\}} \left\{ \sum_{t=0}^{\infty} \beta_L^t \sum_i F(u^i(G_t)) - C(G_t + m_t R - n_t U(G_t)) u^i(G_t) \right\}$$

with the dynamics:

$$G_{t+1} \;=\; G_t \;+\; R \;-\; K u^i(G_t)\;.$$

$\Longrightarrow$ an optimal control problem.

First-order conditions

From the maximum principle:

$$
\begin{aligned}
0 &= \beta_L^t cU(G_t, n_t, m_t)(b + 2cn_t) + (b + (K+1)cn)(q_t - q_{t-1}) \\
&\quad -Kcq_t \\
0 &= \beta_L^t cU(G_t, n_t, m_t)^2(b - (K+1)(b + cn_t)) \\
&\quad +(b + (K+1)cn_t)(\lambda_t^{(n)} - \mu_t^{(n)}) \\
&\quad +Kcq_t(K+1)U(G_t, n_t, m_t) \\
0 &= \beta_L^t cRU(G_t, n_t, m_t)(b + 2cn_t) + (b + (K+1)cn_t)(\lambda_t^{(m)} - \mu_t^{(m)}) \\
&\quad -KcRq_t.
\end{aligned}
$$

where:

- $\lambda_t^{(n)}$, $\mu_t^{(n)}$ multipliers for $n_t \geq 0$ and $n_t \leq 1$, *id* for $\lambda_t^{(m)}$, $\mu_t^{(m)}$.
- $U(\dots)$ is the feedback function of followers.

Stationary situations

If a stationary state and control exists, then:

- optimal m: $m^* = 1$,
- optimal n:

$$n^* = \begin{cases} 0 & \text{if } \beta_L \leq \dfrac{b}{b+c} \\ 1 & \text{if } \beta_L \geq \dfrac{K(b+c)+c}{K(b+2c)+c} \\ \dfrac{K}{K+1} \, \dfrac{\beta(b+c)-b}{(1-\beta)c} & \text{otw.} \end{cases}$$

But we don't know if a stationary state exists.

Stationary situations (ctd.)

Example with corner solution:

$$a = 1, \quad b = 1, \quad c = \frac{6}{10}, \quad z = \frac{9}{10}, \quad R = 1, \quad \beta_L = \frac{95}{100}$$

gives a candidate optimal strationary solution:

$$G^\infty = \frac{9}{2}, \qquad n^* = 1.$$

Satisfies all first-order conditions.

Stationary situations (ctd.)

Example with interior solution:

$$a = 1, \quad b = 1, \quad c = \frac{1}{10}, \quad z = \frac{9}{10}, \quad R = 1, \quad \beta_L = 0.919$$

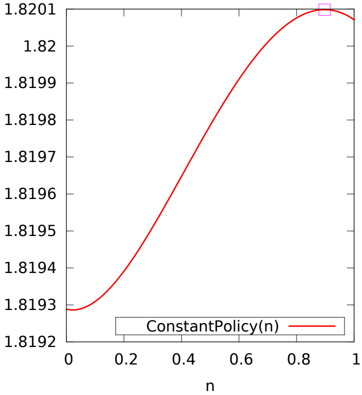
gives a candidate optimal strationary solution:

$$G^\infty \simeq 4.35, \qquad n^* = \frac{218}{243}.$$

Satisfies all first-order conditions.

Stationary situations (ctd.)

The constant-policy value (starting from the expected steady state) is maximum at n^*:



Non-stationary situations

However, convexity plays us tricks.
Example:

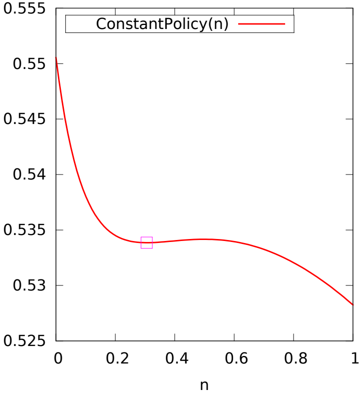
$$a = 1, \quad b = 1, \quad c = \frac{6}{10}, \quad z = \frac{9}{10}, \quad R = 1, \quad \beta_L = \frac{68}{100}$$

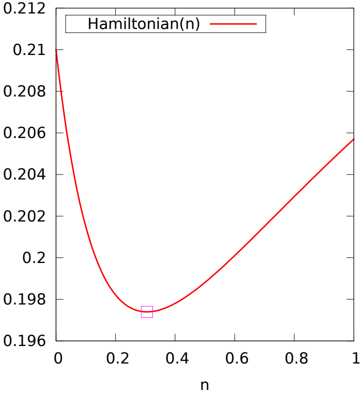
gives a candidate optimal strationary solution:

$$G^\infty = \frac{1}{8}, \qquad n^* = \frac{11}{36}.$$

But...
the optimal constant policy is not there!
And the Hamiltonian has actually a *local minimum* there wrt n.

