Anabolic Persuasion*

Kfir Eliaz and Ran Spiegler†

March 3, 2021

Abstract

We present a model of optimal training of a rational, sluggish agent. A trainer commits to a discrete-time, finite-state Markov process that governs the evolution of training intensity. Subsequently, the agent monitors the state and adjusts his capacity at every period. Adjustments are incremental: the agent's capacity can only change by one unit at a time. The trainer's objective is to maximize the agent's capacity - evaluated according to its lowest value under the invariant distribution - subject to an upper bound on average training intensity. We characterize the trainer's optimal policy, and show how stochastic, time-varying training intensity can dramatically increase the long-run capacity of a rational, sluggish agent. We relate our theoretical findings to "periodization" training techniques in exercise physiology.

KEYWORDS: Exercise physiology, body building, persuasion, principal-agent, sluggish adjustment

*We acknowledge financial support from the Foerder Institute. We thank Martin Cripps, Michael Crystal, Israel Halperin, Nathan Hancart, Ron Peled and Mickey Scheinowitz for helpful comments. We owe special thanks to Oliver Fishman, who produced a fitness video that accompanied this paper (see https://youtu.be/onUrRs3kHIE).

†Eliaz: School of Economics, Tel-Aviv University and David Eccles School of Business, University of Utah. E-mail: kfire@tauex.tau.ac.il. Spiegler: School of Economics, Tel-Aviv University and Economics Dept., University College London and CFM. E-mail: rani@post.tau.ac.il.

1 Introduction

Economists have a long tradition of invading other academic disciplines. Lazear (2000) celebrates this so-called economic imperialism, demonstrating its value for such diverse fields as sociology, criminology and organizational behavior. Proponents of economic imperialism maintain that the ideas of individual rationality, forward-looking behavior, rational expectations and equilibrium analysis help us understand underlying mechanisms behind empirical regularities, and consequently guides policy interventions. Recently, economists applied the imperialistic approach to the field of epidemiology, in the context of the Covid-19 pandemic (Acemoglu et al. (2020)).

This paper carries the imperialistic approach to a new territory: it applies tools from economic theory to the field of exercise physiology, which studies the body's response to exercise and adaptation to exercise training to maximize human physical potential (for an introduction to this field, see Glass et al. (2014)). Specifically, we focus on the question of how the body's muscle mass responds to patterns of physical exercise. We demonstrate that by modeling the body as a forward-looking optimizing agent, we gain insights into the effectiveness of popular physical training strategies.

To describe the body as an optimizing agent, we need to specify its preferences. On the one hand, maintaining muscle mass is costly in terms of energy expenditure. The larger the mass, the higher the cost (Zurlo et al. (1990)). On the other hand, if muscle mass is too low relative to the demands of exercise, the body may incur the energy costs of repairing torn muscle tissue or inflammation (see Frankenfield (2006) and Faulkner et al. (1993)). Moreover, if the body lacks adequate muscle mass, it will not be able to complete the required physical task. It is plausible to assume that the body will record this performance gap as a cost. This cost can be interpreted in terms of psychological motivation, which itself may originate from evolutionary survival pressures (see Sagar and Stoeber (2009) and Lieberman (2015)). A more motivated trainee will record the performance gap as a larger cost relative to the muscle maintenance cost.

In a dynamic environment where the intensity of exercise changes stochastically over time, a key ingredient in modeling the body as an optimizing agent is its expectation of future demands. Here, too, we follow the economist's standard recipe and assume that the body has rational expectations - i.e., it knows the stochastic process that governs the future evolution of physical exercise (possibly as a result of some previous adaptive-learning phase).

Using these ingredients, we construct the following stylized discrete-time model. At period 0, a "trainer" commits to a strategy, which is a stochastic process over exercise intensity. We restrict ourselves to stochastic processes that follow a finite-state Markov chain. The average intensity (according to the chain's invariant distribution) cannot exceed some integer μ, which represents a "budget constraint" that limits the amount of resources that can be devoted to physical training.

Following the trainer's choice of strategy, at every subsequent period, the body (referred to as an "agent") observes the state of the trainer's Markov process and then chooses its muscle mass. We assume that the body can only make *incremental adjustments* to its muscle mass: at every period it can only change the mass by -1, 0 or 1 units. The body is an expected discounted utility maximizer, with a periodic payoff function that trades off the maintenance cost of muscle mass and the excess intensity of current physical exercise relative to current mass. Specifically, the agent's periodic cost when the current mass is m and the current intensity is d, equals $cm + \max\{0, d - m\}$, where c represents the maintenance cost (note that we measure intensity and mass on the same scale). We focus on the non-trivial case in which $c < 1$.

We impose the constraint that the agent has a best-reply to the trainer's strategy that induces a Markov process (over an extended state space that also includes m in the definition of a state) with a *unique* invariant distribution. The trainer's problem is to choose the Markov process to maximize the agent's long-run mass - evaluated according to its *minimal* realization under the invariant distribution. Our use of the max-min criterion is justified by the interpretation of m as a *capacity*: the higher the agent's minimal long-run mass, the higher the intensity he can consistently withstand in the long run.

