The effect of 'smart' financial incentives on driving behaviour of novice drivers.

Duncan Mortimer[1‡], Jasper S. Wijnands[2], Anthony Harris[1], Alan Tapp[3], Mark Stevenson[2,4,5]

Keywords: road safety, behavioural economics, financial incentives, driving behaviour

Word count: 7081 (excluding Abstract, References & Tables).

[1] Centre for Health Economics, Monash Business School, Monash University, Clayton, Australia.

[2] Transport, Health and Urban Design, Melbourne School of Design, University of Melbourne, Melbourne, Australia

[3] Bristol Social Marketing Centre, University of the West of England, Bristol, United Kingdom

[4] Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Australia

[5] Melbourne School of Engineering, University of Melbourne, Melbourne, Australia

[‡]Corresponding author.

Duncan Mortimer, Centre for Health Economics, Monash University, Clayton 3800, AUSTRALIA.

Email: duncan.mortimer@buseco.monash.edu.au, Fax:+613-9905-8344, Ph:+613-9905-0735.

ABSTRACT

Recent studies have demonstrated that financial incentives can improve driving behaviour but high-value incentives are unlikely to be cost-effective and attempts to amplify the impact of low-value incentives have so far proven disappointing. The present study provides experimental evidence to inform the design of 'smart' and potentially more cost-effective incentives for safe driving in novice drivers. Study participants (n=78) were randomised to one of four financial incentives: high-value penalty; low-value penalty; high-value reward; low-value reward; allowing us to compare high-value versus low-value incentives, penalties versus rewards, and to test specific hypotheses regarding motivational crowding out and gain/loss asymmetry. Results suggest that (i) penalties may be more effective than rewards of equal value, (ii) even low-value incentives can deliver net reductions in risky driving behaviours and, (iii) increasing the dollar-value of incentives may not increase their effectiveness. These design principles are currently being used to optimise the design of financial incentives embedded within PAYD insurance, with their impact on the driving behaviour of novice drivers to be evaluated in on-road trials.

1. INTRODUCTION

Despite significant improvements in road safety, road injuries remain the ninth highest cause of burden of disease and are responsible for more than 100,000 deaths per year across high-income countries such as Australia, Japan and the United States (IHME, 2016). For high-income countries, year-on-year improvements in road safety have become increasingly marginal (IHME, 2016) and achieving further reductions in road trauma will require design and scaled delivery of novel road safety measures (Stevenson & Thompson, 2014).

The advent of in-vehicle telematics with GPS-tracking has made possible the accurate and continuous monitoring of risky driving behaviours including distance travelled, speeding, hard acceleration / deceleration, and night-time driving (Horrey et al, 2012). While this technology is now mature (Chang & Fan, 2016) and has found application for monitoring driving behaviour in commercial fleets and pay-as-you-drive (PAYD) insurance for private vehicles (Greaves & Fifer, 2011; NCTCG, 2008), there remains an opportunity for improvements and for delivery-at-scale of road-safety measures designed around in-vehicle telematics systems (Stevenson & Thompson, 2014). These opportunities include linking trip data from in-vehicle telematics with other 'big data' to predict crash risk (McLaughlin & Hankey, 2015), real-time in-vehicle alerts and/or automated throttle control via intelligent speed adaptation systems (e.g. Reagan et al, 2013), delayed or immediate feedback via a smart-phone or web-interface (e.g. Dijksterhuis et al, 2015), and financial incentives that reward good driving behaviours and/or impose penalties for risky driving behaviours (e.g. Bolderdijk et al, 2011).

The evidence regarding the effectiveness of incentives for safe driving has historically been limited to the impact of speed cameras, drink-driving legislation and the associated risks of financial penalty (Avineri et al, 2009). Recent studies have evaluated the impact of direct incentives for safe driving including exchangeable tokens plus feedback for safe on-road driving (Mazureck & van Hettem, 2006), exchangeable tokens plus/minus feedback for decreased speeding in simulated driving scenarios (Mullen et al, 2015), financial incentives plus/minus feedback for reductions in on-road speeding (Reagan et al, 2013), and the effect of behaviour-based and mileage-based PAYD vehicle insurance (or similarly structured incentives) on on-road driving behaviour (Agerholm et al, 2008;

Bolderdijk et al, 2011; Bolderdijk & Steg, 2011; Greaves & Fifer, 2011; Lahrmann et al, 2012; NCTCG, 2008) and in simulated driving scenarios (Dijksterhuis et al, 2015).

Early trials of PAYD incentives offered large monetary rewards in return for changes in driving behaviour (Bolderdijk & Steg, 2011; Greaves & Fifer, 2011). For example, one study offered up to €50 per month for keeping to the speed limit, reductions in distance travelled and reductions in weekend night-time driving; resulting in significant reductions in the percentage of total distance travelled at ≥6% above the local speed limit (Bolderdijk & Steg, 2011). While this suggests that financial incentives can influence driving behaviour, the large monetary rewards used in these studies "may not be economically feasible for insurance companies" (Bolderdijk & Steg, 2011 p18); leading some stakeholders to call for the design of 'smarter' incentives that could achieve similar shifts in behaviour but at a much lower cost.

Recent studies have demonstrated that low-value incentives can be effective when combined with feedback but attempts to amplify the effects of these low-value incentives have proven disappointing (Dijksterhuis et al, 2015). Specifically, Dijksterhuis et al (2015) combined low-value PAYD incentives (capped at €3 per simulator run) and in-car feedback (providing a running total of rewards and penalties during simulator runs) with the aim of increasing the immediacy of financial consequences arising from participants' driving behaviour. While the combination of feedback and low-value PAYD incentives produced significant improvements in driving behaviour when compared to untreated controls, varying the immediacy of feedback made little difference (€0.01/minute difference in payoffs between immediate and delayed feedback groups after feedback, equating to a €0.26 difference in payoffs for an average simulator run). Dijksterhuis et al (2015) concluded that efforts to improve the effectiveness of PAYD incentives may yet prove fruitful but that these efforts should now turn to factors other than the immediacy of feedback (such as certainty of feedback).

