Too hard to get: the role of probabilistic expectations and cognitive complexity in destructive multi-dimensional reference points

Aspen Han

Department of Economics, University of Chicago

Author Note

Aspen Han https://orcid.org/0000-0003-1474-7968

Correspondence concerning this article should be addressed to Aspen Han, Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago, IL 60637, Email: xiangyuhan@uchicago.edu

Abstract

This is the abstract.

Keywords: conflicting multi-dimensional reference points, probabilistic expectations, cognitive complexity

Too hard to get: the role of probabilistic expectations and cognitive complexity in destructive multi-dimensional reference points

1. Introduction

Consider the employee of a firm whose performance is evaluated against targets across various performance dimensions (e.g. production speed, accuracy, quality etc). For example, an assembly line worker in an electronics manufacturing plant could be subject to targets on the number of components made per hour (speed), the proportion of defective components made (accuracy), and the average durability of components made (quality). Similarly in the service sector, an Uber driver could be evaluated on the number of rides provided per month, average mileage per unit time, and the average customer satisfaction rating. It is apparent that tensions between these performance dimensions can surface, which can affect the targets' effectiveness as motivators. The emphasis for consistency and complementarity between different performance dimensions, including targets set in each, is strongly echoed in operations and general management literature (e.g. Hayes, 1984; Hayes & Schmenner, 1978; Skinner, 1974, 1996; Swamidass & Darlow, 2000). I seek to examine this concept within economics. Targets can and have been integrated into the framework of expectations-based reference points (Heath et al., 1999; Von Rechenberg et al., 2016), a growing body of research within behavioral economics. However, empirical studies have mainly examined the effects of reference points uni-dimensionally, though theoretical models encompassing multi-dimensional reference points exist. Thus, I wish to investigate the mechanisms through which reference points interact across dimensions within this economic framework, specifically answering the following research questions:

1. Do probabilistic beliefs about the achievability of reference points across multiple dimensions affect how responsive agents are to said reference points?
2. Does cognitive complexity in reconciling reference points across multiple dimensions affect how responsive agents are to said reference points?

My research is theoretically founded on the Koszegi and Rabin (2006) model of reference-dependence (henceforth KR model). Reference points have redefined preference modeling in economics. Introduced as a core component of Kahneman and Tversky's (1979; 1991) prospect theory, it posits that people evaluate outcomes relative to a reference point rather than on absolute terms and weight losses more than gains. However, they did not identify the source of reference points, which became a source of contention. The KR model endogenises the reference points to be the agent's (rational) expectations, specifically his/her probabilistic beliefs held in the recent past about what will or should happen. This accommodated alternative arguments about the origins of reference points, such as the status quo (e.g. Genesove & Mayer, 2001; Kahneman et al., 1990) and refutes to it (e.g. Plott & Zeiler, 2005; Tversky & Kahneman, 1991). Pinpointing the source of reference points was a major contribution as it allowed for more detailed studies into their effects and design, which motivates my use of the KR model as a theoretical baseline. The KR model also partially reconciles the EU theorem with prospect theory as KR considers the utility of a realized outcome to be the sum of both neoclassical consumption utility (absolute outcome levels) and gain-loss utilities (relative outcome levels), and it weights outcomes by their objective probabilities. This enables the KR model to satisfy internal consistency axioms such as transitivity which strengthens its normative appeal. However, the KR model, similar to most if not all reference point and EU models, also assumes that utilities across different dimensions of consumption are additively separate, which I seek to challenge. Yet, it seems unrealistic to think that people would view reference points in isolation from one another and determine how much to work towards each with complete disregard for the others.

Beyond the hypothetical examples and theoretical framework, my research builds upon empirical studies which have applied the KR model. Crawford and Meng (2011) found that the work patterns of New York taxi drivers could be explained by the KR model with dual reference points in daily wages earned and hours worked. While this is one of few works to consider multi-dimensional reference points, the field context made it difficult to elucidate the reference points, much less the mechanisms through which they could have interacted and affected the drivers' work behavior. Furthermore, since the taxi drivers are independent contractors, their reference points are self-imposed and hence likely consistent by construction, whereas conflicting effects are the focus of my research questions. Abeler et al. (2011) tested and verified the KR model in a laboratory experiment where subjects were set reference points in earnings and then asked to work on a real effort task. The controlled setting allowed the reference points to be exogenously induced so their effects on effort provision could be explicated. However, they only considered a reference point in a single dimension and hence neglected multi-dimensional interaction effects. Synthesizing the laboratory methodology of Abeler et al and the dual reference point model of Crawford and Meng, my undergraduate research sought to test the multi-dimensional version of the KR model. It found that when the two reference points were congruent, they had reinforcing effects, which fits with KR model predictions, but when they were conflicting, they had negating effects in that subjects seemed to ignore the reference points completely instead of compromising between them or prioritizing one over the other as predicted by the KR model. This leads to my research questions, which endeavor to identify the reasons behind this destructive effect between disparate reference points in different dimensions.

