When Training Falls Short: Schmeduling in Economic Decision-Making*

Juliette Sellgren†

April 28, 2025

Abstract

Although marginal analysis is a cornerstone of economic theory, individuals rarely apply it in practice. Using data from an experimental input allocation task, I show that most participants instead rely on average-based heuristics, even after repeated exposure to stable incentives. Cognitive Response Test (CRT) scores robustly predict the use of marginal analysis and higher earnings, while formal economics training does not: economics majors were no more likely to act optimally than others. These results suggest that cognitive ability is the strongest predictor of stable, payoff-maximizing behavior—and that chalk-and-talk economics instruction may fail to instill true marginal thinking.

Keywords: Marginal analysis, melioration, behavioral economics, heuristics, decision-making, teaching economics

JEL Codes: C91, D03, D83

*I'd like to thank Charles Holt, my advisor, and Kerem Cosar, the organizer of the 2025 DMP, for their help throughout the process. I'd also like to thank Jack Nicastro, Lilly Bay, Ritvik Thakur, and Shea Messman for their insightful comments.

†Juliette Sellgren, mnn6tk@virginia.edu, is a fourth year in the Department of Economics at the University of Virginia, Charlottesville 22901.

1 Introduction

Economics teaches that individuals make decisions at the margin. It also teaches the principles of self-interest, rationality, and utility maximization subject to constraints. But observed behavior reveals the limitations of each of these assumptions. Güth, Schmittberger, and Schwarze (1982) show that self-interest does not accurately predict outcomes in the ultimatum game due to preferences for fairness. Herbert Simon (1955) demonstrates the boundedness of rationality through the identification of "cognitive limits." These limitations cause individuals to mentally simplify complex situations, leading to what Simon (1956, 1979) calls "satisficing," or settling for a choice which is just "good enough." Kahneman's (2011)corroboration of Simon's findings hints at the origin of theoretically "suboptimal" behavior: the decision-making process itself.

When it comes to marginal thinking, the most well-documented instance of its misapplication is the American income tax system. Gideon (2017) found that the majority of survey respondents opt to skip questions which ask about their marginal tax rate, while they almost all respond to questions about average tax rates. When respondents do answer marginal tax rate questions, they often do so incorrectly. Other goods with non-linear price schedules—for example, cellular payment plans—exhibit similar characteristics. This occurrence is so common that economists Liebman and Zeckhauser (2004) coined the term "schmeduling" to describe the phenomenon of using average analysis in place of marginal analysis.

Theories of why individuals schmedule abound, although the precise mechanism remains unclear. Liebman and Zeckhauser identify two primary explanations: "ironing" and "spotlighting." When someone irons, their understanding of a price schedule comes from a single piece of information—the average value at their current point of consumption. They don't observe or consider the nuances of the schedule at various points. When someone spotlights, their understanding of the price schedule is only local. While they rely on the marginal incentive, they do not consider how current choices affect the schedule—and their options—later

on. Spotlighters miss the forest for the trees; Ironers may miss the trees for the forest.

In the income tax example, schmeduling leads individuals to ignore the structure of tax brackets, resulting in imperfect decisions about work, leisure, and consumption. Take Johnny, who earns $50,000 annually. He is offered a promotion requiring 10 additional hours per week for an extra $30,000 in annual compensation. His average tax rate is 20%. His marginal tax rate is 25% at $50,000, but 35% at $80,000 (assume for simplicity that the $50,001–$80,000 bracket is taxed at 35%).

If Johnny irons, he assumes he will pay 20%—his average tax rate—on all of his newly earned income. He believes his extra $30,000 will be taxed at 20%, leaving him with $24,000 in post-tax earnings. But this is incorrect: his actual marginal tax rate on that income is 35%, meaning he will only receive $19,500 after taxes. As a result, Johnny overestimates the financial benefit of the promotion. He might accept the offer under the false impression that it nets him more money than it really does. The consequence of ironing is that Johnny works more than he would if he understood the true payoff, sacrificing more leisure time, family obligations, or rest.

If Johnny spotlights, he considers only the marginal tax rate of 35% and concludes that the raise is too heavily taxed to be worthwhile. He focuses on the steepness of the next step, failing to recognize that his overall post-tax income will still increase by $19,500. Johnny might reject the promotion, not because it leaves him worse off, but because his narrow focus on the marginal rate obscures the net gain in income.

Herrnstein and Prelec (1991) offer another framework for understanding schmeduling. Their theory, called "melioration," applies to scenarios in which individuals choose between multiple options and evaluate trade-offs over time. It aligns with the economic concept of average product:

> "Instead, they will choose as if guided by the following pair of rules:
>
> 1. **Value Accounting**: Keep track of the average satisfaction (or value) received per unit invested on each alternative (where the unit of account might be a single

choice occasion, a time interval, or a money unit).

2. **Melioration**: Based on these value accounting calculations, shift behavior (choice, time, money) to alternatives that provide a higher per unit return."

Consider Johnny at a weekly town baseball game. When he gets hungry, he must choose between a hotdog and a pretzel. Each costs the same, but they differ in how much satisfaction (utils) Johnny expects to get. A hotdog gives him 5 utils per dollar today, and a pretzel gives him 10 utils per dollar today. However, melioration suggests Johnny won't decide based solely on today's numbers. Instead, he will rely on value accounting, keeping track of the average satisfaction he has historically received. Suppose hotdogs have averaged 8 utils per dollar over the past month, while pretzels averaged only 7 utils per dollar (perhaps they're too dry). Even if pretzels have a higher current return, Johnny's past experiences bias him toward the hotdog. According to the melioration rule, Johnny will continue shifting toward hotdogs until their average return drops below that of pretzels. This incremental decision-making process helps explain why individuals may persist in suboptimal or path-dependent choices, especially when their tracking is based on short-term or imperfect information.

Field experiments across disciplines have found signs of ironing, spotlighting, and melioration (Herrnstein and Vaughan, 1980; Chetty et al., 2009; Sims et al., 2013; Rees-Jones and Taubinsky, 2020). Finkelstein (2009) argues that the adoption of salience-reducing technologies like E-ZPass increases individuals' reliance on heuristics when making toll payment decisions. Sims et al. (2013) posit that in dynamic and uncertain environments, melioration may be not only prevalent but rational.

Green et al. (Forthcoming) corroborate the general finding that marginal reasoning is not pervasive under uncertainty, but their experimental design holds the environment constant across rounds. In this task, 100 participants made 10 input allocation decisions using fixed production functions. This setup allowed subjects to infer the optimal strategy through repeated feedback. Yet even in the final round, many participants continued to exhibit melioration or satisficing behavior. While Sims et al. suggest that melioration can be an

adaptive response to dynamic uncertainty, Green et al. demonstrate that it also persists in simple, stable environments—highlighting its robustness as a cognitive heuristic rather than a purely strategic choice.

This paper builds on Green et al.'s findings, which show that only about a quarter of participants reached the optimal allocation by the final round—even in the kinked treatment, where the marginal cost of misallocation is steeper. While the previous paper focused on treatment-level patterns, the present study leverages additional demographic and cognitive data collected during the same experimental sessions. In particular, we analyze Cognitive Response Test (CRT) scores for all participants, and, for the kinked treatment, data on subjects' major, academic year, and gender. This richer dataset allows us to test whether specific individual traits predict the use of marginal reasoning.

Analyzing the data with a focus on these covariates, we replicate the central finding from Green et al.: Participants generally rely on average productivity rather than marginal productivity in their decision-making. Heuristic behavior consistent with schmeduling and melioration appears in all treatments, with roughly twice as many subjects exhibiting average-based reasoning as those using marginal analysis. We revisit this result in detail, as it provides the foundation for the novel findings that follow.

The second contribution of this paper is identifying predictors of marginal thinking. CRT scores are the most reliable and consistent: Participants with higher CRT scores are significantly more likely to earn more and reach the optimal allocation. In contrast, we find no evidence that economics majors perform better—despite the field's explicit canonical emphasis on marginal reasoning. The remainder of the paper proceeds as follows: Section II outlines the experimental procedures; Section III presents the main results; Section IV concludes. Supplementary figures and replication data are included in the Appendix.

2 Experimental Procedures

As mentioned above, the experiment in Green et al. consisted of ten rounds of sequential decisions. The decisions were about where to allocate inputs between two production functions, Green Eggs and Ham, with differing productivity schedules. The subjects began the task with no practice rounds and no information about the production functions, save that allocating zero inputs yields zero outputs. Figure 1 shows the production functions as seen by the participants at the beginning of the task.