More generally, designing 'smart' and potentially more cost-effective incentives may be difficult to achieve in practice. Evidence from behavioural economics suggests that offering a low-value reward can have the perverse effect (contrary to that which was intended) of reducing the desired behaviour (Frey & Oberholzer-Gee, 1997; Mellstrom & Johannesson, 2008). Specifically, there is a risk that an individual's intrinsic motivation for safe driving will be 'crowded out' (eroded or displaced) by extrinsic sources of motivation such as monetary rewards. A number of explanations for this 'motivational crowding-out' have been suggested in the literature including information communicated by incentives and the reputational consequences of accepting payment. For example, Gneezy et al (2011) suggests that motivational crowding-out may be linked to the informational content of the reward; where an offer of monetary rewards could be interpreted as a signal that safe driving is difficult or unpleasant and so has to be paid for, or where the magnitude of the reward indicates the (unexpectedly low) social value of the behaviour. Alternatively, motivational crowding-out may be linked to the reputational value that an individual receives from adopting the target behaviour. When monetary rewards are present, drivers who would otherwise strive to maintain a reputation for safe driving as a signal of their concern for others, or of their community mindedness, can no longer distinguish themselves from drivers who adopt safe driving behaviours for more selfish reasons (i.e. payment). Put simply, intrinsically motivated 'safe drivers' may be less motivated to maintain the incentivised behaviour if safe driving carries no reputational value or – worse still – carries the implication that a 'safe driver' is 'in it for the money'. Motivational crowding out is likely to be much more problematic for low-value rewards simply because low-value rewards may be too small to compensate for any loss of intrinsic motivation (Culyer, 1977; Mellstrom & Johannesson, 2008; Titmuss, 1970).

It should be emphasised that the potential for motivational crowding out does not mean that monetary rewards cannot work. For individuals with low or no intrinsic motivation for safe driving, even very low-value rewards may still be effective because "...a crowding-out effect cannot occur... (where) participants have no intrinsic motivation to begin with" (Frey & Jegen, 2001 p597). For individuals with a stronger intrinsic motivation for safe driving, designing 'smart' incentives requires further information regarding the dollar-value that would be required to compensate for any loss of intrinsic motivation. Several studies provide empirical support for the effectiveness of monetary rewards in situations where participation or effort is subject to motivational crowding out, but only if the dollar-value is above the threshold where intrinsic motivation has been completely crowded-out (Gneezy & Rustichini, 2000; Heyman & Ariely, 2004).

Just as further information regarding the presence and extent of motivational crowding-out should assist in fine-tuning financial incentives, there may be scope to vary other features of an incentive to improve cost-effectiveness. Of particular relevance for the present study, incentives may be more effective when they exploit loss aversion and gain/loss asymmetry (Kahneman & Tversky, 1979). Loss aversion and gain/loss asymmetry are pervasive characteristics of preferences (Knetsch & Wong, 2009), with ratios of willingness to accept (WTA) to willingness to pay (WTP) well in excess of unity for private goods such as mugs, chocolate or hockey tickets and for public goods like environmental amenity or public infrastructure (Bischoff, 2008). Gain/loss asymmetry would imply that loss of a discount or upfront payment will have a much larger impact on driving behaviour than a reward or bonus of the same dollar value; with clear implications for the design of 'smarter' incentives. Previous tests of gain / loss asymmetry in PAYD schemes with large monetary rewards (up to €50 per month) found no significant difference between gain and loss frames (Bolderdijk et al, 2011).

This study provides empirical evidence to inform the design of 'smart' and potentially more cost-effective incentives for safe driving in novice drivers. Specifically, the study was designed to evaluate the practical significance of motivational crowding out when offering low-value financial incentives for safe driving, and the extent to which gain/loss asymmetry may be exploited to amplify the effectiveness of low-value financial incentives.

2. MATERIALS & METHODS

2.1 Study design & hypotheses

An experimental design was applied to estimate the effect of (i) financial incentives versus no financial incentive, (ii) higher-value versus lower-value financial incentives, and (iii) penalties versus rewards, on risky driving behaviours among novice drivers in a simulated environment. Here, the term penalties is used to refer to the loss of an upfront payment deposited into a 'safe driving account' (see Table 1). Participants' driving behaviours (including exceeding the posted speed limit, hard braking and excessive swerving) were observed in simulated driving scenarios designed to replicate the experience of driving on local roads under local conditions.

To identify the effect of financial incentives versus no financial incentive, the experiment included a pre/post contrast wherein we observed participants' driving behaviour at baseline under the 'no incentive' condition (baseline simulator run) and then at follow-up under the 'financial incentive' conditions (experimental simulator run). Drivers were randomised to one of four financial incentives: high-value penalty (HP); low-value penalty (LP); high-value reward (HR); low-value reward (LR); allowing us to compare high-value versus low-value incentives, penalties versus rewards and to test specific hypotheses regarding motivational crowding out and gain/loss asymmetry.

To test for motivational crowding-out, we evaluated whether providing financial incentives had the perverse effect of increasing the proportion of the population *participating* in risky behaviours; reflecting a shift in behaviour among initially safe drivers. Here, participation refers to whether we observe a non-zero level of the relevant behaviour (as distinct to non-participation, a zero level of the relevant behaviour). We hypothesised that financial incentives would *increase* participation in risky behaviours (consistent with motivational crowding out) and that this increase would be larger for low-value incentives than for high-value incentives.

To evaluate the practical significance of motivational crowding out, we evaluated the impact of incentives on the *level* of risky driving behaviours. We hypothesised that financial incentives would *decrease* the level of risky driving behaviours when averaged across all novice drivers and that high-value incentives would provoke a larger decrease than low-value incentives. That is to say, we hypothesised that – irrespective of the existence of motivational crowding out – incentives will still have the net effect of reducing risky behaviours when averaged across all novice drivers and that the usual relationship between price and quantity will prevail (pay more, get more).

Finally, to evaluate the extent of gain/loss asymmetry, we compared the impact of penalties (deductions from a 'safe driving account') and otherwise equivalent rewards on the level of risky driving behaviours. We hypothesised that penalties would be more effective than an otherwise equivalent reward (consistent with gain/loss asymmetry) in decreasing the level of risky driving behaviours.

2.2 Recruitment & sample selection

Potential participants were recruited via online classified advertisements and the strategic placement/distribution of flyers (see Supplementary Materials, File S1) at various sites in and around a university campus in Perth, Western Australia. Potential participants were directed to access the participant information sheet (see Supplementary Materials, File S2) via the university website or to contact the study's Project Officer for further information. The flyer and the participant information sheet stated that participants would receive 50 AUD (approximately 40 USD at the time of this study) to cover their time and travel expenses but that there was potential to receive higher amounts depending upon their driving behaviour in a simulated driving scenario. Individuals who contacted the Project Officer were sent a participant information sheet via email or directed to the participant information sheet online. Participants were assessed against inclusion/exclusion criteria based on information provided via online forms or over the phone.

To meet the inclusion criteria, participants needed to be between 17 and 25 years old, hold a provisional Western Australian driving licence, understand explanatory statements, instructions and questionnaires written in English, and be able to participate in a driving simulator session of approximately 45 minutes duration. Potential participants with a previous diagnosis of epilepsy were excluded. Potential participants meeting the inclusion/exclusion criteria were invited to participate and consented over the phone or via email.

