Behavioural and Experimental Economics
Supervision 1

Samuel Lee

Question 1

Part I

(a)

The test subjects wre randomly allocated to the two scenarios, so the risk appetites within the two test groups should be similar. In both scenarios, options A and B yield the same expected return, so the proportion of those who choose the riskier option A/C should be same among the two test groups. As it turns out, 16% of people chose A in scenario I while 69% of people chose C in scenario II.

(b)

Kahneman and Tversky describe a utility function characterised by two 'scales'. The first is a decision weight π, which 'associates with each probability p a decision weight $\pi(p)$' where $\pi(p)$ is a subjective weighting with $\pi(1)$ normalised to one, though π does not necessarily obey the Kolmogorov probability axioms. The second is a value function v which is '(i) defined on deviations from the reference point [reference dependence]; (ii) generally concave for gains and commonly convex for losses [diminishing sensitivity]; (iii) steeper for losses than for gains [loss aversion]'. In order, this means that

$$u(x|r) = v(x - r)$$

where x is the level of some good and r is the reference point; and

$$v''(x - r) < 0 \text{ for } x - r > 0 \text{ and } v''(x - r) > 0 \text{ for } x - r < 0$$

(v may not be differentiable at zero or some other points), and

$$v(|x - r|) < v'(-|x - r|)$$

given that v' exists.

Kahneman and Tversky distinguish between 'regular prospects' and 'strictly positive' or 'strictly negative' prospects. With a regular prospect, we have $x < 0 < y$ or $x > 0 > y$, and the utility associated with the prospect is

$$\pi(p_x)v(x) + \pi(p_y)v(y)$$

With a strictly positive or negative prospect, we have $x > y > 0$ or $x < y < 0$, and the utility is

$$v(y) + \pi(p_x)[v(x) - v(y)]$$

In their words, this is 'the value of the riskless component plus the value-difference between the outcomes, multiplied by the weight associated with the more extreme outcome'. This is only equivalent to the previous utility expression if $\pi(p)+\pi(1-p) = 1$, but as mentioned the weighting function does not necessarily (and in general does not) obey the Kolmogorov axioms. Whether or not we make this distinction changes the answers to the following questions, but since the lectures (and many other sources) don't seem to do so, we stick to the former equation for regular prospects. We are all the better for it since the solutions are otherwise more complicated.

i. If the reference point is the amount of money the subjects had before the start of the experiment, both scenarios involve the same prospects since options A and C result in $x-r = 2000$ with 50% probability and $x - r = 1000$ with 50% probability, while options B and D result in $x - r = 1500$ for sure. So the proportions of people choosing A and B should have the same distribution as the proportions of people choosing C and D. Furthermore, the Kahneman–Tversky value function is concave in gains, and all prospects here involve gains, so we can treat everyone as being risk-averse. If every individual i has decision weights which do not overweight a 50% probability ($\pi_i(0.5) \leq 0.5$), then everyone should be choosing options B/D, since

$$\pi_i(0.5)v_i(2000) + \pi_i(0.5)v_i(1000) \leq 0.5v_i(2000) + 0.5v_i(1000) \leq v_i(1500)$$

where the first inequality is because $\pi_i(0.5) \leq 0.5$ and the second inequality is from the concavity of v_i over gains. Therefore, options B/D are preferred.

If some individuals have decision weights which overweight a 50% probability, then the expected proportions could differ arbitrarily between the two scenarios. But in the benchmark case where we ignore decision weights and assume people maximise expected utility with the Von Neumann–Morgenstern utility function being v_i for every individual i, then the above is true.

ii. If the reference point is the amount of money the subject receives at the start of the experiment, then both prospects in scenario I involve gains (or non-losses) and both prospects in scenario II involve losses. For scenario I,

$$V_i(A) = \pi_i(0.5)v_i(1000) + \pi_i(0.5)v_i(0)$$
$$V_i(B) = v_i(500)$$

and for scenario II,

$$V_i(C) = \pi_i(0.5)v_i(-1000) + \pi_i(0.5)v_i(0)$$
$$V_i(D) = v_i(-500)$$

In the benchmark case where we set $\pi_i(p) = p$ for all i, everyone in scenario I chooses option B over option A since v_i is concave (risk-averse) over gains while everyone in scenario II chooses option D over C since v_i is convex (risk-loving) over losses.