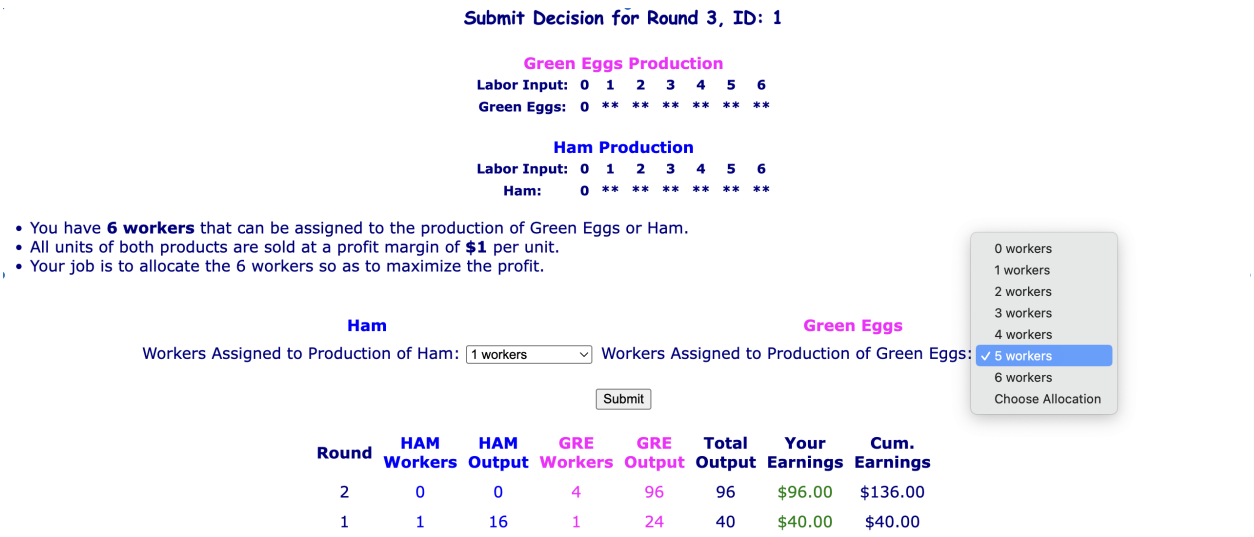
Figure 1: Subjects Are Given No Information About Production Yields

The constrained setup incentivized subjects to test allocation bundles thoughtfully because they could not simply try all bundles. Rather, subjects needed to observe the return associated with their allocation choice each round to maximize earnings.

Each round, the screen prompted subjects to select the number of inputs, "workers," they wanted to allocate to each process. In the first round, participants had two workers to allocate. In each subsequent round, two more workers became available to allocate, until round six, where the number of workers leveled off at 12. Figure 2 shows the submit page for round 3, where subjects allocate 6 workers between Green Eggs and Ham.

For every output unit produced of either type, participants earned one lab dollar, so earnings were maximized by maximizing the total units produced. As subjects allocated through the rounds, they discovered information about the shape of the production functions. This information about output production and round earnings populated on the results page after each round, in a decision history table. Figure 3 shows the results page for round 5, where ID 1 allocated 3 of their 10 available workers to Ham and 7 to Green Eggs, earning

Figure 2: Subjects Choose to Allocate Workers Between Production Functions

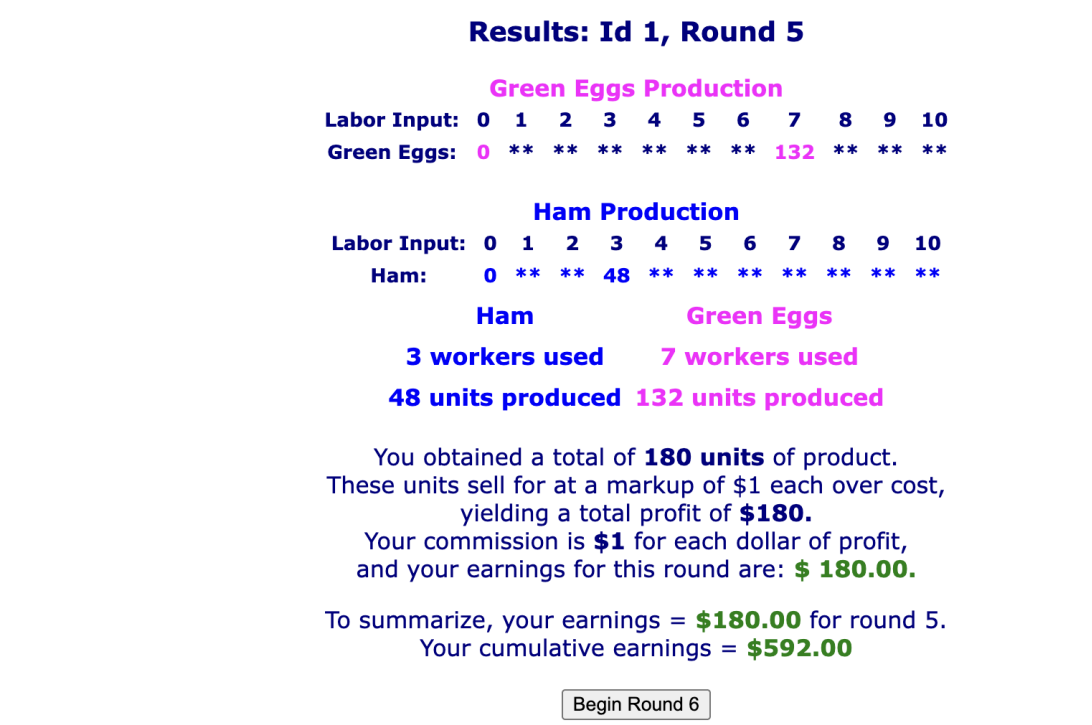


them $180 lab dollars.

There were two treatments in the experiment, the aforementioned Quadratic and Kinked treatments. In both treatments, the Ham production function was linear, with a constant average and marginal product. Every input unit allocated to Ham yielded 16 outputs. In the Quadratic treatment, the Green Eggs production function was quadratic, increasing at a decreasing rate. In the Kinked treatment, the Green Eggs production function was piecewise linear. Before the kink, every input unit allocated to Green Eggs yielded 24 outputs. After the kink, every input unit allocated to Green Eggs yielded 12 outputs. The total production yields for the production functions in each treatment can be seen in Figure 4.

Figure 5 shows the marginal and average product curves for each treatment when the subjects can allocate all 12 workers (rounds 6-10). Each marginal product line (MP) represents the returns from the nth additional worker allocated to that production process. Each average product line (AP) represents the average return of the n workers allocated to that production process. Notice that for Ham, the average and marginal product of every additional unit is equal for both treatments. The earnings-maximizing allocation occurs at the intersection of the dotted marginal product lines. The optimal allocation, in both treatments, is 4 workers to Green Eggs and 8 workers to Ham for the last 5 rounds. The

Figure 3: Subjects Are Shown Their Earnings and Decision Histories



Results: Id 1, Round 5

Green Eggs Production

Labor Input:	0	1	2	3	4	5	6	7	8	9	10
Green Eggs:	0	**	**	**	**	**	**	132	**	**	**

Ham Production

Labor Input:	0	1	2	3	4	5	6	7	8	9	10
Ham:	0	**	**	48	**	**	**	**	**	**	**

Ham	Green Eggs
3 workers used	7 workers used
48 units produced	132 units produced

You obtained a total of **180 units** of product.
These units sell for at a markup of $1 each over cost,
yielding a total profit of **$180.**
Your commission is **$1** for each dollar of profit,
and your earnings for this round are: **$ 180.00.**

To summarize, your earnings = **$180.00** for round 5.
Your cumulative earnings = **$592.00**