2.3 Randomisation, allocation concealment & blinding

After inclusion and consent, an appointment was set for each participant to attend the location of the driving simulator for the purposes of completing a participant survey and simulator session. The Project Officer then requested a randomisation ID and treatment allocation from a team member not located at the simulator site and not engaged in any other aspect of administering the experiment. Study participants were randomised to one of four experimental conditions (HP, HR, LP, LR) using a randomisation sequence generated by a web randomisation service (www.randomization.com), with recruitment and randomisation continuing until the target sample size of 80 participants had been achieved. The team member allocating participants (DM) was blind to all participant characteristics at the point of randomisation. Participants were blind to the set of possible treatments and to the levels of their assigned treatment (high, low; penalty, reward) but it was neither possible (nor desirable) to blind participants to the structure of the financial incentives they faced during the experiment.

2.4 Delivery and data collection

At the start of each appointment, participants were asked to complete a survey (see Supplementary Materials, File S3) that included questions about their age, gender, employment status, education, income, driving experience, and attitudes to driving (adapted from Iversen, 2004). After completing the survey, participants were asked to practise driving in the simulator for 5 minutes. The driving simulator replicates the experience of driving on the road but carries a risk of simulator sickness (Brooks et al, 2010). Participants experiencing symptoms of simulator sickness – as assessed using a standard screening tool after the practice session (Kennedy et al, 1993) or based on self-report at any other stage during the experiment – were excluded from further participation.

After completing the practice session, participants completed a baseline run using the scenario developed for the study (~12.5km drive distance). Each participant's driving behaviour in the baseline run was used in our analysis to adjust for between-group differences in participants' propensity for risky driving that may have remained despite randomisation. Participants were randomly assigned to complete the baseline run of the scenario in one direction (Scenario A) and to complete the experimental run in the reverse direction (Scenario B); so that the experimental run was not simply a re-run of the baseline run. The intent was to reduce practice effects (whereby driving behaviour improved or deteriorated between baseline and experimental runs simply due to familiarity rather than exposure to the experimental treatment) and to avoid order effects (whereby driving behaviour improved or deteriorated between baseline and experimental runs simply due to the higher/lower difficulty of the experimental run).

After completing their baseline run, the Project Officer read participants a script and provided them with a printed hand-out describing the financial incentive for the participant's assigned treatment group. Prior to this point, participants were only aware of the potential to receive additional money for safe-driving behaviour in a simulated driving scenario and were blind to the specifics of the incentives that they would face in the experimental run. Table 1 reproduces text from the printed hand-out for the HP, HR, LP, and LR conditions. To replicate the impact of time pressure on real-world driving behaviour (see Fitzpatrick et al, 2017), all participants faced a time-limit for completing the experimental run and were informed that rewards (in the case of HR & LR conditions) or any remaining balance from the participant's safe-driving account (in the case of HP & LP conditions) would only be paid if the scenario for their experimental run was completed within the time-limit. Unlike participants driving under 'hurried' and 'very hurried' conditions in previous studies (Fitzpatrick et al, 2017), participants in the present study did not receive pop-up notifications comparing time elapsed against drive progress. The time-pressure on participants in the present study can therefore be characterised as relatively weak.

The dollar-values of high-value (up to 15 AUD) and low-value incentives (up to 5 AUD) described in Table 1 are broadly consistent with the dollar-value of low- and high-value incentives in previous simulator studies. For example, Hultkrantz and Lindberg (2011) specified a 'low' penalty of up to 1 Swedish Kroner (SEK) per minute of speeding (up to 20 SEK or around 3.10 AUD per 20 minute simulator run) and a 'high' penalty of up to 2 SEK per minute of speeding (up to 40 SEK or around 6.20 AUD per 20 minute simulator run). Dijksterhuis et al (2015) offered a low-value PAYD incentive "with small rates of gain" (p103) capped at €3 per simulator run (~4.70 AUD per simulator run). Mullen et al (2015) constructed a token economy in which participants could earn up to 10 USD in vouchers per simulator run (~12.90 AUD per simulator run) but made no statement to indicate whether they considered this 'high' or 'low'.

Table 1: Description of financial incentives for each treatment group

	High-value	Low-value
Reward	***Rewards for safe driving behaviours*** *You can earn an extra $15 if you drive safely in the next simulator session. This session will last for 15 minutes and will simulate risky driving situations that drivers commonly face on Australian roads. In order to earn the extra $15, you will need to keep to speed limits, obey traffic signals and give-way rules, and keep a safe distance between you and other road users. You'll also need to complete the entire course within the allocated time.*	***Rewards for safe driving behaviours*** *You can earn an extra $5 if you drive safely in the next simulator session. This session will last for 15 minutes and will simulate risky driving situations that drivers commonly face on Australian roads. In order to earn the extra $5, you will need to keep to speed limits, obey traffic signals and give-way rules, and keep a safe distance between you and other road users. You'll also need to complete the entire course within the allocated time.*
Penalty	***Penalties for risky driving behaviours*** *An extra $15 has been deposited into your 'safe-driving account'. This is separate from the $50 payment you will receive for participating today. You can keep this $15 if you drive safely in the next simulator session but the amount you can keep will get smaller every time you drive in a way that poses a risk to yourself or other road-users. This session will last for 15 minutes and will simulate risky driving situations that drivers commonly face on Australian roads. Penalties will be applied whenever you exceed speed limits, disobey traffic signals or give-way rules, or fail to keep a safe distance between you and other road users. You'll also need to complete the entire course within the allocated time. If you receive too many penalties or don't complete the course, the balance in your 'safe-driving account' will fall to zero and there will be nothing left to withdraw at the end of your session.*	***Penalties for risky driving behaviours*** *An extra $5 has been deposited into your 'safe-driving account'. This is separate from the $50 payment you will receive for participating today. You can keep this $5 if you drive safely in the next simulator session but the amount you can keep will get smaller every time you drive in a way that poses a risk to yourself or other road-users. This session will last for 15 minutes and will simulate risky driving situations that drivers commonly face on Australian roads. Penalties will be applied whenever you exceed speed limits, disobey traffic signals or give-way rules, or fail to keep a safe distance between you and other road users. You'll also need to complete the entire course within the allocated time. If you receive too many penalties or don't complete the course, the balance in your 'safe-driving account' will fall to zero and there will be nothing left to withdraw at the end of your session.*

After confirming that the participant had read and understood the instructions, the participant commenced the experimental run. During the experimental run, the driving simulator continuously tracked speed and road position as participants encountered a range of roadway situations. Simulator data were used to identify risky driving behaviours (such as exceeding the posted speed limit, hard braking and excessive swerving) and to calculate measures of the frequency and severity of these behaviours (defined below). At the conclusion of the experimental session, all participants (regardless of treatment group) were paid the high-value incentive in the full amount (regardless of driving behaviour) plus the participation incentive. The intent was to simplify the administration of payments (which would otherwise have required extraction and analysis of data from the experimental run) and to avoid potential delays between scheduled appointments. Ethical approval for this study was obtained from the University of Melbourne Human Research Ethics Committee (approval number 1646290.1); local ethics and research governance procedures were subsequently completed at the study site (Curtin University Approval Number RDHS-113-16).