The sluggish adjustment of muscle mass is a fundamental assumption in our model. Exercise intensity can fluctuate wildly between periods, but clearly, the body cannot change its muscle mass instantaneously to any level (see DeFreitas et al. (2011) and Counts et al. (2017)). If the body had perfect flexibility in adjusting its muscle mass, our model would be trivialized: at every period, muscle mass would be set such that the excess intensity gap $d - m$ is zero. As a result, under full flexibility, the long-run average mass will be at most μ. This is also the minimal long-run mass that the trainer can attain with a constant-intensity policy. Under such a policy, the distinction between sluggish and flexible agents is irrelevant. The question is whether using some non-degenerate Markov process will enable the trainer to outperform this benchmark when muscle adjustment is sluggish.

The trainer's problem is similar in spirit to Bayesian persuasion (Kamenica and Gentzkow (2011)). In a persuasion problem, the sender wants to increase the receiver's action; in our model the trainer wants to increase the agent's muscle mass. In a persuasion problem, the sender commits to a signal function; in our problem, the trainer commits to a Markov process. In a persuasion problem, the receiver's response to a signal realization is dictated by Bayesian updating; in our model, the agent's response to a realized state is constrained by sluggish adjustment. Finally, in a persuasion problem, the sender's ability to attain his objective is constrained by Bayes plausibility, which requires the average posterior belief to equal the prior; in our model, the trainer is constrained by the average intensity limit μ.

It is unclear how forward-looking the body is when adjusting its muscle mass. Accordingly, our analysis focuses on two extreme cases in terms of the agent's discount factor. We begin by analyzing a myopic agent with a discount factor of zero. In this case, the agent's adjustment rule is independent of the trainer's strategy: the mass moves up (down) a notch when the realized intensity is above (below) the current mass. We show that the trainer cannot attain a minimal long-run muscle mass above $2\mu - 1$. He can implement this upper bound using the following two-state Markov process. Physical intensity has two values, 0 and 2μ. The transition from 0 to 2μ is deterministic, while the transition from 2μ to 0 occurs with a probability that is arbitrarily close to one. This policy ensures that regardless of the initial muscle mass, the agent eventually oscillates between muscle mass levels 2μ and $2\mu - 1$. Thus, a properly designed stochastic training program can vastly outperform the flexible-adjustment benchmark.

We next turn to the case of an arbitrarily patient agent. We show that in this case, the trainer cannot attain a minimal long-run muscle mass above $\mu/c - 1$ (we assume that μ/c is an integer, for convenience). The trainer can approximate this bound arbitrarily well with a two-state Markov process that is similar to the one we construct for the myopic-agent. Physical intensity has two values, 0 and μ/c. The transition in one direction is deterministic, while the transition in the other direction is a function of c, such that the invariant probability of μ/c is almost c. This policy ensures that regardless of the initial muscle mass, the agent eventually oscillates between muscle mass levels μ/c and $\mu/c - 1$. Note that the upper bound on the minimal long-run mass in the patient-agent case is above the upper bound in the myopic-agent case if and only if $c < \frac{1}{2}$.

Our results suggest a rationale for the popular training technique of *periodization*, which structures the training regiment as a cycle with phases of high intensity physical load and recovery phases of low intensity. Since it first began in the 1960s, this methodology has gained popularity and is currently the dominant technique used by professional athletes. Numerous studies have documented the success of periodization in terms of increased muscle mass, increased muscle strength, greater endurance and athletic performance (Bompa and Buzzichelli (2018)).[1] While the physiological literature offers biological explanations for the superiority of a cyclical training design (e.g., see Issurin (2019)), our results provide a complementary perspective: the effectiveness of periodization techniques may stem from the rational yet sluggish adaptation of the body to fluctuations in physical stimuli.

Although our paper strictly follows the exercise-physiology interpretation of the model, the abstraction of the economic-modeling approach enables other interpretations. For example, the variables m and d may represent cognitive capacity and the intensity of cognitive activity, such that our results can be viewed in terms of training programs for building and maintaining cognitive skills. Moving entirely away from physiological or neurological interpretations, we can view the agent as an *organization*, like the military or a fire brigade. The problem is to design a drill program to build and maintain the organization's level of preparedness. The max-min criterion is appealing in this context. Sluggishness is a natural assumption in this setting: these organizations cannot drastically improve their level of preparedness overnight; and likewise, deterioration in preparedness tends to be gradual. Our analysis sheds light on the optimal design of a drill program for such organizations. More generally, we find the optimal design of a stochastic process for a sluggish agent to be an interesting (and, to our knowledge, new) problem from an abstract economic-theory perspective.

2 The Model

We consider a principal-agent model, in which the principal is referred to as a "trainer". We will focus on the interpretation that the agent is a physiological system that is trained to increase its capacity. The trainer commits

[1] For recent discussions of various periodization techniques, see Issurin (2010), Kiely (2012) and Kiely et al. (2019).

ex-ante to a pair (P, f), where P is a discrete-time, Markov process over some finite set of states S, and $f : S \to \mathbb{N}_+$ is an output function that assigns a challenge level to every state $s \in S$. We denote by d_t the challenge level in period t. When there is no risk of confusion we will replace the notation $f(s)$ with $d(s)$.

We impose the following constraints on (P, f). First, P has a unique invariant distribution λ_P. Second,

$$\sum_{s \in S} \lambda_P(s) f(s) \leq \mu$$

where $\mu \geq 1$ is an integer. That is, the long-run average challenge level is at most μ.