With arbitrary π_i, we note the following: loss aversion implies that $v(x) < -v(-x)$ for $x > 0$, and for $x > y > 0$, we have

$$v(y) + v(-y) > v(x) + v(-x)$$

(The above is an assumption, not a result, and it is responsible for the value function being steeper over losses: letting y approach x gives us $v'(x) < v'(-x)$.) Therefore,

$$\begin{aligned}
&v_i(1000) + v_i(-1000) < v_i(500) + v_i(-500) \\
\implies &V_i(A) - V_i(B) + [1 - \pi_i(0.5)]v_i(1000) < V_i(D) - V_i(C) - [1 - \pi_i(0.5)]v_i(-1000) \\
\implies &V_i(A) - V_i(B) < V_i(D) - V_i(C) - [1 - \pi_i(0.5)][v_i(1000) + v_i(-1000)]
\end{aligned}$$

Since $v(x) < -v(-x)$ for $x > 0$, we get $v_i(1000) + v_i(-1000) < 0$ and it is ambiguous whether a person who chooses A over B would choose D over C or vice versa. We might not be able to say much more without additional assumptions, but we have at least shown that the proportions of people choosing A or B need not have the same distribution as the proportions of people choosing C or D.

iii. In both scenarios the expected value of money received in the experiment (including the money given at the start) is I£1500. Therefore,

$$\begin{aligned}
V_i(A) &= V_i(C) = \pi_i(0.5)v_i(500) + \pi_i(0.5)v_i(-500) \\
V_i(B) &= V_i(D) = v_i(0)
\end{aligned}$$

Now the proportions should again follow the same distribution across the two scenarios. Again, in the benchmark case where $\pi_i(p) = p$, everyone should be choosing either B or D. The final outcomes are also compatible with expectations (though they also happen to be compatible if A or C were chosen instead).

Part II

In (i) the diminishing sensitivity assumption allows us to treat everyone as being risk-averse since all prospects involve gains and the value function is concave in gains. For (ii) and (iii), we rely on loss aversion, not diminishing sensitivity, to tell us that $v(x) < -v(-x) \implies v(x) + v(-x) < v(0) = 0$ for $x > 0$, which led us to the predictions we got.

Part III

In the hypothesised case of (ii), the theory is too general and has too many degrees of freedom, whereas in the cases of (i) and (iii) we need to explain the deviations from the predicted outcomes. We probably need some more assumptions to constrain the theory somewhat, so that we can explain heterogeneity in observed choices, and it would be best if such explanations were endogenous rather than by appeal to arbitrarily different reference points, decision weights, or value functions across individuals. Given the different predictions from shifting the reference point, choosing a reference point based on maximum likelihood might mean that any test of the model has low power. It might be good to have a mechanism to determine what is the appropriate reference point, either by modelling reference points endogenously or perhaps relying on some criterion.

Question 2

(a)

With observational data, we have to identify changes in wages which arise from shifts in demand since wages are jointly determined by labour supply and demand. There are also institutional constraints which get in the way of identification, for example the limited ability of workers to vary their hours of work. Another problem is common to studies using either experimental or non-experimental data: even with perfect identification, the variation observed is typically the uncompensated elasticity of labour supply, and one must decompose this into the income and substitution effect. This is less of a problem for small and transitory wage changes, in which case the effect on lifetime wealth is small, but most wage changes are not transitory and therefore result in significant income effects.

Fehr and Goette randomly assign a wage increase to bicycle messengers in the treatment group, so it is the labour supply elasticity which is identified. The messengers were free to choose how many five-hour shifts to work during the week, there is no minimum number of shifts they must work, and there was typically at least one unfilled shift at all times, so institutional constraints on labour supply are much less present in their setup. The randomised wage increases were also transitory, which made it easier to assume away income effects.

In the studies of taxi drivers, it was less obvious whether changes in the wage rate were supply-driven or demand-driven. Camerer, Babcock, Loewenstein, and Thaler had to assume that differences in wages across days were due to demand shocks, justifying this by observing that wages were uncorrelated across days. One similarity with Fehr and Goette is that the taxi drivers were also free to choose the number of hours to work each day, though there are some constraints for drivers who rent their taxi cars: they can drive as long as they like during a continuous twelve-hour shift, which means there is a limit to how much they can allocate hours away from low-wage days into high-wage days. The wage shocks examined in the taxi driver studies are also transitory.

One drawback of the Fehr and Goette study is that shifts always comprised five hours, so while the messengers were also able to vary their number of hours within the week, they could not vary the hours as finely as taxi drivers could. In fact nobody worked more than one shift in a day. They could however vary their effort within a shift, by choosing how much they respond to calls on the dispatch radio. The authors seem to want to frame this as an advantage ('no other study that we are aware of can look at these two dimensions simultaneously'), but this potentially introduces some complications. For example, it could be difficult to interpret whether an increase in number of shifts worked accompanied by a reduction in effort per shift represents an increase or decrease in labour supply. The answer seems to be defined by whether we focus on a day or a week as the relevant unit of analysis. With this many free parameters it might be easy to accidentally generate self-fulfilling results: focusing on each day might make it seem like there is a decrease in labour supply following a wage increase, which would be compatible with a daily reference point. Also, the taxi driver studies have the advantage of including uncertainty about wages, whereas the wage increase for the treatment group in the messenger study was fully anticipated. When there is uncertainty about wages there may be a degree of precautionary working even when wages are low; this and similar considerations seem to be more relevant to real-life labour markets than a pre-announced transitory wage increase in an experimental setting.