2.5 Facilities

The advanced driving simulator laboratory at Curtin University in Perth, Australia, offered a safe and controlled environment in which to examine the effect of incentives on risky driving behaviours. The simulator used for the experiment is a state-of-the-art CKAS Mechatronics driving simulator. The simulator incorporates a complete Holden Commodore (a four-door family sedan with a significant share of the Australian market), allowing participants to drive using the same controls, instruments and driving systems, and in the same seat and cabin, they would use in a Holden Commodore on the road. The full motion system recreates forces, loads and sounds consistent with the participant's driving behaviour, replicating the feel of real-world driving. Participants have a full 360° view of the road and surroundings via a 'windscreen', 'side windows' and mirrors. When combined with a driving scenario that replicates street signage, street furniture, road-side surroundings, and traffic conditions of local roads, simulator runs can provide a close analogue to on-road driving.

The simulated driving scenario used in this experiment was designed specifically for this study to replicate the experience of driving on local roads under local conditions. A schematic of the simulated driving scenario including the location of intersections, speed limit advisory signs, traffic lights, pedestrians, slow-moving vehicles and other hazards has been provided in Supplementary Materials (Supplementary Materials, File S4, Figure S4.1). Stills / screenshots from video of a windscreen view and an above-following view from a drive-through of the simulated driving scenario have also been provided in Supplementary Materials (Supplementary Materials, File S4, Figures S4.2-7).

Empirical evidence demonstrates close correspondence between simulator and on-road environments for *changes* in driving behaviour due to changes in driving conditions (relative validity) (Godley et al, 2002; Mullen et al, 2015; Yan et al, 2008) but not for the *level* of risky driving behaviours (absolute validity) (Godley et al, 2002). This offers some reassurance that the *changes* in driving behaviour we observe in the simulator, will provide a good indication of the relative magnitude of *changes* we can expect in on-road driving behaviour.

2.6 Outcome measures

Data obtained from the baseline and experimental simulator runs were used to identify risky driving behaviours (exceeding the posted speed limit, hard braking and excessive swerving) and to calculate measures of the frequency and severity of these behaviours. The primary outcome for evaluating the effect of financial incentives on risky driving behaviour was total seconds exceeding the posted speed limit of 80km/h. Secondary outcomes included (i) measures of speeding at higher thresholds, namely, total seconds exceeding the posted speed limit by ≥3km/h and by ≥6km/h, (ii) measures of hard braking, namely, total seconds over two braking thresholds defined as a decelerations of at least −0.4g and −0.5g measured in units of g-force, and (iii) measures of swerving, namely, total seconds over two swerving thresholds defined as movements in the y-dimension of at least 0.05g and 0.10g measured in units of g-force.

Previous studies have defined qualitative labels for swerving, braking and acceleration events wherein a 'red event' is characterised as a "very aggressive driving manoeuvre that could result in injury or cause vehicle passengers or cargo that are not securely restrained to be shifted within the vehicle" and a 'yellow event' as "involving sufficient forces to cause passenger discomfort" (US Department of Transportation, 2014 p9). The US Department of Transportation estimated the threshold between a yellow event and a red event at around 10–15 ft/s^2, equivalent to ±0.31–0.46g, which is consistent with the values selected for our two braking thresholds (US Department of Transportation, 2014). Qualitative labels defined by the US Department of Transportation suggest that the lower thresholds adopted for swerving in the present study are more indicative of frequent lane changing than 'red' or 'yellow' evasive manoeuvres.

For all primary and secondary outcomes, the corresponding measure can take a wide range of possible values; varying from zero seconds (threshold never exceeded during the 15 minute scenario) through to 900 seconds (threshold exceeded for every second of the 15 minute scenario). To test different hypotheses, we modelled *participation* or *level* as appropriate. Recall that *participation* refers to whether we observe a non-zero level of the relevant behaviour (as distinct to non-participation, a zero level of the relevant behaviour). For participation, we dichotomised the outcome measures (some versus none) to emphasise changes between zero levels of the relevant behaviour and some positive level of the relevant behaviour. When modelling levels, we treated the outcome measure as a continuous indicator of the relevant behaviour. While participation and level emphasise different shifts in behaviour, both are drawn from the same underlying data and so should not be interpreted as independent tests of the same hypothesis.

2.7 Analysis

Estimation of treatment effects relied on between-group comparisons of risky driving behaviours during the experimental run but exploited the panel structure of the data to control for any between-group differences in participant characteristics. The simplest implementation of this empirical strategy would be to estimate Equation (1) below:

$$\log(Y_{it}) = \alpha_i + \beta_1 LP_{it} + \beta_2 HP_{it} + \beta_3 LR_{it} + \beta_4 HR_{it} + \varepsilon_{it} \quad (1)$$

Where Y_{it} captures the primary or secondary outcome for participant *i* in period *t* over two time-periods: pre (baseline run) and post (experimental run). A set of treatment variables identifies the incentive structure under which participants completed their experimental run: LP_{it} (low penalty), HP_{it} (high penalty), LR_{it} (low reward), and HR_{it} (high reward).[1] Participant fixed effects (α_i) control for any between-group differences in time-invariant participant characteristics.

In this equation, coefficients on the treatment variables give the average main effects of assignment to treatment group relative to baseline. Contrasts between treatments can then be obtained from comparisons of coefficients to quantify the relative effectiveness of different treatments. Equation (1)

[1] The treatment variables are dummy variables identifying whether or not a participant had been exposed to a high-value incentive at time *t*. For example, HP_{it}=1 for participants completing their experimental run under the high-value penalty; HP_{it}=0 for all participants during the baseline run and for participants completing their experimental run under the low-value reward (LR), high-value reward (HR) or low-value penalty (LP).

allows us to test for gain / loss asymmetry. If a high-value penalty is more effective than a high-value reward, we would expect $|\beta_2 - \beta_4|$ to be significantly greater than zero. If a low-value penalty is more effective than a low-value reward, we would expect $|\beta_1 - \beta_3|$ to be significantly greater than zero.