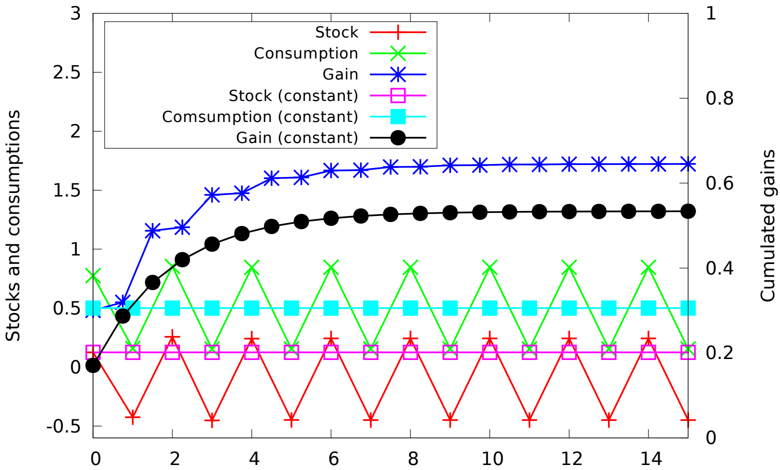
Non-stationary situations (ctd.)





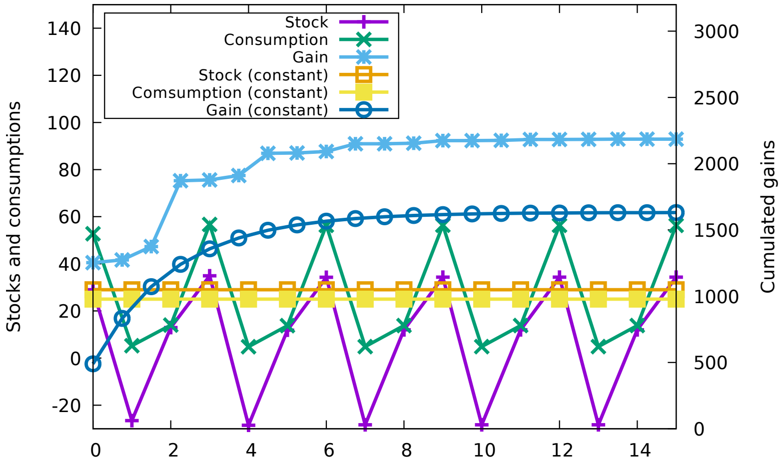
Hamiltonian is for constant trajectory from the expected steady state.

Non-stationary situations (ctd.)



Non-stationary situations (ctd.)

Same values except $R = 50$, $\beta = 7/10$:



A heuristic policy

These periodic policies were identified by using the following heuristic:

$$\begin{aligned} m(G) &= 1 \\ n(G) &= \text{argmax}\{V_{n1}(G) \mid 0 \leq n \leq 1\}. \end{aligned}$$

i.e. use the best static policy... but as state feedback.

Stackelberg with non-myopic followers

Followers being non-myopic, at time t they react to the sequence of "announced" regulations:

$$\{(n_t, m_t), (n_{t+1}, m_{t+1}), \ldots\}$$

while playing Nash!

$\Longrightarrow$ very complicated control law...

$\Longrightarrow$ not time-consistent?

Stackelberg with non-myopic followers (ctd)

A reasonable formulation: the supervisor announces feedback laws

$$G \mapsto n(G), \qquad G \mapsto m(G).$$

Followers play Nash Feedback with criterion:

$$\max_{\{u_t^i\}} \sum_0^\infty \beta_F^t [F_i(u_t^i) - C_i(G_t + m(G_t)R - n(G_t)\sum_k u_t^k)\ u_t^i],$$

and dynamics

$$G_{t+1} = G_t + R - \sum_k u_t^k, \quad G_0, \quad \text{given.}$$

Again an optimal control problem... to be studied.

Conclusions and extensions

Conclusions:

- regulator charges users in function of their behavior, not just in function of the level of resource
- introduce strategic interaction where there was none, in case of myopic agents ($n \neq 0$)
- difficult optimal control problem

Extensions under investigation

- simulations for the optimal control problem of the regulator
  $\rightarrow$ numerical methods, value iteration, policy iteration...
- non-myopic followers in the case of constant regulator policies
- stochastic case
  $\rightarrow$ learning algorithms under test
  $\rightarrow$ importance of m