The agent knows the trainer's choice of (P, f). At every period t, he observes the realized state s_t and then chooses a non-negative capacity level $m_t \in \{m_{t-1} - 1, m_{t-1}, m_{t-1} + 1\}$. Henceforth, we refer to m_t as the agent's "mass" at time t. Let $m_0 \in \mathbb{N}_+$ be the agent's initial mass. The restricted choice set for m_t reflects the sluggish adaptation of the agent's mass.

The agent is an expected discounted utility maximizer with discount factor δ. His payoff at period t is

$$-[cm_t + \max(0, d_t - m_t)]$$

where $c \in (0, 1)$ and $d_t = f(s_t)$, where s_t is the state of the Markov process P at period t.

Note that given (P, f), the agent faces a Markov decision problem over an extended state space, where the state at period t is the pair (s_t, m_{t-1}). We impose the following additional constraint on the trainer: the extended Markov process over (s_t, m_{t-1}) that is induced by the agent's best-reply to (P, f) has a unique invariant distribution $\lambda^*_{(P,f)}$. This ensures that the long-run average mass, as well as the minimal long-run mass, are well-defined and independent of the initial condition m_0.

The trainer aims to maximize the agent's lowest muscle mass in the support of the invariant distribution $\lambda^*_{(P,f)}$. Formally, the trainer's problem can be stated as follows:

$$\sup_{(P,f)} \min\{m \mid \lambda^*_{(P,f)}(s,m) > 0 \text{ for some } s \in S\}$$

subject to the feasibility constraint

$$\sum_{(s,m)} \lambda^*_{(P,f)}(s,m) f(s) \leq \mu$$

Note that we use sup rather than max to state the trainer's problem. The reason is that the set of pairs (P, f) that induce an extended Markov process with a unique invariant distribution is not closed.

Interpretation

For the sake of expositional focus, we adopt the interpretation of the agent as the human body where m represents muscle mass. The trainer's objective is to maintain a high long-run level of muscle mass. The sup min criterion means that the trainer looks for the highest capacity that the agent's body *consistently* maintains in the long run. The constraint that the long-run average of d is at most μ captures a resource constraint in terms of the average amount of training per unit of time that the agent can afford. For example, the trainee can only devote a certain fraction of his time for training.

Exercise intensity d can be interpreted in terms of duration (e.g. the number of repetitions of a given exercise), load (e.g. lifting weight) or effort (e.g. running speed).[2] Of course, the stylized nature of our model abstracts from such fine distinctions. However, the interpretation of the resource constraint does depend on the meaning of exercise intensity. If d represents exercise duration, then μ is the average amount of time per period that the trainee

[2] See Steele (2014) and Steele et al. (2017) for discussions of these different notions of intensity.

can devote to physical exercise. If, however, d represents load or effort, μ is perhaps better viewed as a parameter of the trainer's problem than an exogenous constraint.

We view the human body as an optimizing system that minimizes costs. The sluggishness assumption captures the fact that the body cannot instantaneously change its body mass to any level, but does so in increments (even though the challenge levels may fluctuate dramatically between periods). The body's periodic cost function incorporates two factors. First, cm_t is the caloric maintenance cost of muscle mass m_t. Second, the gap between m_t and d_t (when the latter is higher) represents a performance shortfall because the agent's capacity is lower than the challenge it faces. This cost can be purely physical reflecting body stress, inflammation or the repair of torn tissues. Alternatively, it may be viewed as motivational: the body internalizes the trainee's psychological motivation to complete challenges. Thus, the higher the agent's minimal long-run mass, the higher the challenge level that he is guaranteed to meet in the long run.

Under this interpretation, our model endows the human body with rational expectations because it has knowledge of (P, f) when making a decision. The rationale for this assumption is that the body forms adaptive expectations based on a long memory. Hence, it is reasonable to assume that in the long-run it will be able to learn Markov processes with sufficiently low dimensionality. This is one of our reasons for restricting the trainer to finite-state Markov processes (the other reason being expositional simplicity).

The adaptive-expectations rationale also underlies our restriction that the trainer does not condition d_t on past realizations of m. Under rational expectations, if the trainer could condition d on past choices of m, he would be able to incentivize the agent using off-equilibrium threats. For instance, he could incentivize to increase muscle mass using a policy of zero challenges sustained by a threat to switch to extreme challenges if m fails to increase. We find such policies absurd in the physiological context, and therefore, rule them out by assuming that the trainer does not condition on m. Under our alternative interpretation of the agent as an organization, it is questionable whether the trainer can monitor m, which represents the organization's level of preparedness.[3]

Benchmark: Completely flexible adjustment

Suppose the agent could choose any $m_t \in \mathbb{N}_+$ at every period. Then, since $c \in (0, 1)$, it would be optimal for him to choose $m_t = d_t$ at every t. This means that the long-run average of m_t would coincide with the long-run average of d_t, which by assumption cannot exceed μ. Therefore, the best the trainer can do according to his $\sup\min$ criterion is to play a constant strategy $d_t = \mu$ at every period, such that this flexible agent's mass will be μ as well. The same deterministic process attains the same long-run mass of μ also when the agent is sluggish (because the agent will eventually reach this mass and stay there indefinitely). The question is whether the trainer can outperform this benchmark with a non-degenerate Markov process.

3 A Myopic Agent

In this section we analyze the trainer's problem when $\delta = 0$ - i.e., the agent is myopic. We show that by using a non-degenerate Markov process over d, the trainer can increase the agent's minimal long-run mass by a large margin (nearly twice for large μ), relative to the flexible-agent benchmark.