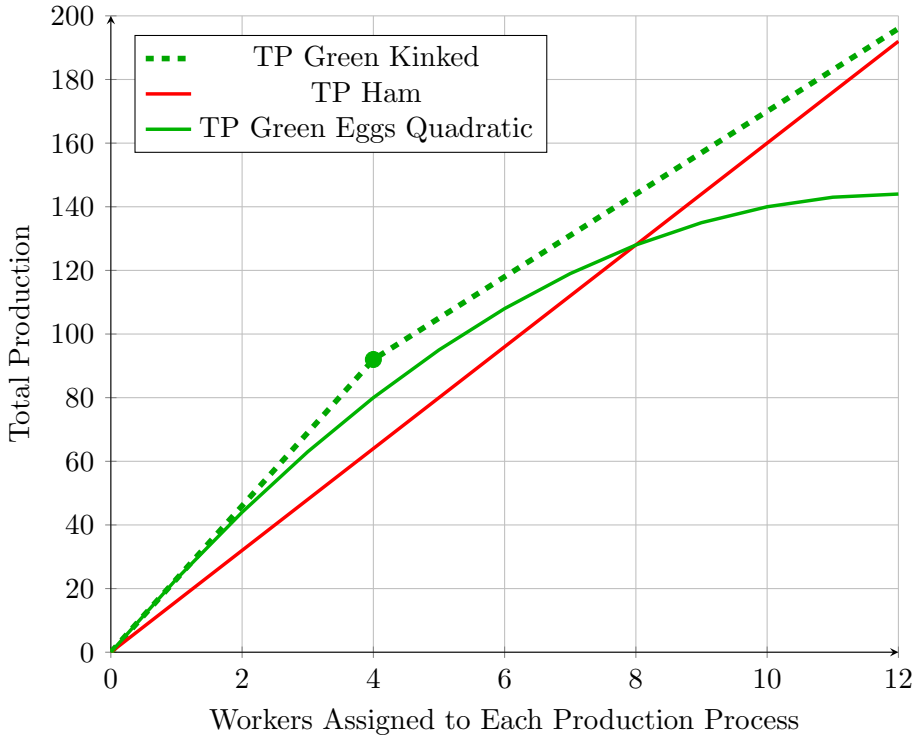
Begin Round 6

Round	HAM Workers	HAM Output	GRE Workers	GRE Output	Total Output	Your Earnings	Cum. Earnings
5	3	48	7	132	180	$180.00	$592.00
4	2	32	6	120	152	$152.00	$412.00
3	1	16	5	108	124	$124.00	$260.00
2	0	0	4	96	96	$96.00	$136.00
1	1	16	1	24	40	$40.00	$40.00

meliorating allocation, using average analysis instead of marginal, occurs at the intersection of the red dotted line (Ham average product and marginal product) and the green solid average product line. In the quadratic treatments, this equal-AP allocation. In the Kinked treatment, this allocation is 12 to Green Eggs and 0 to Ham.

The 100 subjects in this experiment were students from the University of Virginia. The experiment itself was conducted with groups of fewer than 15 at a time. After the consent process was completed, instructions were read aloud so that participants progressed through and heard the information at the same time. Including a $10 reward for showing up to the experiment, the earnings were converted to $1 for each $100 lab dollars, resulting in take-

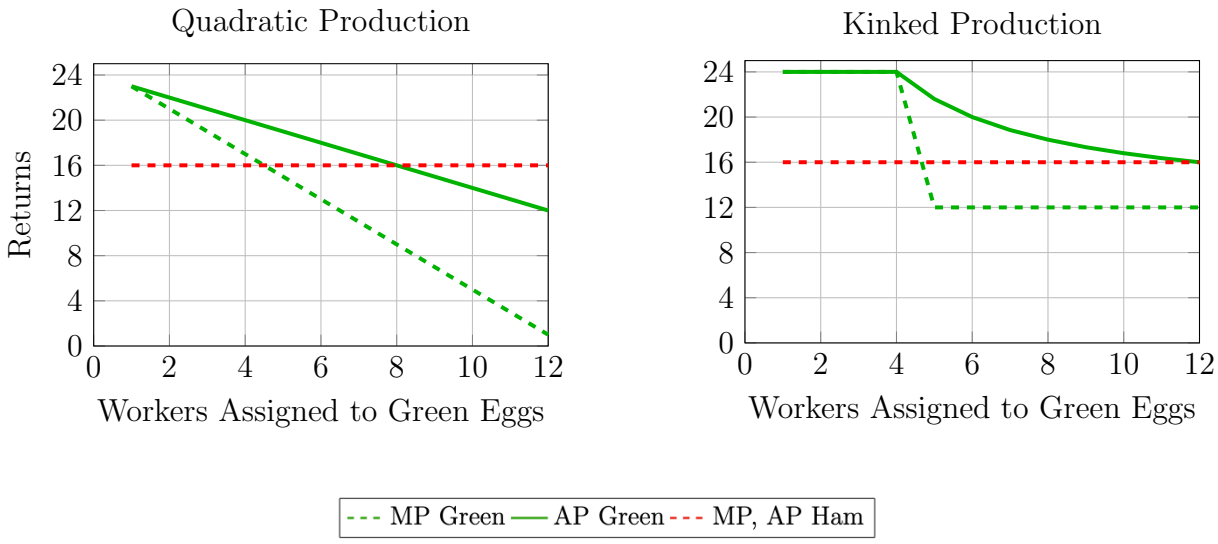
Figure 4: Total Production of Green Eggs and Ham in Both Treatments



home pay in the range of $25-30. The tasks altogether took less than an hour to complete. After completing this ten-round task, the subjects of both treatments took an unincentivized, 3-question Cognitive Response Test (CRT). [1] The possible scores ranged from 0 (answering none correctly) to 3 (answering all correctly). Of the subject pool, 50 subjects were in the Quadratic treatment group and the other 50 were in the Kinked treatment group. Those in the Kinked treatment also answered a short demographic survey after the CRT test, which pertained to major, year in college, and gender. The full instructions, CRT test questions, and demographic data responses are included in Appendix B. The allocation datasets for

[1]To assess participants' cognitive reflection, we employed a short version of the Cognitive Reflection Test (CRT), which is the best suited cognitive measure for our purposes because each question presents an intuitively appealing but incorrect response. Our version of the test included two of the original three questions developed by Frederick (2005)—the "bat and ball" and "lily pad" problems—along with an additional speed-related question adapted from Holt, Porzio, and Song (2017). The third question similarly includes a compelling but incorrect intuitive response. Including a novel question helps mitigate familiarity effects from potential previous exposure. The CRT was administered without financial incentives, streamlining the session and preserving the integrity of the questions for future experimental use. Prior work suggests that even hugely increased monetary incentives have little effect on CRT performance (Enke et al., 2023), supporting the minimal importance of incentivizing it alone.

Figure 5: Marginal and Average Product of Workers in Both Treatments



both treatments can be found on Github. [2]

3 Results

Although economic theory posits that individuals make decisions at the margin, our findings in Green et al. (Forthcoming) suggest that this principle does not consistently explain, nor does it drive, participants' allocation choices. When 12 workers are available for allocation (Rounds 6–10), the profit-maximizing choice is to send 4 workers to Green Eggs and 8 to Ham. However, only a subset (about a quarter) of participants settled this optimal allocation by the final round. We categorized subjects into three decision-making types, classifying them by their round 10 allocations. The "Marginal Thinkers," allocated 4 workers to Green Eggs and 8 workers to Ham in the final round, reaching the optimal allocation. The "Average Chasers," over-allocated more than 4 workers to Green Eggs, the production with higher average productivity. Finally, the "Overreactors," under-allocated to Green Eggs, the production function with higher average and marginal productivity.

Across treatments, subjects of each decision-making profile exhibited certain character-

[2] https://github.com/Veconlab/Marginal-Average

istic trends in allocation choices. Table 1 shows a stereotypical Marginal Thinker from the Quadratic treatment, who allocates exactly 4 workers to Green Eggs in the final round. This person notes the higher return to Green Eggs in earlier rounds and allocates all 10 available workers there by round 5. But this allocation yields total earnings of 140 lab dollars, lower earnings than the 144 earned previously. Though marginal product was not observed, the decrease in total earnings signaled to the subject to allocate more workers to Ham. In round 6, they allocate 4 workers to Ham and 8 to Green Eggs, raising total earnings to 192. Here, the average products of both functions are equal, and an Average Chaser or Meliorator would stay at this allocation. But this "slow walker" continually transferred one worker at a time to Ham in the final four rounds. These incremental changes reveal marginal product. Whether this participant noted and responded to explicit marginal product calculations or the overall increase in earnings, they leveraged marginal change to maximize earnings.

Table 1: Quadratic Treatment Example of a Marginal Thinker

Round	Ham Workers	Green Eggs Workers	Ham Produced	Green Eggs Produced	Earnings
1	1	1	16	23	39
2	2	2	32	44	76
3	2	4	32	80	112
4	4	4	64	80	144
5	0	10	0	140	140
6^{A}	4	8	64	128	192
7	**5**	**7**	**80**	**119**	**199**
8	**6**	**6**	**96**	**108**	**204**
9	**7**	**5**	**112**	**95**	**207**
$\mathbf{10^{M}}$	**8**	**4**	**128**	**80**	**208**

Notes: This person experiments with allocation decisions, transferring one worker at a time to Ham in the final 4 rounds. The superscript "A" indicates the allocation that equates average products, and the superscript "M" indicates the allocation that equates marginal products.