Equation (1) also allows us to evaluate the net effect of incentives on the *level* of the relevant behaviour. If financial incentives have the net effect of *decreasing* the level of risky behaviours, then coefficients on all four treatment variables: β_1, β_2, β_3, and β_4, should be negative and significant. If a high-value penalty is more effective than and a low-value penalty, we would expect $|\beta_2 – \beta_1|$ to be significantly greater than zero. If a high-reward penalty is more effective than a low-value reward, we would expect $|\beta_4 - \beta_3|$ to be significantly greater than zero.

For the present study, participation in risky driving behaviours was of independent interest as a test for motivational crowding out (whether incentives had the perverse effect of increasing the proportion of the study sample with non-zero levels of the relevant behaviour). To achieve this aim, we estimate the following equation:

$$D_{it} = \gamma_i + \delta_1 LP_{it} + \delta_2 HP_{it} + \delta_3 LR_{it} + \delta_4 HR_{it} + \omega_{it} \tag{2}$$

$$\text{where } D_{it} = \begin{cases} 1 \text{ if } Y_{it} > 0; \\ 0 \text{ if } Y_{it} = 0. \end{cases}$$

D_{it} is a dummy indicator for participation (some versus none) in the relevant behaviour for participant *i* in period *t* over two time-periods: pre (baseline run) and post (experimental run). Estimating the participation equation via logit or probit results in loss of more than half of our sample because the participant fixed effect predicts failure / success perfectly for participants with no pre-post change in participation. For this reason, we estimate our participation equation using a linear probability model (LPM) but re-estimate using more familiar logit models in sensitivity analyses.

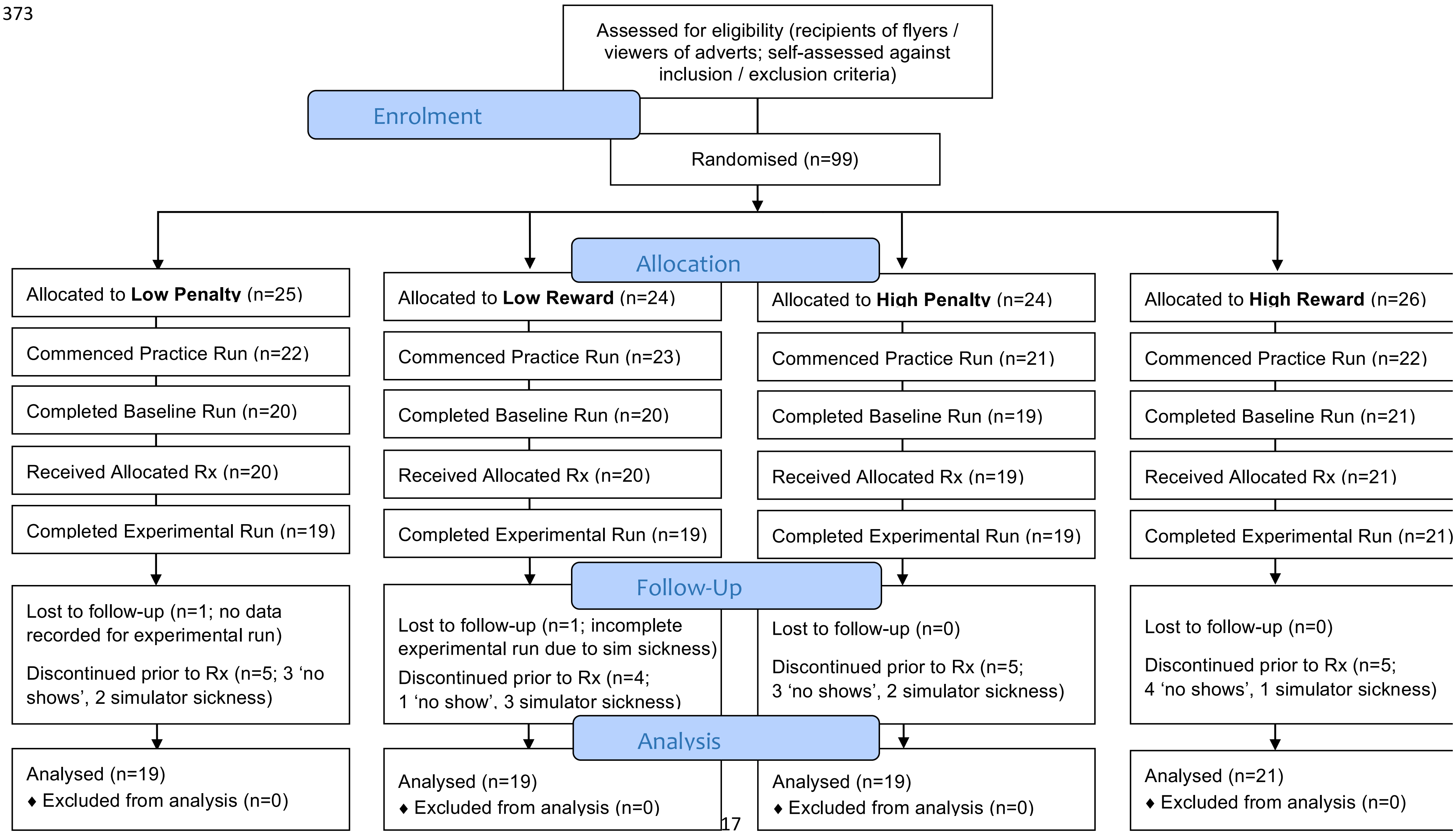
For the participation equation (Equation 2), the main parameters of interest are the coefficients on the treatment variables: δ_1, δ_2, δ_3, and δ_4. If financial incentives have the effect of *increasing* participation in risky behaviours (consistent with motivational crowding out), then coefficients on all four treatment variables should be positive and significant. If low-value rewards increase participation more than high-value rewards (consistent with motivational crowding out), then δ_3 - δ_4 should be positive and significantly greater than zero.

3. RESULTS

3.1 Recruitment & Study Sample

Recruitment commenced in September 2016 and we randomised the final participant (n=99) in mid-December 2016. After randomisation but prior to receipt of the assigned treatment, a number of participants were excluded from the study sample due to failure to attend (n=11). For participants who failed to attend their initial appointment, every attempt was made to schedule another appointment. Participants who failed to attend initial *and* subsequent appointments (n=11) were distributed equally across the four treatment groups (χ^2=1.72, p=.63).

Figure 1: Participant Flow Diagram



Similarly, a small number of participants (n=8) were excluded prior to treatment due to simulator sickness, with no significant differences in discontinuations between groups (χ^2=1.26, p=.74). Recruitment and randomisation continued until we achieved our target sample size (n=80), after replacement of no-shows and exclusions prior to treatment. Figure 1 documents the flow of participants through the experiment; including the number discontinued from each treatment group at each stage of the experiment.