Proposition 1 *Let $\delta = 0$. Then:*

(i) For any trainer strategy, the minimal long-run mass induced by the agent's best-reply is bounded from above by $2\mu - 1$.

(ii) This upper bound can be implemented by the following (P, f). The Markov

[3] We conjecture that if the trainer can condition d_t on m_{t-1}, the results in our paper will not change.

process P has two states, H and L and a transition matrix given by

$\Pr(s_t \to s_{t+1})$	L	H
L	0	1
H	β	$1-\beta$

where β is arbitrarily close to 1. The output function is $f(H) = 2\mu$ and $f(L) = 0$. In the $\beta \to 1$ limit, the invariant mass distribution assigns probability $\frac{1}{2}$ to $m = 2\mu$ and $m = 2\mu - 1$.

Thus, we see that in the presence of a sluggish agent, a slightly perturbed cyclic training program can dramatically increase the minimal long-run mass. The trainer's "training regime" approximately consists of alternating periods of high intensity ($d = 2\mu$) and rest ($d = 0$). However, after a period of high intensity training, there is an arbitrarily small chance $1 - \beta$ that the high-intensity episode will be repeated. This stochastic perturbation ensures that the set of mass values $\{2\mu, 2\mu - 1\}$ is absorbing: the agent will reach it in finite time with probability one, regardless of m_0.

Proof of part (i) of Proposition 1

The proof proceeds by a series of steps. Recall that we use the notation $d(s)$ as a substitute for $f(s)$.

Step 1: The agent's strategy

Consider the agent's move at period t, given the extended state (s_t, m_{t-1}). A myopic agent will choose m_t to minimize $cm_t + \max(0, d(s_t) - m_t)$. Therefore, we can immediately pin down the agent's behavior, independently of the trainer's strategy. Since $c \in (0, 1)$, we obtain the following: if $d(s_t) > m_{t-1}$, the agent will choose $m_t = m_{t-1} + 1$; if $d(s_t) < m_{t-1}$, the agent will choose $m_t = m_{t-1} - 1$; and if $d(s_t) = m_{t-1}$, the agent will choose $m_t = m_{t-1}$. That is, the agent will always adjust his mass in the direction of the current level of d. $\square$

Consider an arbitrary strategy for the trainer, which induces an extended Markov process with a unique invariant distribution. Let $(m_{t-1}, d_t)_{t=1,2,\ldots}$ be a possible sample path that results from the extended process. By the unique-invariant-distribution requirement, the extended process is ergodic. Therefore, the long-run frequency of every (m, d) in the sample path coincides with the probability of this pair according to the invariant distribution. Let $\lambda(m, d)$ denote the probability of (m, d) according to the invariant distribution, as well as the frequency of (m, d) in the sample path. Let X be the set of recurrent pairs (m, d) in the sample path (omitting the time subscripts $t-1$ and t from m and d). Partition X into three classes:

$$\begin{aligned} X^+ &= \{(m, d) \in X \mid d > m\} \\ X^- &= \{(m, d) \in X \mid d < m\} \\ X^0 &= \{(m, d) \in X \mid d = m\} \end{aligned}$$

Step 2: Showing that

$$\sum_{(m,d)\in X^+} \lambda(m, d)(m + 1) = \sum_{(m,d)\in X^-} \lambda(m, d)m \tag{1}$$

Consider some period t along the sample path such that $(m_t, d_{t+1}) \in X^+$. By definition, this pair is recurrent. Therefore, it must be visited again in some later period. Let $t' + 1$ be the earliest such period. Since m moves only in one-unit increments, it must be the case that $(m_{t'}, d_{t'+1}) \in X^-$ and $m_{t'} = m_t + 1$. We have thus defined a one-to-one mapping from periods t for which $(m_t, d_{t+1}) \in X^+$ to periods t' for which $(m_{t'}, d_{t'+1}) \in X^-$, such that $m_{t'} = m_t + 1$. In a similar way, we can define a one-to-one mapping from periods t for which $(m_t, d_{t+1}) \in X^-$ to periods t' for which $(m_{t'}, d_{t'+1}) \in X^+$, such that $m_{t'} = m_t - 1$. It follows that

$$\lim_{T\to\infty} \frac{\sum_{t=1}^{T} \mathbf{1}[(m_t, d_{t+1}) \in X^+] \cdot (m_t + 1)}{T} = \lim_{T\to\infty} \frac{\sum_{t=1}^{T} \mathbf{1}[(m_t, d_{t+1}) \in X^-] \cdot m_t}{T}$$

which can be rewritten as (1). □

Step 3: Showing the long-run *average* of m is at most 2μ

The long-run average of m induced by the trainer's strategy can be written as

$$E(m) = \sum_{(m,d)\in X^+} \lambda(m,d)m + \sum_{(m,d)\in X^-} \lambda(m,d)m + \sum_{(m,d)\in X^0} \lambda(m,d)m \qquad (2)$$

By the feasibility constraint,

$$\sum_{(m,d)\in X^+} \lambda(m,d)d + \sum_{(m,d)\in X^-} \lambda(m,d)d + \sum_{(m,d)\in X^0} \lambda(m,d)d \leq \mu$$