Table 2 shows a low-earning Average Chaser from the Kinked production treatment. In round 1, this subject allocates one worker to each production function. At this allocation, both the marginal and average products of Green Eggs surpass those of Ham. In response, the subject allocates 3 workers to Green Eggs, and only 1 to Ham. By round 3, the subject allocates all workers to the Green Eggs production function and continues to do so in all

subsequent rounds. In round 7, they experiment, allocating 1 worker to Ham once more, earning a total of 196 lab dollars in this round. Their total earnings increased from the previous allocation by 4 dollars, signaling to the subject that their earlier allocation choice was suboptimal. But in this portion of the production schedules, the average product of Green Eggs surpasses the average product of Ham. So the subject, allocating by the rule of melioration, returns to an earlier allocation of 12 workers to Green Eggs and 0 to Ham by the final round. By meliorating, the subject earned lower total and round earnings than subjects classified as Marginal Thinkers.

Table 2: Kinked Treatment Example of an Average Chaser

Round	Ham Workers	Green Eggs Workers	Ham Produced	Green Eggs Produced	Earnings
1	**1**	**1**	**16**	**24**	**40**
2	**1**	**3**	**16**	**72**	**88**
3	**0**	**6**	**0**	**120**	**120**
4	0	8	0	144	144
5	0	10	0	168	168
6	0	12	0	192	192
7	1	11	16	180	196
8–10	0	12	0	192	192

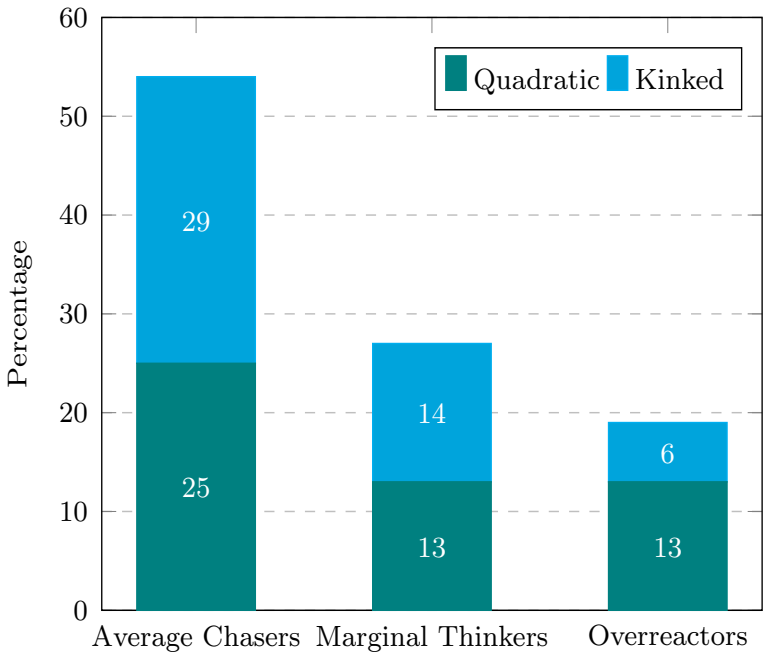
Notes: The participant in Table 2 puts "all eggs in one basket," Green Eggs, by relying a the decision-making rule of melioration. The rule was based on early analysis (Green Eggs is better than Ham) without reevaluating. This occurrence was prevalent even with the steeper marginal differences in returns associated with the kinked treatment.

Most, but not all, subjects adhere to the rules of thumb associated with their classification. But some of the Marginal Thinkers make quicker progress, allocating by increments of 2 workers at a time instead of 1. And others experiment with melioration or allocation at the extremes before settling on marginal analysis. Similarly, Average Chasers themselves can dabble in marginal analysis before being wooed by the short-term appeal of average analysis. Although there is some variation in round-by-round decision-making within each classification, a striking trend emerged overall. This result is our primary finding in Green et al. Forthcoming:

Result 1: Most individuals consistently use average analysis instead of marginal analysis.

Average Chasers comprise the majority of the subject pool across both treatments. Even in the kinked treatment, where increased marginal cost further incentivizes against analysis by average product. There were twice as many subjects classified as Average Chasers than there were Marginal Thinkers. Figure 6 below shows, in percentage terms, the distribution of participants by classification. Across both treatments, there were 54% Average Chasers, 27% Marginal Thinkers, and 19% Overreactors.

Figure 6: Average Chasing is Twice as Prominent as Marginal Thinking

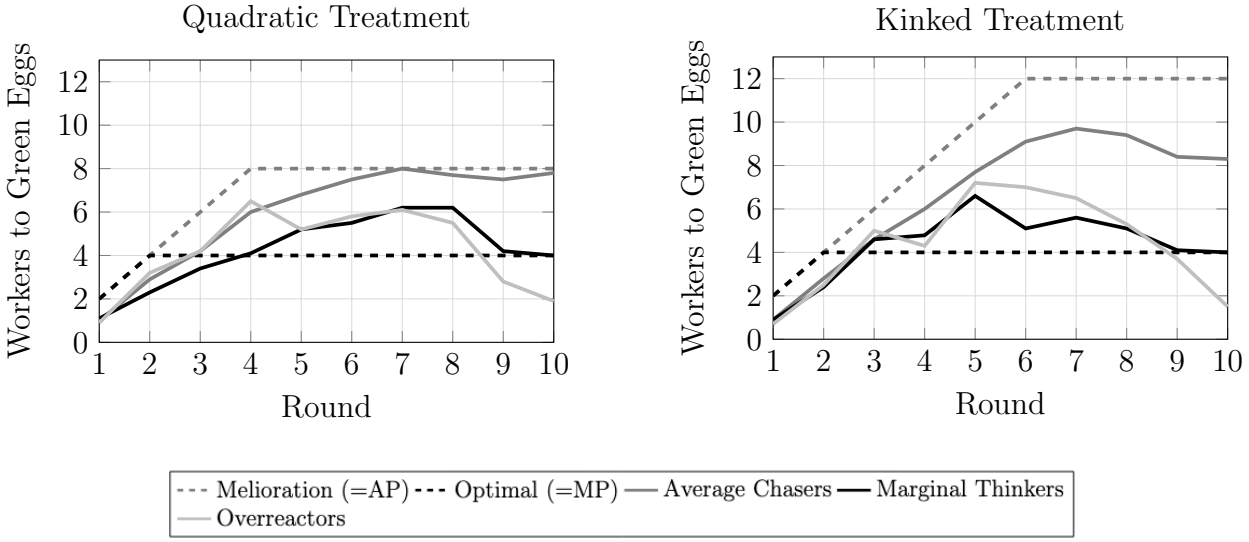


Notes: Each bar represents one of the following decision-making types observed in both treatments: 54% are Average Chasers, 27% are Marginal Thinkers, and 19% are Overreactors. Across all treatments, the percentage of average chasers is nearly double that of marginal thinkers.

Figure 7 shows the average number of workers allocated to Green Eggs across rounds in both treatments. The dashed, dark gray line shows the optimal allocation in each round, if the decision-making rule is marginal analysis. The dashed, light gray line shows the melioration allocation in each round, if the decision-making rule is average analysis. Notably, Marginal Thinkers (solid, dark gray line) tend to converge toward the marginal product line by the final rounds, while Average Chasers (solid, medium gray line) do not, hugging the

average product line. The Overreactors (solid, light gray line) appear to fluctuate between the two other groups until the end, when they consistently allocate fewer workers to Green Eggs than equating averages or marginals would suggest.

Figure 7: Average Allocation of Workers to Green Eggs by Round



Note: Each line represents the average number of workers allocated to Green Eggs per round by a behavioral type. The horizontal dashed line at 4 represents the marginally optimal (profit-maximizing) allocation; the upward-sloping dashed line shows the average product-maximizing level.

