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

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 propose two main explanations: agents are unresponsive to reference points when they perceive the probability of being able to achieve them concurrently to be low, and/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 through additional parameters which scale the gain-loss utility components, which would alter the first-order conditions predicting optimal effort provision such that they align with the experimental results.

I test these propositions with a laboratory experiment. I elected for a experimental methodology as I wanted to clearly identify the decision-making mechanisms which integrate multi-dimensional reference points, and this is most clearly elicited in the controlled experiments and difficult to establish with observational data where the reference points are elusive and there are many potential confounds. While I have linked my research motivations to the workplace, the foremost step would be to uncover general ways in which people perceive and respond to multi-dimensional reference points which are applicable to various contexts, so the abstract setting of the laboratory experiments is well-suited for it. It also provides a less costly way to verify the hypothesized mechanisms at work given the logistical constraints.

In the experiment, subjects worked on a real effort task where they had to drag sliders along a scale of 0 to 100 to designated numbers. They were evaluated on speed as measured by the number of slider sets completed and accuracy by the number completed correctly, and set targets for each metric. These two performance dimensions had inherent trade-offs as improving in accuracy necessitated spending more time on each slider to position it correctly and thus compromising on speed. The treatments varied the difficulty of achieving the targets, which augmented the probabilistic expectations of simultaneous target achievement, and the extent of explanation about the relationship between the two performance dimensions and their targets, which affected the cognitive complexity of reconciling them.

Methods

Design and implementation

The experiment was divided into two parts: subjects were asked to work on a real effort task and then complete a survey. The former provided the main data to answer the research questions, whereas the latter provided covariate data for heterogeneity and robustness analysis. A fixed amount of compensation was provided to incentivise participation. Ideally, there would have been additional incentives for effort exertion in the real effort task but this was not feasible due to budgetary constraints.

The slider task was selected to be the real effor task as because which consisted of a series of slider sets, with each set containing three sliders which could be moved over a scale of 0 to 100. To complete a set, subjects just had to drag all sliders to or past the “50” point mark, but to correctly complete it, subjects needed to correctly position every slider at its designated number (which was always weakly greater than 50).

Subjects were randomly assigned to four treatment groups, which varied in terms the reference points (i.e. targets) and how their work was assessed. Reference points were set in the two task performance dimensions: speed (tasks completed per minute) and accuracy (proportion of recorded mistakes). There were two assessment critera: strict which recorded all actual mistakes made and lenient which recorded only a quarter; these affected the likelihood of achieving both reference points concurrently. To reinforce this, subjects were primed to think that "achieving both targets [was] manageable under a lenient criterion but highly challenging under a strict criterion". Reference points were also either presented as is or explained in greater detail by mapping tasks completed per minute to maximum actual mistakes made such that both targets were met, which affected the cognitive complexity of reconciling both 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: 9 tasks completed per minute and 10% recorded mistakes, and primed. Treatments 2 and 4 had a 75% probability of getting a lenient assessment criteria and 25% probability of strict, whereas treatment 3 had the inverse. Treatments 2 and 3 had the reference points explained in greater detail, whereas treatment 4 did not. The control allows for verification of the existence of reference point effects, which is a prerequisite to identifying any changes in those effects. Comparing treatments 2 and 3 demonstrates the role of low probabilistic expectations of achievement in attenuating reference point effects, whereas comparing treatments 2 and 4 elicits the role of cognitive complexity.

Sample and context

There are 21 observations.

Theoretical specification and hypotheses

In the experiment, the agent works on a task where he/she 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. $\mu_1[(n(e_1) - N)\mathbb{I}(n > N) + \lambda_1(n(e_1) - N)\mathbb{I}(n \leq N)]$ is the gain-loss utility in the speed dimension, where $\mu_1 0$ is the gain-loss parameter, $\lambda 1$ is the loss aversion parameter, and $\mathbb{I}(.)$ is an indicator function equaling 1 when the condition in the bracket holds and 0 otherwise. $\mu_2[(Q - q(e_2))\mathbb{I}(q < Q) + \lambda_2(Q - q(e_2))\mathbb{I}(q \geq Q)])]$ 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

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

The first additional term $E[\mathbb{P}(\{n \geq N - \varepsilon\} \cap \{q \leq Q + \varepsilon\}]$ 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 $\theta \geq 0$ is a parameter decreasing in the cognitive complexity required to integrate the multiple reference points, so greater cognitive complexity attenuates reference point effects.

Extending the two models to the experimental context with strict and lenient assessment criteria, the two models provide distinct predictions for optimal effort provision in the real effort experiment[1]

KR model predictions:

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

Appended model predictions:

- A1: Treatment 3 will have lower positive effect for achieving any target than treatment 2.
- A2: Treatment 4 will have a lower positive effect for achieving any targets than treatment 2.

Results

Data overview

Effort exertion and target achievement

Relationship with covariates

Discussion

Robustness checks

Limitations

Conclusion

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

References

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

Derivation of first-order conditions for experimental theoretical specification