By definition, $d \geq m+1$ for every $(m,d) \in X^+$, $d \geq 0$ for every $(m,d) \in X$, and $d = m$ for every $(m,d) \in X^0$. Therefore,

$$\sum_{(m,d)\in X^+} \lambda(m,d)(m+1) + \sum_{(m,d)\in X^-} \lambda(m,d)\cdot 0 + \sum_{(m,d)\in X^0} \lambda(m,d)m \leq \mu$$

This means that

$$\sum_{(m,d)\in X^+} \lambda(m,d)m \leq \sum_{(m,d)\in X^+} \lambda(m,d)(m+1) \leq \mu - \sum_{(m,d)\in X^0} \lambda(m,d)m$$

By (1), it follows that

$$\sum_{(m,d)\in X^-} \lambda(m,d)m \leq \mu - \sum_{(m,d)\in X^0} \lambda(m,d)m$$

as well. Plugging the last two inequalities in (2), we obtain

$$E(m) \leq 2\mu - \sum_{(m,d)\in X^0} \lambda(m,d)m \leq 2\mu$$

□

Step 4: Showing the minimal long-run value of m is at most $2\mu - 1$
Suppose that the long-run distribution over d is degenerate at some d^*. Therefore, $d^* \leq \mu$. The agent's myopic best-reply implies that eventually, his mass will coincide with d^*. It follows that in order to reach a minimal long-run mass above μ, the long-run distribution over d must assign positive probability to at least two values. This means that there will be infinitely many periods t in which $d_t \neq m_{t-1}$. By the agent's myopic best-replying, this precludes the possibility that the long-run distribution over m is degenerate. Since the long-run average of m is at most 2μ, there must be infinitely many periods t in which $m_t \leq 2\mu - 1$. This completes the proof of part (i). □

Proof of part (ii) of Proposition 1
Consider the trainer's strategy described in part (ii) of the statement of the result. As long as $\beta \in (0, 1)$, the Markov process over m that is induced by the strategy and the agent's best-reply (given by Step 1) has a unique invariant distribution, with $m = 2\mu$ and $m = 2\mu - 1$ being the only recurrent mass values. The reason is that if $m_t > 2\mu$, $m_{t+1} = m_t - 1$ with certainty; if $m_t < 2\mu - 1$, there is a positive probability that there will be a streak of realizations $d = 2\mu$ such that m will keep adjusting upward until it reaches $m = 2\mu$; and finally, if $d_t = 0$ then $d_{t+1} = 2\mu$ for sure, which means that once m hits 2μ and later goes down to $2\mu - 1$, it will return to 2μ immediately in the next period. Finally, in the $\beta \to 1$ limit, the invariant distribution over m assigns probability $\frac{1}{2}$ to each of the values $m = 2\mu$ and $m = 2\mu - 1$. This completes the proof of part (ii). ■

4 A Patient Agent

In this section we characterize the solution to the trainer's problem when the agent is forward-looking and arbitrarily patient. For expositional convenience, we assume that μ/c is an integer.

Proposition 2 *Let δ be arbitrarily close to 1. Then:*

(i) The minimal long-run mass at the solution to the trainer's problem is bounded from above by $\mu/c - 1$.

(ii) This upper bound can be approximated arbitrarily well by (P, f) with the following properties. The Markov process P has two states, H and L, and a transition matrix given by

$\Pr(s_t \to s_{t+1})$	L	H
L	$1-\alpha$	α
H	β	$1-\beta$

where $\alpha = 1$ if $c \geq \frac{1}{2}$, $\beta = 1$ if $c < \frac{1}{2}$, and $\alpha/(\alpha+\beta)$ is arbitrarily close to c from above. The output function is $f(H) = \mu/c$ and $f(L) = 0$. In the $\alpha/(\alpha+\beta) \to c$ limit, the invariant mass distribution assigns probability c to $m = \mu/c$ and probability $1-c$ to $m = \mu/c - 1$.

The upper bound on the agent's minimal long-run mass is higher than in the myopic benchmark whenever $c < \frac{1}{2}$. Moreover, it gets arbitrarily high when $c \to 0$. In contrast, when c is close to one, the highest minimal long-run mass is close to the flexible-agent benchmark μ.[4]

The structure of the Markov process that approximates the upper bound is similar to the one we constructed for the myopic-agent case. The main difference is that persistence of one of the two states occurs with non-vanishing probability. When $c < \frac{1}{2}$, a "rest period" (corresponding to the state L) is followed by another one with probability approximately equal to $(1-2c)/(1-c)$. When $c > \frac{1}{2}$, a high-intensity training period (corresponding to the state H) is followed by another one with probability $(2c-1)/c$.

Compare this result with the myopic case of the previous section. The myopic agent only responds to the realizations of d, disregarding the trainer's

[4]Because μ/c is an integer, we rule out the possibility that c is arbitrarily close to one. In that case, the trainer cannot outperform the flexible-agent benchmark of μ.

overall strategy. In contrast, the patient agent reacts to the trainer's strategy. In particular, when $c < \frac{1}{2}$, the trainer's program allows for a streak of $d = 0$ realizations. When this happens, the patient agent does not lower his mass below $\mu/c - 1$, because he takes into account the future loss $d - m$ in the event that d switches to $d = \mu/c$. The trainer designs the transition probabilities such that the patient agent's intertemporal trade-offs lead him to be indifferent between lowering his mass and remaining at $m = \mu/c - 1$. In contrast, the myopic agent cannot be made indifferent when faced with a streak $d = 0$ realizations: he repeatedly lowers his mass. This difference enables the trainer to achieve a higher minimal long-run mass when the agent is patient, as long as $c < \frac{1}{2}$.