Due to stark heterogeneity in decision-making rules between groups, we estimated the following linear regression using ordinary least squares (OLS), to distinguish between reliance on melioration and marginal analysis for individual subjects. The model estimates changes in the proportion of workers assigned to Green Eggs, ΔGE_{it}, between rounds. The two key explanatory variables are: the difference in average product between Green Eggs and Ham, Equation (1), and an approximation of the difference marginal product of Green Eggs, Equation (2). They were calculated for each round as follows:

$$AP_{\text{difference}} = \text{AP}_{\text{Green},\,t-1} - \text{AP}_{\text{Ham},\,t-1} \tag{1}$$

$$\Delta GreenMP_{\text{approximation}} = \frac{TP_{\text{Green},\,t-1} - TP_{\text{Green},\,t-2}}{Workers_{\text{Green},\,t-1} - Workers_{\text{Green},\,t-2}} \tag{2}$$

Ham is not included in the regression because its productivity remains constant at 16 throughout all rounds. This regression was estimated using data from rounds 3 through 10, for all treatments and each separate treatment. The regression specification is presented in Equation (3) and the full results in Table 3.

$$\Delta GE_{it} = \beta_0 + \beta_1 AP_{\text{difference}} + \beta_2 \Delta GreenMP_{\text{approximation}} + \epsilon_{it} \tag{3}$$

Table 3: OLS Regression for Change in Share of Workers Assigned to Green Eggs

	Pooled	**Quadratic**	**Kinked**
Constant	-0.115***	-0.111***	-0.162***
	(0.020)	(0.032)	(0.028)
AP difference:	0.030***	0.036***	0.035***
	(0.004)	(0.009)	(0.006)
MP_{Green} *approximation:*	0.002	-0.001	0.002
	(0.003)	(0.005)	(0.003)
Observations	890	449	441
Number of Individuals	100	50	50
Overall R-squared	0.154	0.144	0.186

Notes: If there was no change in Green Eggs allocations in the prior period, the approximated difference in marginal products was based on the most recently observed change. Observations are omitted if, for example, subjects do not have any information on average product of Green Eggs, since they had yet to allocate any workers to its production. Standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The results that follow extend beyond those reported in Green et al. (Forthcoming) by incorporating analyses of demographic characteristics and CRT scores collected in post-experiment surveys. These additional data allow for an investigation into individual-level predictors of decision-making strategies within the same experimental framework.

Result 2: A high CRT score is the only significant predictor of marginal analysis and optimal performance.

While Result 1 pools observations across both the Quadratic and Kinked treatments, the following analyses draw on additional demographic data collected only in the Kinked treatment. Specifically, data on participants' majors, years in college, and gender were collected in a post-experiment survey administered only to subjects in the Kinked group. However, Cognitive Reflection Test (CRT) scores were collected for participants in both treatments, and are analyzed accordingly where relevant. Given the consistency of behavioral patterns across treatments in the main result—namely, the comparable prevalence of average-based reasoning—it is reasonable to interpret the demographic findings from the Kinked treatment as indicative of broader patterns across both groups.

To evaluate the relationship between cognitive reflection and decision-making in the Kinked treatment, I implement a series of nonparametric permutation tests. This statistical procedure compares the observed difference in means between groups to a reference distribution generated by randomly reassigning group labels—here, behavioral classification. The p-value reflects the proportion of simulated reallocations that produce a difference in means as large or larger than the observed one. Where relevant, treatment labels are permuted within strata defined by fixed values of CRT score, gender, or major. This stratified approach ensures that any observed group differences are not attributable to those covariates.

In the Kinked treatment, results from the permutation tests reveal a strong association between CRT scores and behavioral type. Marginal thinkers have a significantly higher average CRT score (2.64 correct answers) than both average chasers (1.53) and overreactors

(1.83). The difference between Marginal Thinkers and Average Chasers is highly significant ($p = 0.0044$). The comparison of Marginal Thinkers with Overreactors is also statistically significant ($p = 0.046$). Table 4 presents further evidence using only Marginal Thinkers and Average Chasers in the Kinked treatment. Overreactors were ommitted from this analysis to isolate the effect of melioration associated with Average-Chasing allocations. Those with higher CRT scores (2 or 3) assign more workers to the optimal marginal allocation (4), consistently across demographic subgroups. For instance, within the row of female non-econ majors, a participant with a low CRT score might assign 12 workers to Green Eggs in the final round—triple the marginal optimum—while a high-CRT participant might assign 4, indicating a sharper sensitivity to marginal costs.

Table 4: High CRT Scores Predict Optimal Final Round Allocation Decisions

	Green Workers Assigned for Low CRT (0 or 1)	**Green Workers Assigned for High CRT (2 or 3)**
5 Male Econ/Business Majors	7	4, 4, 4, 12
12 Male Other Majors	7, 8, 9	4, 4, 4, 4, 4, 6, 8, 8, 8
12 Female Econ/Business Majors	4, 5, 8, 9, 9	4, 4, 6, 6, 7, 10, 12
15 Female Other Majors	6, 7, 8, 8, 9, 12, 12, 12	4, 4, 4, 5, 6, 6, 12
44 Avg. Chasers or Marginal Thinkers:	**Avg. = 8.24, Median = 8**	**Avg. = 6.07, Median = 5**

Notes: The table lists final-period assignments for 44 subjects by CRT score (low or high). The Overreactors in each category have been omitted to isolate the effect of average analysis (allocation of workers ¿ 4) relative to marginal analysis (allocation of workers = 4).

The correlations between CRT scores and behavioral types are suggestive: the intuitive thinking that leads to incorrect CRT responses may also drive biases like neglecting marginal comparisons. Average-based reasoning relies on simple current-period ratios, while marginal analysis requires more complex, less intuitive comparisons across time. As Frederick (2005) notes, CRT scores often correlate so strongly with behavioral biases that it is difficult to disentangle cognitive reflection from bounded rationality. Demographic patterns further complicate this, with females tending to have lower CRT scores than males (Frederick, 2005; Holt et al., 2017).

Table 4 provides a stratified view of CRT effects under controlled demographic conditions,

while Table 5 summarizes these results and compares CRT to gender and major. While it originally seems that major, gender, and CRT are correlated with final Green Eggs allocation, the relationship with CRT is the only significant one. High-CRT individuals assign just over two fewer workers to Green Eggs than their low-CRT counterparts, suggesting sharper sensitivity to marginal returns. Gender differences, by contrast, are smaller and not significant once CRT is held constant. For example, a male and female with the same CRT score and major are equally likely to reach the optimal allocation—indicating that gender alone does not reliably predict marginal thinking in this setting.

Table 5: CRT Significantly Predicts Allocation Accuracy; Gender and Major Do Not

Indicator	**Mean Group 1**	**Mean Group 2**	**Difference**	**P-Value**
CRT (High vs. Low)	6.07	8.24	-2.17	0.018
Gender (Male vs. Female)	6.18	7.37	-1.19	0.24
Major (Economics vs. Other)	6.76	7.37	-0.61	0.42

Notes: Results from permutation tests on 44 subjects in the Kinked Treatment using clustered stratification by CRT score, gender, and major. 6 Overreactors are omitted to isolate the extent of average analysis. P-values are for a stratified permutation test (Holt and Sullivan, 2023), which randomly permutes the assignment vector within strata (50,000 random reassignments).

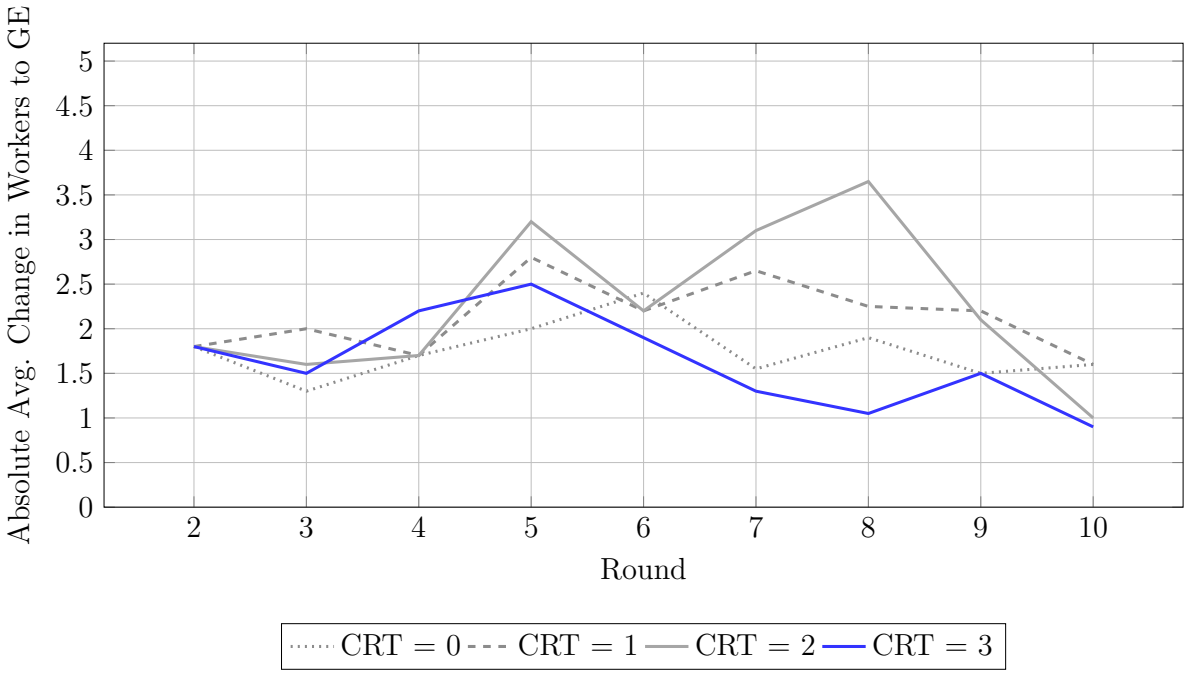
These findings extend beyond the final round. Appendix A reports the results of a supplementary permutation test, which examines the relationship between CRT scores and average total earnings across all rounds. The result reinforces and extends the earlier analysis: participants with higher CRT scores earn significantly more over the course of the experiment than those with lower CRT scores ($p = 0.0089$). This result suggests that the advantages of reflective reasoning are not confined to reaching the correct allocation by the final round, but persist throughout the rounds. Higher CRT individuals not only tend to converge toward the optimal allocation but do so with greater consistency, avoiding costly fluctuations across rounds. In this sense, CRT predicts both accuracy and stability—both critical components of long-run payoff maximization in dynamic environments.