(b)

A worker that chooses a fixed shift has to commit to that shift for at least 6 months. Therefore, they could not change their shifts following the commencement of the experiment or during the experiment. Fehr and Goette chose to focus on changes in effort during fixed shifts since the experiment could not have induced any selection effect with respect to the fixed shifts.

(c)

They assume that there is some daily income target serving as a reference point. The reasons they give for this is that the messengers can keep track of their earnings for every completed delivery, but must turn in their receipts at the end of the shift which makes it difficult for them to track their earnings over several shifts. With that it might be more likely that the reference point is based on daily earnings rather than weekly or monthly earnings. Furthermore the possibility of a daily hours-based reference point is unlikely given each shift is a 5-hour commitment, and is anyway irrelevant in the sample where nobody works more than one shift per day. There might, however, be a weekly target for the number of shifts worked.

(d)

If utility is not time-separable, then a neoclassical model could also explain a reduction in effort per shift following a wage increase. Fehr and Goette make the argument that if the reference dependent model were true, we would expect treatment effects to be heterogeneous across different degrees of loss aversion. They carry out an experiment to measure the messengers' loss aversion 8 months after the wage increase, and find that the treatment effect was negative among messengers with loss aversion and slightly positive otherwise, as one would expect with reference dependence.

Question 3

Köszegi and Rabin assume the following risk-free sub-utility function

$$u(c|r) = m(c) + n(c|r)$$

where $m(c)$ is utility in goods defined on absolute levels, and $n(c|r)$ is 'gain-loss' utility with respect to a reference point r. Utility is equal to the above function if c and r are deterministic. If c is stochastic with probability measure F, then utility is given by

$$\int_{c\in\mathbb{R}^k} u(c|r)\ \mathrm{d}F(c)$$

and if r is also stochastic with probability measure G, then utility is given by

$$\int_{c\in\mathbb{R}^k}\int_{r\in\mathbb{R}^k} u(c|r)\ \mathrm{d}G(r)\mathrm{d}F(c)$$

(a)

As an informal definition, suppose we plan to follow some state-contingent strategy, where we pick one outcome out of a choice set for every possible choice set that might become available. Setting the outcomes of this strategy as reference points, if the original state-contingent strategy we chose is optimal in every state of the world, then expectations are rational and we have a personal equilibrium.

For the shoe example, Köszegi and Rabin dispense with diminishing sensitivity and assume a linear value function. We have $m(c) = c_1 + c_2$, where $c_1 \in \{0, 1\}$ is the number of shoes and $c_2 \in \mathbb{R}_+$ is wealth which we normalise to zero when no shoes are bought. For the gain-loss utility, we have $m(c|r) = \mu(c_1 - r_1) + \mu(c_2 - r_2)$ where

$$\mu(c_i - r_i) = \begin{cases} \eta \times (c_i - r_i) & \text{if } c_i - r_i > 0 \\ \eta\lambda \times (c_i - r_i) & \text{if } c_i - r_i < 0 \end{cases}$$

with $\eta > 0$ and $\lambda > 1$. For the strategy 'Always buy', the reference point $r = (r_1, r_2)$ is stochastic and should be $(1, -p_L)$ when prices are low and $(1, -p_H)$ when prices are high. Given this reference point, we must check if buying is optimal at both $p = p_L$ and $p = p_H$. Given some arbitrary $p_M \in [p_L, p_H]$, the utility from buying is

$$U(1, -p_M|p_M, r) = 1 - p_M - q_L\eta\lambda(p_M - p_L) + (1 - q_L)\eta(p_H - p_M)$$

Not buying involves a loss of 1 shoe and a gain in p_L or p_H, so

$$U(0, 0|p, r) = -\eta\lambda + q_L\eta p_L + (1 - q_L)\eta p_H$$

For a personal equilibrium, buying must be optimal when we set $p_M = p_L$ or $p_M = p_H$. Therefore, the following conditions must be met:

$$1 - p_L + (1 - q_L)\eta(p_H - p_L) \geq -\eta\lambda + q_L\eta p_L + (1 - q_L)\eta p_H \quad (1)$$
$$1 - p_H - q_L\eta\lambda(p_H - p_L) \geq -\eta\lambda + q_L\eta p_L + (1 - q_L)\eta p_H \quad (2)$$

Condition (2) implies (1) since

$$1 - p_L + (1 - q_L)\eta(p_H - p_L) > 1 - p_H + (1 - q_L)\eta(p_H - p_L) > 1 - p_H - q_L\eta\lambda(p_H - p_L)$$

Therefore, the condition on λ is

$$\eta\lambda[1 - q_L(p_H - p_L)] \geq p_H - 1 + \eta[q_L p_L + (1 - q_L)p_H] \quad (3)$$

In the personal equilibrium, there is gain/loss utility with respect to wealth even though one is acting according to expectations, since the reference point for wealth is stochastic and the final outcome involves a partial gain of $p_H - p_L$ with probability q_L and a partial loss of the same magnitude otherwise.