We now turn to the proof of Proposition 2. In our proof of part (i), we actually prove a somewhat stronger result: in order to attain a strictly positive minimal long-run mass, the *average* long-run mass cannot exceed $\mu/c - 1 + c$. The Markov process we construct in part (ii) approximates this upper bound. This means that among all trainer strategies that attain the minimal long-run mass of $\mu/c - 1$, this process is the best in terms of average mass.

Proof of part (i) of Proposition 2

Let p be the joint invariant distribution over (d, m) that results from the trainer's strategy and the agent's best-replying strategy. Recall that the existence of such a distribution is a constraint on the trainer's problem. We abuse notation and write $p(d)$, $p(m)$ and $p(d \mid m)$ to represent marginal and conditional distributions induced by p. As in the myopic-agent case, we first derive an upper bound on the expected mass according to p, which we then use to derive the upper bound on the minimal long-run mass. Finally, we show how to implement this upper bound.

In Section 2, we saw that the trainer can implement a minimal long-run mass of at least μ (by playing $d = \mu$ at every period). Therefore, we take it for granted that the minimal value of m in the support of p is at least $\mu \geq 1$.

Step 1: $p(d>0)\geq c$
Consider the following deviation by the agent. Pick some period-t history for which $m_{t-1}\geq 1$ is at the lowest recurrent value according to p. Therefore, $m_t=m\in\{m_{t-1},m_{t-1}+1\}$. At this history, the agent deviates to $m_t'=m-1$. Subsequently, the agent behaves according to his original strategy *as if the deviation did not occur.*

This deviating strategy induces an invariant distribution p' such that for every (d,m) in the support of p, $p'(d,m-1)=p(d,m)$. Therefore, the deviation saves c at every period, but raises costs by one unit per period whenever $d\geq m$ under the original strategy. In order for this deviation to be unprofitable for an arbitrarily patient agent, it must be the case that $p(d\geq m)\geq c$. Since $m>0$ with probability one, $p(d>0)\geq p(d\geq m)$, hence $p(d>0)\geq c$**.** □

Step 2: The expectation of m **according to** p **is at most** $\mu/c-1+c$**.** Assume the contrary. Then, the agent's average long-run cost exceeds

$$c\cdot[\frac{\mu}{c}-1+c]=\mu-c(1-c)$$

Now consider a deviation to the following strategy. Descend from m_0 to $m=0$, and then implement the following rule: $m_t=0$ whenever $d_t=0$, and $m_t=1$ whenever $d_t>0$. When the agent is arbitrarily patient, the average long-run cost from this strategy is approximately

$$\begin{aligned} p(d &= 0)\cdot 0+p(d>0)\cdot[c+\sum_{d>0}p(d\mid d>0)d-1] \\ &\leq p(d>0)(c-1)+\mu \end{aligned}$$

By Step 1,

$$p(d>0)(c-1)+\mu<\mu-c(1-c)$$

such that the deviation is profitable, a contradiction. □

Step 3: The minimal long-run mass is at most $\mu/c - 1$.
By assumption, μ/c is an integer. Therefore, $\mu/c - 1 + c$ is not an integer. Therefore, in order for average long-run cost to be weakly below $\mu/c - 1 + c$, the minimal long-run mass cannot exceed $\mu/c - 1$.[5] This completes the proof of part (i). □

Proof of part (ii) of Proposition 2
Consider the trainer's strategy that is described in the statement of part (ii). Our objective is to show that given this strategy, there is a best-reply for the agent that induces the following joint invariant distribution over d and m: $m = \mu/c$ whenever $s = H$ and $m = \mu/c - 1$ whenever $s = L$.

Since the agent faces a Markovian decision problem with an extended state space (s, m), there exists a best-reply that is Markovian with respect to this state space. To derive such a best reply, we proceed in four steps.

Step 1: *There is no best-reply in which the invariant distribution assigns probability one to a single m.*
Proof. Assume the contrary. If $m < \mu/c$, then it is profitable for the agent to deviate to a strategy that plays $m + 1$ whenever $s = H$ and m whenever $s = L$. Likewise, if $m > 0$, it is profitable for the agent to deviate to a strategy that plays m whenever $s = H$ and $m - 1$ whenever $s = L$. □

Step 2: The set of recurrent values of m (according to the unique invariant distribution induced by the two parties' strategies) is a set of consecutive numbers $\underline{m}, \underline{m} + 1, ..., \overline{m}$, where $\overline{m} \leq \mu/c$.
Proof. The agent's sluggishness implies that if the agent visits two non-adjacent masses m and m', then he must also visit every m'' between them. Therefore, if m and m' are recurrent, so is m''. Suppose $\overline{m} > \mu/c$. Then, there is a profitable deviation from the agent that instructs the agent to remain at $\overline{m} - 1$ whenever the original strategy instructs him to switch to

[5] The proof of this step utilizes the convenient assumption that μ/c is an integer. An alternative proof that does not rely on this assumption is analogous to Step 4 in the proof of Proposition 1.

$\overline{m}$. $\square$

Step 3: *There is a best-reply that induces an invariant distribution that assigns positive probability to exactly two values of* m.