To further assess the behavioral mechanism behind CRT performance, I examine alloca-

tion stability across the experiment, focusing on both the Quadratic and Kinked Treatments. Figure 9 plots the average absolute change in the number of workers allocated to Green Eggs between rounds, grouped by CRT score. Stability is captured by smaller within-subject fluctuations in allocation from one round to the next.

The figure reveals a clear pattern: subjects with the highest CRT score (3) display lower average changes in allocation. In the final three rounds—when the production function and strategies are most transparent—CRT 3 participants exhibit especially low volatility. For example, in round 8, the average CRT 3 subject changed their Green Eggs allocation by 1 worker, compared to 1.8 for CRT 0, 2.3 for CRT 1, and 3.6 for CRT 2. This lower volatility suggests that CRT 3 participants either begin closer to the optimal allocation or adjust more cautiously and precisely in response to marginal signals. Overall, CRT 3 participants exhibit smaller average fluctuations in worker allocation, indicating increased marginal thinking.

Figure 9: Highest CRT score Exhibits Lower Absolute Average Change in Allocation of Workers to Green Eggs Across Rounds



Notes: This figure plots the average absolute change in Green Eggs worker allocations between consecutive rounds, by CRT score. Participants with a CRT score of 3 consistently exhibit less fluctuation in their allocations, suggesting more deliberate, marginal decision-making.

To further test whether CRT scores predict stability in marginally-informed decision-making, I estimate a linear model of changes in Green Eggs allocation across rounds. The model examines whether higher CRT scores are associated with more deliberate, incremental adjustments—distinct from both erratic overreaction and passivity—after controlling for signals from average and marginal product measures. The regression is run separately for the Quadratic and Kinked treatments using data from rounds 5 through 10, where average and marginal productivity diverge most clearly. Because Ham's productivity remains constant at 16 in all rounds, only Green Eggs variation is included in the analysis.

CRT scores are coded categorically, with CRT = 0 as the reference group. Each CRT coefficient therefore represents the average difference in allocation behavior between individuals with that CRT score and those with no correct answers. In both treatments, CRT = 2 is associated with significantly more stable behavior—roughly a 9 to 11 percentage point smaller change in allocation relative to CRT = 0. CRT = 3 subjects also show significantly smaller adjustments in the Kinked treatment, consistent with more restrained and consistent learning. These effects reinforce the emerging pattern: higher CRT scores predict not only better alignment with marginal incentives, but also smoother behavioral trajectories over time. The regression specification is provided in Equation (4), and full results are shown in Table 6.

$$\Delta GE_{it} = \beta_0 + \beta_1 AP_{\text{difference}} + \beta_2 \Delta GreenMP_{\text{approximation}} + \sum_{k=1}^{3} \beta_k \cdot CRT_i^k + \epsilon_{it} \qquad (4)$$

Table 6: OLS Regression for Change in Share of Workers Assigned to Green Eggs with CRT Score Levels

	Pooled	**Quadratic**	**Kinked**
Constant	-0.025	-0.003	-0.062*
	(0.029)	(0.054)	(0.034)
$AP_{Green(t-1)} - AP_{Ham(t-1)}$	0.027***	0.029***	0.035***
	(0.005)	(0.010)	(0.006)
MP_{Green} *approximation:*	0.006*	0.005	0.009**
$\frac{TP_{Green(t-1)} - TP_{Green(t-2)}}{Workers_{Green(t-1)} - Workers_{Green(t-2)}}$	(0.004)	(0.007)	(0.004)
CRT = 1	-0.063**	-0.094**	-0.035
	(0.028)	(0.047)	(0.037)
CRT = 2	-0.096***	-0.113***	-0.112**
	(0.027)	(0.039)	(0.052)
CRT = 3	-0.057**	-0.042	-0.087**
	(0.028)	(0.043)	(0.035)
Observations	599	300	299
Number of Individuals	100	50	50
Overall R-squared	0.139	0.129	0.190

Notes: If there was no change in Green Eggs allocations in the prior period, we approximate the difference in marginal products based on the most recently observed change in a prior period. A few observations are omitted if, for example, subjects do not have any information on average product of Green Eggs, since they had yet to allocate any workers to its production. Regression output for CRT variables are relative to a CRT baseline of 0. Standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Result 3: The economics major does not teach individuals to use marginal analysis in decision-making.

An analysis of the demographic major data reveals that economics and business majors—despite formal training in marginal analysis—do not significantly outperform their peers. Because demographic information was collected only for the Kinked treatment, the following analysis is limited to that condition. Of the 22 economics majors in the Kinked treatment, over half are classified as Average Chasers. On average, economics majors allocate only 0.6 fewer workers to Green Eggs in the final round compared to non-economics majors, and they earn just 2.47% more in total earnings. Both differences are statistically insignificant based on two-tailed permutation tests that control for gender and CRT score. These results are summarized in Table 7, with full test statistics in Appendix A. The same stratified permutation procedure described above is used here to hold gender and CRT constant.

Table 7: Effect of Economics/Business Major on Key Decision Outcomes for the Kinked Production Treatment

Dependent Variable	**Mean$_{\text{Econ}}$**	**Mean$_{\text{Other}}$**	**Difference**	**P-Value**
Final GE Allocation	6.76	7.37	-0.61	0.42
Final Earnings	1619.06	1616.59	2.47	0.48

Notes: Results from permutation tests using clustered stratification of gender and CRT score. P-values are for a stratified permutation test (Holt and Sullivan, 2023), which randomly permutes the assignment vector within strata (50,000 random reassignments). Full tables with individual-level data and subgroup averages can be found in Appendix A.

This finding is particularly revealing when considered alongside the earlier results on CRT. As shown in Tables 4 and 5, CRT scores are the only robust predictor of both final allocations and overall earnings. In contrast, economics training does not meaningfully improve decision-making at the margin. Even those with economics backgrounds appear to rely on intuitive, average-based heuristics—highlighting a disconnect between classroom theory and applied reasoning under pressure. In short, while economics majors may be

taught marginal analysis, they do not seem to apply it in practice.

A natural concern with interpreting major-based effects is that self-declared economics or business majors may vary widely in their exposure to formal training. Some participants may have been just weeks into their first economics course, while others could be near degree completion. To explore this variation, Table 8 disaggregates final-period Green Eggs assignments for 22 economics and business majors in the Kinked treatment by stage in major (Years 1–2 vs. Years 3–4), holding CRT score and gender constant across four comparison groups. Overreactors are excluded to isolate meaningful differences between Average Chasers and Marginal Thinkers.

A clear pattern fails to emerge: participants further along in their economics coursework do not assign workers more optimally than those earlier in the major. In fact, their average and median allocations are slightly higher, suggesting a greater tendency to chase average product. A nonparametric permutation test comparing the two groups yields no significant difference in final allocation decisions ($p = 1.0$).

Table 8: Advanced Economics Majors Do Not Allocate Workers More Optimally Than Early Majors

	Green Workers Assigned for Years 1 and 2	**Green Workers Assigned for Years 3, 4, and 5**
6 High CRT Male	4	8, 8, 8, 4, 4
2 Low CRT Male	4	9
8 High CRT Female	6, 6, 4, 7, 10	4, 6, 4
6 Low CRT Female	9, 4, 8	8, 12, 12
24 Avg. Chasers or Marginal Thinkers:	**Avg. = 6.08, Median = 6**	**Avg. = 7.67, Median = 8**

Notes: The table lists final period assignments for 22 subjects. The Overreactors in each category have been omitted so that the remaining numbers measure a tendency to over-allocate workers to Green Eggs relative to the optimal level of 4. There is no significant difference in allocation tendencies between majors in their first and second halves of college.