I proposed two main explanations: agents are unresponsive to reference points when they perceive the probability of being able to achieve them all in tandem to be low, or when they find it cognitively complex to reconcile the reference points, and these problems arise when reference points across multiple dimensions conflict. We can easily append these features to the KR model, which would alter the first-order conditions predicting optimal effort provision such that they align with the experimental results.

I designed an experiment to test these propositions. I elected for a laboratory experimental methodology as I want to clearly identify the decision-making mechanisms through which reference points affect effort provision, which is difficult to accomplish with observational data where there are many confounds and more clearly elicited in the controlled environment of an experiment, and resource constraints restrict it to a laboratory setting.

<limitations of design, implementation, analysis, and findings

Preliminary:

A portion of the experimental sample will be drawn from undergraduate classes in the spring quarter, for which incentive will likely be class credit instead of money due to budgetary constraints. All participants are awarded a fixed amount of credit just for participation in the experiment regardless of the amount of effort exerted. This is not a problem within the experiment as it is concerned with intrinsic effort motivation from gain-loss utilities associated with the reference points which is independent of external incentives. However, it will weaken the external validity of the findings since they draw upon a specific population and context (i.e. undergraduate students completing the task for course credit). Nevertheless, such an experiment can still provide interesting insights on the effort provision decision with respect to multi-dimensional target achievement.

Other limitations of the lab context>

<implications of findings and relations to broader literature

Preliminary: Overall, I believe my research can provide a deeper understanding of how agents integrate reference points across different dimensions into their decision-making. This is highly relevant since people are often confronted with multiple rather than singular reference points, whether when making decisions about labor supply in the workplace, school enrolment, or consumption allocation more generally. Returning to the opening example, the study can directly inform the design of targets-based incentive structures in the workplace to more effectively induce higher productivity (according to given organisational objectives). Targets should push people to strive for more but should be realistic and comprehensible. Validation of appended model (1) would suggest that management should be mindful of complementarities and conflicts across different performance dimensions when setting targets in each, weighting how important each is to the organization and how much trade-off in other dimensions each entails to arrive at some optimal matrix of targets. It also shows why and how visualization and planning exercises for target achievement can make targets more effective motivators by raising the expected probability of achievement. Verification of appended model (2) would suggest that management should set targets with clear demands, and when they are in discrete dimensions, explain how they line up so that workers can easily understand and respond. My work will also have broader implications for policies using reference points to guide decision-making in other markets such as nudge-based policies and marketing.>

2. Methodology

A. Experimental design

The experiment was conducted on Qualtrics. Subjects were asked to work on a slider task.

Subjects were randomly assigned to four treatment groups, which varied in terms of the presence of reference points (i.e. targets) in speed and accuracy, the probabilities of their work being assessed by strict or lenient assessment criterion, which respectively caught all or caught only half of mistakes made, and the extent of explanation of the reference points. Treatment 1 was the control with no reference points (and hence no explanation) and certainty of being assessed by a strict criterion. Treatments 2, 3, and 4 were set the same reference points and primed. Treatment 2 had a 75% probability of getting a lenient assessment criteria and 25% probability of strict, whereas treatment 3 had the inverse. Additionally, these two treatments explained to subjects how they should work to achieve both targets concurrently. Treatment 4 will have the same criteria assignment probabilities as treatment 2 but without the explanation of reference points. Subjects were primed to think that it would be manageable to concurrently achieve both targets under the lenient criterion but difficult to do so under the strict criterion.

B. Experimental sample and data

There are 5 number of observations.

C. Theoretical specification and hypotheses

In the experiment, the agent works on a task where he has to exert effort e, and has reference points N for the number of tasks completed per minute, and Q for the percentage of mistakes made. e is split into e_1, effort in speed, and e_2, effort in accuracy. First, consider a simplified version where outcomes are deterministic, reference points are degenerate, and gain-loss utilities are linear with constant loss aversion. Under the KR model, expected utility from effort across two dimensions is given by the KR model as

$$
\begin{aligned}
U =& p(e_1, e_2) - c(e_1, e_2) + \\
& \mu_1[(n(e_1) - N)\mathbb{I}(n > N) + \lambda_1(n(e_1) - N)\mathbb{I}(n \leq N)] + \\
& \mu_2[(Q - q(e_2))\mathbb{I}(q < Q) + \lambda_2(Q - q(e_2))\mathbb{I}(q \geq Q)]
\end{aligned}
$$

$p(e)$ is the level payoff from effort exertion, summed across both dimensions. $c(e)$ is the cost of effort. $_1[(n(e_1) - N)\mathbb{I}(n < N) + {}_1(n(e_1) - N)\mathbb{I}(nN)]$ is the gain-loss utility in the speed dimension, where $_1 0$ is the gain-loss parameter, 1 is the loss aversion parameter, and $\mathbb{I}(.)$ is an indicator function equaling 1 when the condition in the bracket holds and 0 otherwise. $_2[(Q - q(e_2))\mathbb{I}(q < Q) + {}_2(Q - q(e_2))\mathbb{I}(qQ)])]$ is analogously defined for the accuracy dimension.