Proof. Consider the invariant distribution over (d, m) induced by the trainer's strategy and the agent's best-reply. By Step 1, $\overline{m} - \underline{m} \geq 1$. If $\overline{m} - \underline{m} = 1$, we are done. Therefore, assume $\overline{m} - \underline{m} > 1$. There are two cases to consider.

First, let $\alpha = 1$ (this fits the case of $c \geq 1/2$). This means that whenever $s = L$, the state switches immediately to $s = H$ in the next period. Consider the top two values of m in the invariant distribution, namely $\overline{m}$ and $\overline{m}-1$. By Step 2, $\overline{m} \leq \mu/c$. Moreover, when $s = L$ (at which d attains its lowest value according to the trainer's strategy), the agent strictly prefers $\overline{m} - 1$ to $\overline{m}$. Consider some t for which $m_t = \overline{m}$ (there are infinitely such periods because $\overline{m}$ is recurrent). If $s_{t+1} = L$, the agent necessarily switches to $m_{t+1} = \overline{m} - 1$. If, on the other hand, $s_{t+1} = H$, we need to consider two possibilities.

- Suppose that when $s_{t+1} = H$, it is not optimal for the agent to play $m_{t+1} = \overline{m}$. That is, the agent switches from $m_t = \overline{m}$ to $m_{t+1} = \overline{m} - 1$ for *any* realization of s_{t+1}. But this also means that if $m_{t'} = \overline{m} - 1$ at some period t' and $s_{t'+1} = H$, it cannot be optimal for the agent to switch to $m_{t'+1} = \overline{m}$. The reason is that by revealed preference, the agent prefers being at $\overline{m} - 1$ to being at $\overline{m}$ when the state is H. And since we already saw that the agent prefers being at $\overline{m} - 1$ to being at $\overline{m}$ when the state is L, this means that the agent will *never* switch from $\overline{m} - 1$ to $\overline{m}$, contradicting the definition of $\overline{m}$ as a recurrent state.

- Suppose that when $s_{t+1} = H$, it is optimal for the agent to play $m_{t+1} = \overline{m}$. This reveals a weak preference for $\overline{m}$ over $\overline{m}-1$ when the state is H. Therefore, there is a best-reply for the agent that prescribes $m_{t+1} = \overline{m}$ whenever the extended state (s_{t+1}, m_t) is $(H, \overline{m} - 1)$ or $(H, \overline{m})$. We already saw that when the extended state is $(L, \overline{m})$, the agent switches to $\overline{m} - 1$. Since $\alpha = 1$, this means that we have constructed a best-reply for the agent such that once he reaches $\overline{m}$, he will only visit $\overline{m}$ and $\overline{m}-1$ from that period on, contradicting the assumption that there are additional recurrent values of m.

Thus, we have ruled out the possibility that $\overline{m} - \underline{m} > 1$ when $\alpha = 1$. Now suppose $\beta = 1$ (this fits the case of $c \leq 1/2$). An analogous argument establishes that there is a best-reply for the agent that induces an invariant distribution with only two recurrent mass values, $\underline{m}$ and $\underline{m} + 1$.

It follows that we can restrict attention to strategies of the agent that induce an invariant distribution which assigns positive probability to precisely two consecutive mass values, m and $m - 1$, where $0 < m \leq \mu/c$. □

Step 4: *There is a best-reply for the agent that induces an invariant distribution on the mass values μ/c and $\mu/c - 1$.*
Proof. Given Step 3, it is clearly optimal for the agent to be at m when $s = H$ and at $m-1$ when $s = L$. In addition, when $m > \mu/c$ ($m < \mu/c - 1$), the agent clearly wants to move downward (upward).

The invariant distribution of the trainer's two-state Markov process assigns probability $\alpha/(\alpha+\beta)$ to state H and $\beta/(\alpha+\beta)$ to state L. Therefore, since the agent is arbitrarily patient, his long-run expected payoff is approximately

$$-\frac{\alpha}{\alpha+\beta} \cdot (cm + \frac{\mu}{c} - m) - \frac{\beta}{\alpha+\beta} \cdot c(m-1)$$

It is now easy to see that given that $\alpha/(\alpha+\beta) > c$, this expression increases with m, such that the optimal value of m is μ/c. The expected value of m according to this strategy is

$$\frac{\alpha}{\alpha+\beta} \cdot \frac{\mu}{c} + \frac{\beta}{\alpha+\beta} \cdot (\frac{\mu}{c} - 1)$$

which is arbitrarily close to the upper bound. ■

5 Comment: The Trainer's $\sup\min$ Criterion

In our model, the trainer's criterion is the agent's $\sup\min$ long-run mass. Alternatively, we could use the long-run *average* mass as a criterion. However, this criterion is less attractive in our context because it does not reflect the idea of "preparedness" - namely, that the body should be able to perform at a *consistently* high level. In particular, the average criterion allows zero to be a recurrent value for the agent's mass (and consequently, his level of preparedness).

A by-product of our analysis in Section 3 is that in the myopic-agent case, 2μ is an upper bound on the average long-run mass that the trainer can attain. It can be shown that this upper bound can be approximated arbitrarily well, but this must come at the price of arbitrarily long recurrent stretches of $m_t = 0$ realizations (which are compensated for by periods in which m_t reaches arbitrarily high values). Obviously, such paths imply that the agent cannot consistently meet positive challenge levels. By comparison, the process we constructed in Section 3 induces an average long-run mass of approximately $2\mu - \frac{1}{2}$, and a $\sup\min$ long-run mass of $2\mu - 1$.