Of the 80 participants who received their assigned treatment and commenced the experimental run, two participants were lost to follow-up due to failure to complete the experimental run. For one participant in the LP group, the practice scenario (different road map, no traffic, and shorter duration) was loaded instead of the scenario completed by all other participants in the experimental run. In the LR group, one participant commenced the experimental run but pulled over and stopped driving prior to completing the scenario. This participant reported symptoms of simulator sickness and was excluded as per our inclusion / exclusion criteria. As recruitment had been finalised, we were unable to replace this participant.

Table 2 summarises participant characteristics for our final study sample (n=78). Participants ranged in age from 17 to 24 years, were more likely to be male (60%), and more likely to have been born in Australia (69%). Participants were novice drivers with an average of 12.5 weeks of driving experience and an average of 6.5 hours of driving experience per week since receiving their Probationary Licence. Despite their lack of experience, participant responses on the attitudes to risky driving (ATRD) scale suggested low levels of endorsement of risky driving behaviours (mean: 28.6, range: 17-46). Participants were drawn from postcode areas with relatively low levels of socio-economic disadvantage, populated by households and individuals with relatively high levels of education and working in high-skill occupations.

Table 2: Characteristics of the Study Sample

Participant characteristics	LP (N=19)	LR (N=19)	HP (N=19)	HR (N=21)	All (N=78)
	n/N (%) or Mean (Min-Max)				
Age (years)	18.3 (17-21)	18.3 (17-23)	18.4 (17-24)	17.8 (17-20)	18.2 (17-24)
Gender (% Male)	10 / 19 (53%)	8 / 19 (42%)	12 / 19 (63%)	16 / 20 (80%)	46 / 77 (60%)
Born in Australia	14 / 19 (74%)	15 / 18 (83%)	14 / 19 (74%)	10 / 21 (48%)	53 / 77 (69%)
Weeks since probationary licence (P-plates)?	12.6 (1-26)	10.0 (0-21)	14.5 (2-27)	13.8 (1-25)	12.5 (0-27)
Ave hrs per wk of driving since awarded P-plates?	7.7 (2-15)	5.9 (1-18)	7.8 (2-21)	5.4 (1-14)	6.7 (1-21)
SEIFA Index by postcode, decile rank in Australia					
Socio-economic Disadv#	8.0 (4-10)	8.2 (5-10)	9.2 (6-10)	8.7 (6-10)	8.5 (4-10)
Education & Occupation^	7.5 (3-10)	8.1 (3-10)	8.6 (6-10)	8.4 (4-10)	8.2 (3-10)
Attitudes to risky driving ~	27.9 (18-45)	28.5 (17-46)	29.5 (20-39)	28.7 (20-42)	28.6 (17-46)

#The SEIFA Index of Education & Occupation (ABS, 2011) describes the education and occupation of individuals and households resident in a postcode area. Higher index scores indicate areas with many individuals with higher qualifications, employed in high-skill occupations. The top (bottom) decile is comprised of areas with the highest (lowest) index scores.
^The SEIFA Index of Socio-economic Disadvantage (ABS, 2011) describes the economic and social disadvantage of individuals and households resident in a postcode area. Higher index scores indicate areas with a relative lack of disadvantage. The top (bottom) decile would be comprised of areas with the lowest (highest) level of disadvantage.
~Attitudes to risky driving (ATRD) were evaluated on a five-point Likert scale over 16 items and three main factors capturing attitudes toward rule violations and speeding, attitudes toward careless driving of others, and attitudes toward drink-driving (Iversen, 2004). In our sample, the 16 ATRD items had a Cronbach's alpha of 0.8169; suggesting that these items combine to provide an internally consistent measure of the same underlying construct. Summary scores were calculated as the sum of all item-scores, after (re)coding all response data so that higher item-scores indicated endorsement of higher-risk behaviours. ATRD summary scores had a possible range of 16 to 80, with higher scores indicating more frequent endorsement of higher-risk behaviours.

Table 2 also compares participant characteristics across the four treatment groups. At a significance level of 0.05, the four groups were not significantly different with respect to any participant characteristic. The largest between-group differences were for gender (χ^2=6.36, p=.10) and country of birth (χ^2=5.61, p=.13). This is addressed by using participant fixed effects in our empirical models to control for between-group differences in observed and unobserved participant characteristics.

3.2 Motivational Crowding Out

Pre-post comparisons suggest that incentives have no statistically significant impact on *participation* in speeding (χ^2=0.24, p=.63). Table 3 summarises the pre-post changes in participation by treatment group. Here, the pre-post change in participation reflects the net effect of movement by some participants from zero levels at baseline to positive levels in the experimental run and movement by a second group of participants from positive levels at baseline to zero levels in the experimental run. For speeding, 16 / 78 (21%) participants had zero levels at baseline but positive levels in the experimental run but this increase in participation was offset by the 19 / 78 (24%) participants with positive levels at baseline but zero levels in the experimental run.

Table 3: Participation and levels of risky driving behaviours

Outcome measure	**Participation: n%**					**Levels: Mean(SD) seconds, Min-Max**				
Period	LP	HP	LR	HR	All	LP	HP	LR	HR	All
Speeding+										
Pre	63	74	37	57	58	25(38), 0-103	30(40), 0-118	22(58), 0-226	23(33), 0-105	25(42), 0-226
Post	53	63	42	57	54	10(19), 0-79	9(14), 0-42	18(44), 0-168	7(14), 0-48	11(26), 0-168
Speeding+3km/h										
Pre	32	47	21	43	36	15(27), 0-86	19(34), 0-112	18(48), 0-185	7(23), 0-155	14(33), 0-185
Post	32	37	26	19	28	6(16), 0-68	5(10), 0-38	14(41), 0-155	3(9), 0-42	7(23), 0-155
Speeding+6km/h										
Pre	26	32	16	14	22	11(21), 0-76	15(30), 0-102	15(41), 0-165	3(12), 0-56	11(28), 0-165
Post	16	21	11	10	14	4(13), 0-53	3(9), 0-34	12(37), 0-141	2(7), 0-33	5(20), 0-141
Braking−0.4g										
Pre	100	95	100	100	99	7(6), 1-25	6(5), 0-20	5(5), >0-19	5(3), 1-12	6(5), 0-25
Post	100	95	95	96	96	6(5), 2-20	5(4), 0-19	5(5), 0-20	4(3), 1-13	5(4), 0-20
Braking−0.5g										
Pre	89	89	84	81	86	4(4), 0-19	3(2), 0-9	2(3), 0-14	2(2), 0-7	3(3), 0-19
Post	100	89	79	67	83	3(3), >0-14	2(2), 0-10	2(4), 0-15	1(2), 0-7	2(3), 0-15
Swerving±0.05g										
Pre	100	100	100	100	100	17(16), 1-72	17(11), 2-43	11(10), 1-37	14(9), 3-38	15(12), 1-72
Post	100	100	100	100	100	8(6), >0-21	10(6), 2-23	10(9), 1-37	10(7), 1-28	9(7), >0-37
Swerving±0.10g										
Pre	63	68	63	71	67	4(11), 0-49	3(5), 0-18	1(2), 0-6	2(2), 0-9	3(6), 0-49
Post	68	74	74	52	67	1(1), 0-3	1(1), 0-4	1(3), 0-14	1(2), 0-8	1(2), 0-14