Taken together, these results reinforce the earlier finding that formal economics education does not predict marginal reasoning in practice. Even among upper-level economics students—who have presumably completed coursework in microeconomic theory—final allo-

cations remain indistinguishable from those of introductory students. Despite being taught to think at the margin, students appear not to internalize or apply marginal analysis in real-time decision-making tasks.

4 Conclusion

Most individuals do not think at the margin, even with repeated exposure to a simple, stable decision-making environment. In an experiment with just two inputs to choose between and ten total decisions, the majority of participants consistently relied on heuristics rather than marginal analysis. The most prevalent of these, melioration and ironing, depend on observing average returns or local cues, signals more readily available than marginal returns. Marginal returns, without an intentional increase in allocation by one or two units, can be difficult if not impossible to observe. Average and absolute returns, on the other hand, can be easily calculated every round no matter the allocation choices made.

Decision-making by these simplified rules of thumb, instead of marginal analysis, dominated across treatments and most demographics. Even when the marginal cost of misallocation increased sharply, as in the Kinked treatment, few participants adjusted their behavior accordingly.

Although gender and major appear to be correlated with marginal analysis, they are not. The strongest and most consistent predictor of success was not formal training in economics or business, nor demographic traits like gender, but rather cognitive reflection. High CRT scores predicted both greater total earnings and a higher likelihood of reaching the optimal allocation by the final round. However, these individuals were not necessarily marginal thinkers in the formal sense. Their incremental approach to updating decisions shielded them from the high costs associated with overreaction and volatility, even if swayed by high average returns. By contrast, lower CRT individuals were more prone to large swings in allocation, yielding lower total earnings and allocations chasing higher average returns.

Surprisingly-and ironically for a subject grounded in the principle of marginal analysis-the economics major had no significant effect on performance. Participants who majored in economics were no more likely to use marginal analysis or find the optimal allocation than those who majored in unrelated fields. In the kinked treatment, one explicitly designed to increase marginal incentives, even those trained to notice the margin didn't. While one might speculate that this reflects weaknesses specific to the economics program at the University of Virginia, the result is more plausibly interpreted as a general failure of non-experiential instruction: lectures do not translate into practice. If marginal analysis is to matter, it must be trained by experience.

That failure may also stem from a deeper issue: Our cognitive machinery evolved to deal with uncertain, reactive environments where average performance often dominates marginal trade-offs. As Barkow, Cosmides, and Tooby (1992) argue, humans are not adapted for abstract optimization, but for satisficing under pressure. In uncertain environments, like hunting or foraging, optimizing on the margin may be both infeasible and maladaptive. From this perspective, the persistence of average-based decision rules is not a flaw but a feature of human cognition.

And yet, our environments have changed. We now face decisions about retirement, education, and taxation that require planning over long horizons, and where marginal trade-offs accumulate into meaningful differences in income, leisure, and well-being. Misunderstanding tax brackets, failing to save for retirement, and misallocating productive inputs are all modern analogs of ancient missteps. In this sense, schmeduling is a vestige of our earlier and less complex lives.

While heuristics offer speed and simplicity, they are not costless. As Benjamin Franklin famously noted, "nothing is certain except death and taxes," and the non-linear structure of the U.S. tax system may be the clearest, most consequential instance of how misunderstanding the margin results in real, measurable losses. If we are serious about promoting economic agency, we must be equally serious about teaching marginal analysis.

References

Barkow, J. H., Cosmides, L., & Tooby, J. (Eds.). (1992). *The adapted mind: Evolutionary psychology and the generation of culture*. Oxford University Press.

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, *99*(4), 1145–1177. https://doi.org/10.1257/aer.99.4.1145

Finkelstein, A. (2009). E-ztax: Tax salience and tax rates. *The Quarterly Journal of Economics*, *124*(3), 969–1010. https://doi.org/10.1162/qjec.2009.124.3.969

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42.

Gideon, M. (2017). Do individuals perceive income tax rates correctly? *Public Finance Review*, *45*(1), 97–117. https://doi.org/10.1177/1091142115602062

Green, M., Holt, C. A., Sellgren, J., & Sprott, E. (Forthcoming). Missing the margin: Behavioral biases in incremental decision making [Forthcoming]. *Review of Behavioral Economics*.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, *3*(4), 367–388. https://doi.org/10.1016/0167-2681(82)90011-7

Herrnstein, R. J., & Prelec, D. (1991). Melioration: A theory of distributed choice. *The Journal of Economic Perspectives*, *5*(3), 137–156. https://doi.org/10.1257/jep.5.3.137

Herrnstein, R. J., & Vaughan, W. J. (1980). Melioration and behavioral allocation. In J. E. R. Staddon (Ed.), *Limits to action: The allocation of individual behavior* (pp. 143–176). Academic Press. https://www.researchgate.net/publication/243771444_Melioration_and_behavioral_allocation

Holt, C. A., Porzio, M., & Song, M. (2017). Price bubbles, gender, and expectations in experimental asset markets. *European Economic Review*, *100*, 72–94.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus; Giroux.

Liebman, J. B., & Zeckhauser, R. J. (2004). *Schmeduling* [NBER Working Paper], Harvard University.

Rees-Jones, A., & Taubinsky, D. (2020). Measuring 'schmeduling'. *The Review of Economic Studies*, *87*(5), 2399–2438. https://doi.org/10.1093/restud/rdaa007

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99–118. https://doi.org/10.2307/1884852

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138. https://doi.org/10.1037/h0042769

Simon, H. A. (1979). Rational decision making in business organizations. *American Economic Review*, *69*(4), 493–513.

Sims, C. R., Neth, H., Jacobs, R. A., & Gray, W. D. (2013). Melioration as rational choice: Sequential decision making in uncertain environments. *Psychological Review*, *120*(1), 139–154. https://doi.org/10.1037/a0030850

A Appendix: Additional Figures and Tables

Table 9: Undergraduate Training in Economics Has No Effect on Final Round Allocation Choices in the Kinked Treatment

	Green Workers Assigned for Economics/Business Majors	**Green Workers Assigned for Other Majors**
13 High CRT Males	4,4,4,12	4,4,4,4,4,6,8,8,8
4 Low CRT Males	7	7, 8, 9
14 High CRT Females	4, 4, 6, 6, 7, 10, 12	4, 4, 4, 5, 6, 6, 12
13 Low CRT Females	4, 5, 8, 9, 9	6, 7, 8, 8, 9, 12, 12, 12
44 Avg. Chasers or Marginal Thinkers:	**Avg. = 6.76, Median = 6**	**Avg. = 7.37, Median = 7**

Note: The table lists final period assignments for 44 subjects in the Kinked treatment. The Overreactors have been omitted so that the remaining observations measure a tendency to over-allocate workers to Green Eggs relative to the optimal level of 4.

Table 10: Permutation Test Results: Effect of Major on Final GE Allocation in the Kinked Treatment

	Treatment 1 (Econ/Business)	**Treatment 2 (Other Majors)**
Number of Observations	17	27
Average Final GE Allocation	6.76	7.00
Median Final GE Allocation	6	7
Difference (Treatment 1 - Treatment 2)	6.76 - 7.00 = -0.235	
Permutation Test Results		
P-Value (2-tailed)	0.8124	
P-Value (1-tailed)	0.4123	

Notes: Based on 50,000 simulations. I used a stratified permutation test (Holt and Sullivan, 2023), which randomly permutes assignment vectors within strata, in this case gender and CRT score (see table 9). There is an insignificant difference between economics and business majors and other training in predicting final round allocation.

Table 11: Undergraduate Training in Economics Has No Effect on Final Round Earnings in the Kinked Treatment

	Total Earnings for Economics/Business Majors	**Total Earnings for Other Majors**
13 High CRT Males	1700, 1676, 1556, 1696	1696, 1552, 1568, 1588, 1660, 1656, 1636, 1716, 1620
4 Low CRT Male	1604	1540, 1600, 1580
14 High CRT Females	1656, 1640, 1644, 1524, 1592, 1632, 1580	1548, 1704, 1604, 1668, 1616, 1676, 1708
13 Low CRT Females	1624, 1608, 1588, 1608, 1596	1640, 1532, 1604, 1656, 1616, 1552, 1584, 1528
44 Avg. Chasers or Marginal Thinkers:	**Avg. = 1617.1, Median = 1624**	**Avg. = 1614.1, Median = 1616**

Note: The table lists total earnings for 44 subjects in the Kinked treatment. The Overreactors have been omitted so that the remaining observations measure a tendency to over-allocate workers to Green Eggs relative to the optimal level of 4.