To account for the role of probabilistic expectations and cognitive complexity in reference point effects, I propose appended model

$$
\begin{aligned}
U =& p(e_1, e_2) - c(e_1, e_2) + \\
& E[\mathbb{P}(\{n \geq N - \varepsilon\} \cap \{q \leq Q + \}] \times \theta \times \\
& \{\mu_1[(n(e_1) - N)\mathbb{I}(n > N) + \lambda_1(n(e_1) - N)\mathbb{I}(n \leq N)] + \\
& \mu_2[(Q - q(e_2))\mathbb{I}(q < Q) + \lambda_2(Q - q(e_2))\mathbb{I}(q \geq Q)]\}
\end{aligned}
$$

The first additional term $E[\mathbb{P}(\{n \geq N - \varepsilon\} \cap \{q \leq Q + \}]$ captures the agent's expected probability of simultaneously achieving (within some bandwidth ε of) all reference points. When this expected probability is lower, the agent weights the gain-loss utilities less and hence is less responsive to the reference points. The second additional parameter $\theta 0$ is a parameter decreasing in the cognitive complexity required to integrate the multiple reference points, so greater cognitive complexity attenuates reference point effects.

Now extending the two models to the experimental context with strict and lenient assessment criteria…

Thus, the two models provide distinct predictions for optimal effort provision in the real effort experiment[1]

KR model predictions:

- Treatments 2 and 4 will have similar positive effects on the probability of achieving any target. — (KR1)
- Treatment 3 should have a larger positive effect than treatment 2 (and 4) for achieving both targets or for achieving Q but not N. — (KR2)

Appended model predictions:

- Treatment 3 should have lower positive effect for achieving any target than treatment 2. — (A1)
- Treatment 4 should have a lower positive effect for achieving any targets than treatment 2. — (A2)

3. Analysis and results

A. Data overview

B. Probability analysis

C. Reduced form and structural estimation of loss aversion

4. Discussion

Relationship with covariates

Robustness checks

Limitations

5. Conclusion

[1] Refer to appendix for formal derivation of the first-order conditions

Bibliography

Abeler, J., Falk, A., Goette, L., & Huffman, D. (2011). Reference Points and Effort Provision. *American Economic Review*, *101*(2), 470–492. https://doi.org/10.1257/aer.101.2.470

Crawford, V. P., & Meng, J. (2011). New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income. *American Economic Review*, *101*(5), 1912–1932. https://doi.org/10.1257/aer.101.5.1912

Genesove, D., & Mayer, C. (2001). Loss Aversion and Seller Behavior: Evidence from the Housing Market. *The Quarterly Journal of Economics*, *116*(4), 1233–1260. https://doi.org/10.1162/003355301753265561

Hayes, R. H. (1984). *Restoring our competitive edge : Competing through manufacturing* (S. C. Wheelwright, Ed.). Wiley.

Hayes, R. H., & Schmenner, R. W. (1978). How should you organize manufacturing. *Harvard Business Review*, *56*(1), 105–118.

Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as Reference Points. *Cognitive Psychology*, *38*(1), 79–109. https://doi.org/10.1006/cogp.1998.0708

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental Tests of the Endowment Effect and the Coase Theorem. *Journal of Political Economy*, *98*(6), 1325–1348. https://doi.org/10.1086/261737

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263. https://doi.org/10.2307/1914185

Koszegi, B., & Rabin, M. (2006). A Model of Reference-Dependent Preferences. *The Quarterly Journal of Economics*, *121*(4), 1133–1165. https://doi.org/10.1093/qje/121.4.1133

Plott, C. R., & Zeiler, K. (2005). The Willingness to Pay–Willingness to Accept Gap, the "Endowment Effect," Subject Misconceptions, and Experimental Procedures for Eliciting Valuations. *American Economic Review*, *95*(3), 530–545. https://doi.org/10.1257/0002828054201387

Skinner, W. (1974). The focused factory. *Harvard Business Review*, *52*, 113–121.

Skinner, W. (1996). MANUFACTURING STRATEGY ON THE "S" CURVE. *Production and Operations Management*, *5*(1), 3–14. https://doi.org/10.1111/j.1937-5956.1996.tb00381.x

Swamidass, P. M., & Darlow, N. R. (2000). FOCUSED FACTORY. In P. M. Swamidass (Ed.), *Encyclopedia of Production and Manufacturing Management* (pp. 219–224). Springer US. https://doi.org/10.1007/1-4020-0612-8_355

Tversky, A., & Kahneman, D. (1991). Loss Aversion in Riskless Choice: A Reference-Dependent Model. *The Quarterly Journal of Economics*, *106*(4), 1039–1061. https://doi.org/10.2307/2937956

Von Rechenberg, T., Gutt, D., & Kundisch, D. (2016). Goals as Reference Points: Empirical Evidence from a Virtual Reward System. *Decision Analysis*, *13*(2), 153–171. https://doi.org/10.1287/deca.2016.0331

Appendix

I. Derivation of first-order conditions for experimental theoretical specification