Table 4 reports estimated treatment effects from the regression models for (net) participation in the primary outcome relative to baseline, after controlling for participant fixed effects. Results for the full sample from the main LPM confirm that treatment effects with respect to participation are not significantly different from zero for any of the four treatments; with no evidence of a trend towards *increased* participation in speeding due to motivational crowding out. Results from more familiar logit models exclude data for a large number of participants with no pre-post change in participation (because the participant fixed effect predicts failure / success perfectly). Results from these models (not reported but available upon request) were qualitatively consistent with results for the main model and from a LPM estimated in the same sub-sample used for estimation of logit models. For participation in speeding, differences between coefficients on low-value and high-value rewards were not significant in any model (p≥.68).

Table 4: Effect of financial incentives on the primary outcome

	Participation: D(Speeding+)	Level: Speeding+	
	LPM, FE[a] (1)	OLS, FE: Log(Y_{it})[a] (2)	OLS, FE: Y_{it} (3)
Low Penalty (LP)	-0.105	-0.919	-15.71*
	(0.157)	(0.729)	(8.36)
High Penalty (HP)	-0.105	-1.030	-20.60**
	(0.157)	(0.729)	(8.36)
Low Reward (LR)	0.053	0.285	-4.73
	(0.157)	(0.729)	(8.36)
High Reward (HR)	0.000	-0.988	-15.58*
	(0.149)	(0.693)	(7.95)
constant	0.577***	0.634**	25.09***
	(0.055)	(0.254)	(2.92)
Observations (N*T)	156	156	156
Participants (N)	78	78	78

Beta coefficients with standard errors in parentheses. * p < .10, ** p < .05, *** p < .01

[a] Main model. Participant fixed effects included but results omitted for brevity.

LPM: Linear Probability Model. FE: fixed-effects within estimator.

Table 4 also reports the effect of incentives on *levels* of the primary outcome (total seconds exceeding the posted speed limit of 80km/h) relative to baseline. Results from the regression of treatment variables on the untransformed speeding data (model 3) suggest that LP, HP, and HR provoked potentially important reductions in the level of speeding which reached significance at conventional levels for the HP condition (p=.02) and approached significance for LP (p=.06) and HR (p=.054) conditions. For the main *levels* regression on log-transformed data (model 2), reductions were again observed for LP, HP and HR conditions but with a somewhat higher probability that these reductions were due to chance (p≥.16). To test if high-value incentives were more effective than low-value incentives in reducing levels of speeding, we contrasted coefficients on low-value and high-value penalties and low-value and high-value rewards; differences between coefficients for low-value and high-value incentives were not significantly different in any model (p≥.21).

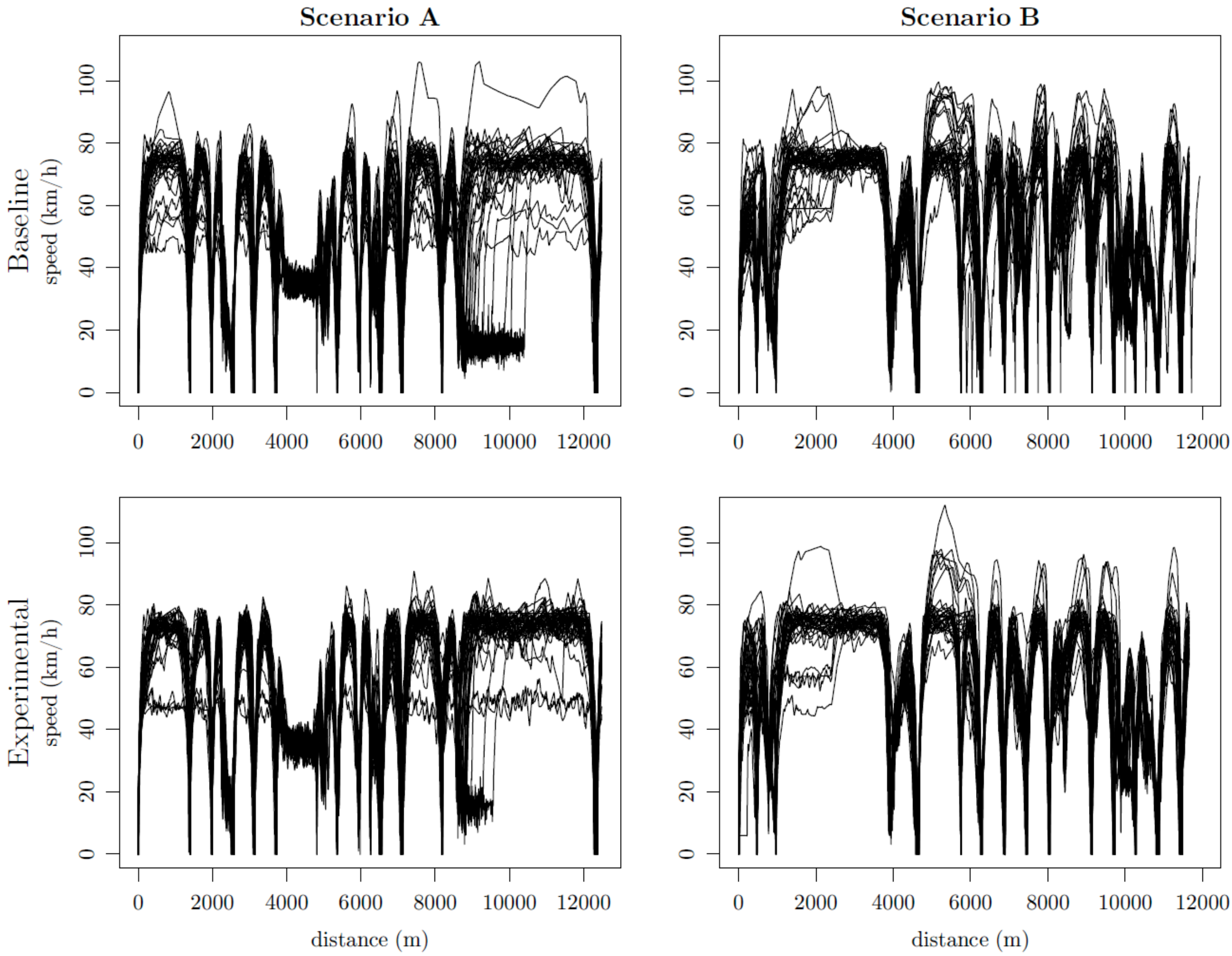


Figure 2: Pre-post change in speed by distance travelled (per participant)