(b)

The left-hand sides of (1) and (2) are the equilibrium utilities in the two states of the world, and they are both decreasing in q_L. This is because a higher q_L raises the reference point (people expect not to have to pay that much on average), which makes it especially depressing when prices turn out to be high. This goes through in both states of the world. When prices are low, there is some gain utility since we pay p_L when we expected to have to pay p_H some of the time. But this gain utility shrinks as q_L increases and the weight on the high-price state of the world in the reference point goes to zero. In other words, we don't feel as relieved when the low price is realised since we were mostly expecting it anyway. When prices are high, there is some loss utility since we had some hope that we could have paid a lower price. This loss utility increases in q_L; we were banking so much on getting a low price that we become devastated when the high price shows up instead. This might be why people sometimes try to talk themselves into having low expectations when applying for schools or jobs.

(c)

Plugging all these numbers into (3), we get

$$\lambda\left[1-\frac{1}{2}\left(\frac{9}{8}-\frac{1}{4}\right)\right]=\frac{9}{16}\lambda \geq \frac{9}{8}-1+\frac{1}{2}\times\frac{1}{4}+\frac{1}{2}\times\frac{9}{8}=\frac{13}{16} \implies \lambda \geq \frac{13}{9}$$

We have $p_{min}=\frac{1+\eta}{1+\eta\lambda}=\frac{2}{1+\lambda}$. We have the following:

$$p_L < p_{min} \iff \frac{1}{4} < \frac{2}{1+\lambda} \iff \lambda < 7$$

and any personal equilibrium involves buying shoes at p_L. There's probably not much to improve on Köszegi and Rabin's explanation, so we copy it below:

> Whatever the consumer expected, she assesses a price p paid for shoes as some combination of loss and forgone gain. Adding this gain–loss sensation to her consumption value for money, her disutility from spending on the shoes is between $(1+\eta)p$—her disutility if she had expected to pay p or more—and $(1+\eta\lambda)p$—her disutility if she had expected to pay nothing. By a similar argument, the consumer's total utility from getting the shoes is between $1+\eta$ and $1+\eta\lambda$. Hence, the expectations most conducive to buying induce a disutility of $(1+\eta)p$ from spending money and a utility of $1+\eta\lambda$ from getting the shoes, so that no matter her expectations the consumer would never buy for prices $p > p_{max} \equiv (1+\eta\lambda)/(1+\eta)$. Conversely, even given the expectations least conducive to buying, the consumer buys for all prices $p < p_{min} \equiv (1+\eta)/(1+\eta\lambda)$.

(d)

The ex-ante expected utility of the 'Always buy' equilibrium is probability-weighted sum left-hand sides of (1) and (2). We can see that the left-hand side of (2) is linearly decreasing in λ, while that of (1) is independent of λ. So the ex-ante expected utility of the 'Always buy' equilibrium decreases without bound as $\lambda \to \infty$. On the other hand, the ex-ante expected utility of never buying shoes is zero. So at some point as λ increases, the ex-ante expected utility of the 'Always buy' equilibrium will cross into negative territory, and never buying brings a higher level of utility. Again, we borrow Köszegi and Rabin's explanation:

> Intuitively, since the consumer values the shoes, she buys whenever the price is very low. The attachment to the good induced by realizing that she will do this, however, changes her attitudes toward the purchase decision. If the price turns out to be higher, she must choose between a loss of money and a loss of shoes. While buying is her best response to her expectations, it is still worse than if she could have avoided the risk of loss by avoiding the expectation of getting the shoes in the first place. More generally, because the consumer does not internalize the effect of her ex post behaviour on ex ante expectations, the strategy that maximizes ex ante expected utility is often not a PE.

If the price is always $9/8$, then the reference point when one decides to always buy is deterministic and coincides with the outcome, and there is never any gain-loss utility. The utility from adhering to the personal equilibrium is therefore just the consumption value $1-9/8 = -1/8$. If one chooses not to buy after having expected to buy, they enjoy no consumption value and they get a gain-loss utility of $-\eta\lambda + \eta\frac{9}{8}$. Therefore, buying is only an equilibrium if

$$-\frac{1}{8} \geq -\eta\lambda + \eta\frac{9}{8} = -\lambda + \frac{9}{8} \implies \lambda \geq \frac{10}{8}$$