Table 12: Permutation Test Results: Effect of Major on Final Earnings in the Kinked Treatment

	Treatment 1 (Economics)	**Treatment 2 (Other Majors)**
Number of Observations	17	27
Average Final Earnings	1619.06	1616.59
Median Final Earnings	1608	1616
Difference (Treatment 1 - Treatment 2)	1619.06 - 1616.59 = 2.466	
Permutation Test Results		
P-Value (2-tailed)	0.8842	
P-Value (1-tailed)	0.4805	

Notes: Based on 50,000 simulations. I used a stratified permutation test (Holt and Sullivan, 2023), which randomly permutes assignment vectors within strata, in this case gender and CRT score (see Table 11). There is an insignificant difference between economics majors and other majors in predicting total earnings.

Table 13: Gender has no Effect on Final Round Green Eggs Allocation in the Kinked Treatment

	Green Workers Assigned for Males	**Green Workers Assigned for Females**
17 High CRT Econ Majors	4, 4, 4, 4, 4, 4, 4, 6, 8	6, 6, 4, 4, 12, 4, 7, 10
6 Low CRT Econ Majors	7	5, 9, 4, 8, 9
16 High CRT Other Majors	8, 8, 8, 8, 4, 4, 6, 6	12, 4, 6, 4, 6, 5, 4, 4
18 Low CRT Other Majors	8, 7, 9, 6, 4, 4, 4, 4, 4, 6	8, 12, 7, 12, 8, 9, 12, 6
44 Avg. Chasers or Marginal Thinkers:	**Avg. = 6.18, Median = 6**	**Avg. = 7.37, Median = 7**

Notes: The table lists final round allocation of Green Eggs workers for 44 subjects in the Kinked treatment. The Overreactors have been omitted so that the remaining observations measure a tendency to over-allocate workers to Green Eggs relative to the optimal level of 4.

Table 14: Permutation Test Results: Effect of Gender on Final Round Earnings in the Kinked Treatment

	Treatment 1 (Male)	**Treatment 2 (Female)**
Number of Observations	17	27
Average Final GE Allocation	6.18	7.37
Median Final GE Allocation	6	7
Difference (Treatment 1 - Treatment 2)	6.18 - 7.37 = -1.194	
Permutation Test Results		
P-Value (2-tailed)	0.2370	
P-Value (1-tailed)	0.2269	

Notes: Based on 50,000 simulations. I used a stratified permutation test (Holt and Sullivan, 2023), which randomly permutes assignment vectors within strata, in this case CRT score and major (see Table 13). There is no statistically significant difference between male and female participants in final Green Eggs allocation.

Table 15: Permutation Test Results: Effect of CRT Score on Final Earnings in the Both Treatments

	Treatment 1 (High CRT)	**Treatment 2 (Low CRT)**
Number of Observations	27	17
Average Final Earnings	1633.78	1591.76
Median Final Earnings	1640	1600
Difference (Treatment 1 - Treatment 2)	1633.78 - 1591.76 = 42.014	
Permutation Test Results		
P-Value (2-tailed)	0.0089	
P-Value (1-tailed)	0.0075	

Notes: Based on 50,000 simulations. I used a stratified permutation test (Holt and Sullivan, 2023), which randomly permutes assignment vectors within strata, in this case gender and major (see Table 4). High CRT participants earn significantly more than low CRT participants in across all treatments.

Table 16: Permutation Test Results: Effect of Year on Final Green Eggs Allocation Among Economics Majors in the Kinked Treatment

	Treatment 1 (1st/2nd Year Econ)	**Treatment 2 (3rd/4th Year Econ)**
Number of Observations	21	23
Average Final GE Allocation	6.90	6.91
Median Final GE Allocation	7	6
Difference (Treatment 1 - Treatment 2)	6.90 - 6.91 = -0.008	
Permutation Test Results		
P-Value (2-tailed)	1.0000	
P-Value (1-tailed)	0.5847	

Notes: Based on 50,000 simulations. I used a stratified permutation test (Holt and Sullivan, 2023), which randomly permutes assignment vectors within strata, in this case gender and CRT score (see Table 8). There is no significant difference between early- and late-stage economics majors in final allocation behavior for participants in the Kinked Treatment.

Table 17: OLS Regression for Change in Share of Workers Assigned to Green Eggs with CRT levels in all Rounds

	Pooled	**Quadratic**	**Kinked**
Constant	-0.077***	-0.064*	-0.135***
	(0.027)	(0.048)	(0.033)
$AP_{Green(t-1)} - AP_{Ham(t-1)}$	0.030***	0.035***	0.038***
	(0.005)	(0.010)	(0.006)
MP_{Green} *approximation:*	0.002	0	0.001
$\frac{TP_{Green(t-1)} - TP_{Green(t-2)}}{Workers_{Green(t-1)} - Workers_{Green(t-2)}}$	(0.003)	(.005)	(0.003)
CRT = 1	-0.035*	-0.061**	-0.010
	(0.021)	(0.035)	(0.025)
CRT = 2	-0.061***	-0.068**	-0.082*
	(0.023)	(0.034)	(0.036)
CRT = 3	-0.043*	-0.037	-0.059*
	(0.022)	(0.036)	(0.025)
Observations	890	449	441
Number of Individuals	100	50	50
Overall R-squared	0.159	0.150	0.198

Notes: If there was no change in Green Eggs allocations in the prior period, we approximate the difference in marginal products based on the most recently observed change in a prior period. A few observations are omitted if, for example, subjects do not have any information on average product of Green Eggs, since they had yet to allocate any workers to its production. Regression output for CRT variables are relative to a CRT baseline of 0. Standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B Appendix: Data

Table B.1: Subject-Level Data for Quadratic Treatment (CRT Scores)

Session	ID	CRT
1	1	3
1	2	2
1	3	3
1	4	2
1	5	0
1	6	1
1	7	3
1	8	3
1	9	1
1	10	3
1	11	2
1	12	0
2	1	3
2	2	3
2	3	2
2	4	2
2	5	1
2	6	3
2	7	2
2	8	1
2	9	2
2	10	3

Note: Subject-level allocation data is available on Github:

https://github.com/Veconlab/Marginal-Average.

Session	ID	CRT
2	11	1
2	12	0
3	1	3
3	2	2
3	3	3
3	4	0
3	5	3
3	6	2
3	7	3
3	9	1
3	10	2
3	11	3
3	12	2
3	13	0
3	14	1
3	15	3
4	1	3
4	2	3
4	3	0
4	4	2
4	5	2
4	6	2
4	7	2
4	8	3
4	9	3
4	10	3
4	11	3
4	12	0

Table B.2: Subject-Level Data for Kinked Treatment (Demographics and CRT Scores)

Session	ID	GE Effort	CRT	Gender	Year	Bus	Econ
6	8	0	2	0	2	0	0
10	10	0	1	1	1	1	0
8	4	1	3	0	2	0	0
12	4	2	1	0	2	0	1
8	6	3	1	1	2	0	1
12	13	3	3	1	3	0	0
6	1	4	3	1	2	1	0
6	3	4	2	1	4	0	0
6	9	4	3	0	4	0	1
6	11	4	3	1	2	0	0
8	12	4	3	1	2	0	0
10	2	4	3	1	4	1	1
10	4	4	3	0	1	0	0
10	9	4	2	0	2	0	0
12	1	4	3	1	4	0	1
12	3	4	1	0	2	1	0
12	7	4	2	0	2	1	1
12	8	4	3	1	3	0	0
12	10	4	3	1	3	0	0
12	15	4	3	0	3	0	0
8	1	5	0	0	3	1	0
12	11	5	3	0	3	0	0

Note: Subject-level allocation data is available on Github:
https://github.com/Veconlab/Marginal-Average.

Session	ID	GE Effort	CRT	Gender	Year	Bus	Econ
6	4	6	3	0	2	1	0
6	10	6	1	0	5	0	0
8	7	6	2	0	2	0	0
8	10	6	2	0	3	0	0
10	12	6	3	1	5	0	0
12	6	6	2	0	2	0	1
6	5	7	1	1	3	0	1
8	9	7	3	0	2	0	1
10	5	7	0	0	3	0	0
10	6	7	1	1	3	0	0
6	2	8	1	1	5	0	0
6	6	8	3	1	2	0	0
6	7	8	3	1	2	0	0
6	12	8	0	0	1	1	0
8	8	8	0	0	3	0	0
10	1	8	0	0	1	0	0
10	3	8	3	1	2	0	0
8	3	9	1	0	1	0	1
10	11	9	1	1	2	0	0
12	5	9	0	0	3	0	1
12	12	9	1	0	4	0	0
12	9	10	2	0	2	0	1
8	2	12	3	1	4	0	1
8	5	12	0	0	2	0	0
8	11	12	1	0	2	0	0
10	8	12	3	0	3	0	1
12	2	12	3	0	3	0	0
12	14	12	0	0	4	0	0