A similar diagnosis pertains to the patient-agent case of Section 4. An upper bound on the average long-run mass is μ/c. The reason is that if the average mass exceeds this value, it implies that the agent's average long-run cost is above μ. However, the agent can ensure an average cost of μ by always playing $m = 0$, hence a long-run mass in excess of μ/c is inconsistent with the agent's best-replying. We believe that as in the myopic-agent case, this upper bound can be approximated arbitrarily well. However, as in the myopic-agent case, it can be shown that recurrent stretches of $m_t = 0$ realizations are necessary for this - which, once again, fail the $\sup\min$ criterion miserably. By comparison, the process we constructed in Section 4 induces an average long-run mass of approximately $\mu/c - 1 + c$, and a $\sup\min$ long-run mass of $\mu/c - 1$.

6 Conclusion

In this paper we presented a theoretical approach to the subject of exercise physiology, based on the view of the human body as a forward-looking optimizing agent which is nevertheless constrained by sluggish adjustment. We saw that this very sluggishness is actually a boon to physical trainers: using a stochastic training strategy that resembles popular "periodization" techniques, the trainer can achieve a significantly higher long-run muscle mass than if the body could instantaneously adjust its mass to physical stress.

We believe that thanks to its abstraction, our modeling approach can be extended to related problems, such as the optimal design of dynamic dieting regimes. A model that describes the body's metabolism as a consequence of dynamic *sluggish* optimization with rational expectations may shed light on prevalent dieting programs such as carb cycles. We hope to pursue this approach before the next pandemic.

References

[1] Acemoglu, D., Chernozhukov, V., Werning, I., and Whinston, M. D. (2020), A multi-risk SIR model with optimally targeted lockdown (no. w27102), National Bureau of Economic Research.

[2] Bompa, T. and Buzzichelli, C. (2018), Periodization: theory and methodology of training. Human kinetics.

[3] Cheval, B., Radel, R., Neva, J. L., Boyd, L. A., Swinnen, S. P., Sander, D. and Boisgontier, M. P. (2018), Behavioral and neural evidence of the rewarding value of exercise behaviors: a systematic review, Sports Medicine 48(6), 1389–1404.

[4] R. Counts, B.R., S.L. Buckner, J.G. Mouser, S.J. Dankel, M.B. Jessee, K.T. Mattocks, and J.P. Loenneke (2017), Muscle growth: To infinity and beyond? Muscle and Nerve 56(6), 1022-1030.

[5] DeFreitas, J.M., T.W. Beck, M.S. Stock, M.A. Dillon and P.R. Kasishke II (2011), An examination of the time course of training-induced skeletal muscle hypertrophy, European Journal of applied Physiology 111, 2785–2790

[6] J. A. Faulkner, S.V. Brooks and J.A Opiteck (1993), Injury to skeletal muscle fibers during contractions: conditions of occurrence and prevention, Physical Therapy, 73(12), 911–921.

[7] Frankenfield, D. (2006), Energy expenditure and protein requirements after traumatic injury, Nutrition in clinical practice, 21(5), 430-437.

[8] Kamenica, E. and Gentzkow, M. (2011), Bayesian persuasion, American Economic Review 101(6), 2590-2615.

[9] Kiely, J. (2012), Periodization paradigms in the 21st century: evidence-led or tradition-driven?. International journal of sports physiology and performance, 7(3), 242-250.

[10] Kiely, J., Pickering, C. and Halperin, I. (2019). Comment on "Biological background of block periodized endurance training: a review". Sports Medicine, 49(9), 1475-1477.

[11] Glass, S., Hatzel, B. and Albrecht, R. (2014), Kinesiology For Dummies, John Wiley & Sons.

[12] Issurin, V.B. (2010), New horizons for the methodology and physiology of training periodization,Sports Medicine, 40(3), 189-206.

[13] Issurin, V.B. (2019), Biological background of block periodized endurance training: A review. Sports Medicine 49, 31–39.

[14] Lazear, E. P. (2000), Economic imperialism, Quarterly Journal of economics 115(1), 99-146.

[15] Lieberman, D. E. (2015), Is exercise really medicine? an evolutionary perspective, Current sports medicine reports 14(4), 313–319.

[16] Sagar, S.S. and J. Stoeber (2009), Perfectionism, fear of failure, and affective responses to success and failure: The central role of fear of experiencing shame and embarrassment. Journal of Sport and Exercise Psychology, 31(5), 602-627.

[17] Steele, J. (2014), Intensity; in-ten-si-ty; noun. 1. Often used ambiguously within resistance training. 2. Is it time to drop the term altogether? British Journal of Sports Medicine, 48(22), 1586-1588.

[18] Steele, J., Fisher, J., Giessing, J., & Gentil, P. (2017), Clarity in reporting terminology and definitions of set endpoints in resistance training, Muscle & Nerve, 56(3), 368-374.

[19] Zurlo, F., Larson, K., Bogardus, C., Ravussin, E. et al. (1990), Skeletal muscle metabolism is a major determinant of resting energy expenditure, Journal of clinical investigation 86(5), 1423–1427.