Table 5 reports results for participation and levels models on secondary speeding outcomes: speeding in excess of 3km/h and 6km/h over the posted speed limit of 80km/h. While it is recognised that speeding at higher severity thresholds for even short durations of time may equate to substantial increments in risk, participation rates for speeding at higher severity thresholds were low even at baseline; leaving little room for response to treatment. Figure 2 graphs speed against distance travelled, summarising minimum/maximum speed and duration of higher severity speeding for each participant, for the baseline (top panels) and experimental runs (bottom panels), and by order of completion (Scenario A then B, or Scenario B then A).

Table 5: Effect of financial incentives on secondary speeding outcomes

	Participation: D(Speed+3km/h)	Participation: D(Speed+6km/h)	Level: Speed+3km/h		Level: Speed+6km/h	
	LPM, FE (4)	LPM, FE (5)	$Log(Y_{it})$ (6)	Y_{it} (7)	$Log(Y_{it})$ (8)	Y_{it} (9)
Low Penalty (LP)	0.000	-0.105	-0.685	-9.09	-0.733	-7.03
	(0.132)	(0.105)	(0.659)	(6.73)	(0.606)	(5.74)
High Penalty (HP)	-0.105	-0.105	-1.050	-14.26**	-0.840	-11.75**
	(0.132)	(0.105)	(0.659)	(6.73)	(0.606)	(5.74)
Low Reward (LR)	0.053	-0.053	-0.032	-4.01	-0.319	-3.01
	(0.157)	(0.105)	(0.659)	(6.73)	(0.606)	(5.74)
High Reward (HR)	-0.238	-0.048	-0.886	-4.08	-0.190	-1.30
	(0.126)	(0.100)	(0.627)	(6.40)	(0.577)	(5.46)
constant	0.359***	0.218***	-0.478**	14.44***	-1.050***	10.87***
	(0.046)	(0.037)	(0.230)	(2.35)	(0.211)	(2.00)
Observations (N*T)	156	156	156	156	156	156
Participants (N)	78	78	78	78	78	78

Beta coefficients with standard errors in parentheses. * p < .10, ** p < .05, *** p < .01
Participant fixed effects included but results omitted for brevity. LPM: Linear Probability Model. FE: fixed-effects within estimator.

As with the primary outcome, pre-post comparisons suggest that incentives have no statistically significant impact on participation in speeding at either the 3km/h (χ^2=1.06, p=.30) or 6km/h (χ^2=1.57, p=.21) threshold but these are net effects that comprise movements in to and out of participation. For speeding in excess of 3km/h over the posted speed limit, 10 / 78 (13%) participants had zero levels at baseline but positive levels in the experimental run; 12 / 78 (15%) participants had positive levels at baseline but zero levels in the experimental run. For speeding in excess of 6km/h over the posted speed limit, 5 / 78 (6%) participants had zero levels at baseline but positive levels in the experimental run; 6 / 78 (8%) participants had positive levels at baseline but zero levels in the experimental run. Results from the main LPMs (models 4 & 5) confirm that none of the four treatments provoked a change in net participation; with no suggestion of any trend towards *increased* participation as we would expect if financial incentives had the effect of crowding out a pervasive intrinsic motivation for safe driving.

Results from the regression of treatment variables on the untransformed speeding data (models 7 & 9) again suggest that HP provoked a significant reduction in the *level* of speeding. For the higher severity thresholds, reductions in speeding in response to LP and LR conditions were no longer significant ($p \geq .18$) but the direction and relative magnitudes of treatment effects remain broadly consistent with results for our primary outcome. Results from regressions on log-transformed speeding (models 6 & 8) follow a similar pattern, with HP provoking the largest magnitude reduction in speeding ($p \geq .12$) followed by LP ($p \geq .23$) or HR ($p \geq .16$), depending upon the severity threshold, but with none of these effects achieving statistical significance at conventional levels.

For secondary braking and swerving outcomes, baseline participation rates were too high to permit meaningful testing of hypotheses regarding motivational crowding out. For secondary braking outcomes at baseline, 77 participants (99%) had non-zero levels of braking at the -0.4g threshold and 67 participants (86%) had non-zero levels of braking at the -0.5g threshold. For secondary swerving outcomes at baseline, 78 participants (100%) had non-zero levels of swerving at the 0.05g threshold and 52 participants (67%) had non-zero levels of swerving at the 0.10g threshold, with no pre-post difference in participation for either threshold.

3.3 Gain/Loss Asymmetry

Table 4 reports the effect of incentives on *levels* of the primary outcome (total seconds exceeding the posted speed limit of 80km/h) relative to baseline. As reported above, potentially important reductions in speeding were observed for LP, HP and HR conditions; with these reductions reaching or approaching significance in some models. To test if penalties were more effective than otherwise equivalent rewards in reducing levels of speeding (gain / loss asymmetry), we contrasted coefficients on low-value penalties and low-value rewards and on high-value penalties and high-value rewards. Relativities between coefficients were broadly consistent across models; with the high-penalty provoking a larger reduction in speeding than the high-reward and the low-penalty provoking a larger reduction than the low-reward. For the primary outcome, these differences were not statistically significant in any model (p≥.25).

Table 5 reports the effect of incentives on *levels* of the secondary speeding outcomes: speeding in excess of 3km/h over the posted speed limit and speeding in excess of 6km/h over the posted speed limit. For the secondary speeding outcomes, penalties again provoked larger reductions in speeding than an equal value reward but differences failed to reach significance at conventional levels (p≥.19).

Table 6: Effect of financial incentives on secondary braking and swerving outcomes

	Level: Braking−0.4g		Level: Braking−0.5g		Level: Swerving±0.05g		Level: Swerving±0.10g	
	$Log(Y_{it})$ (10)	Y_{it} (11)	$Log(Y_{it})$ (12)	Y_{it} (13)	$Log(Y_{it})$ (14)	Y_{it} (15)	$Log(Y_{it})$ (16)	Y_{it} (17)
Low Penalty (LP)	-0.123	-1.137	-0.379	-0.987**	-0.737***	-8.661***	-0.434*	-3.342**
	(0.207)	(0.742)	(0.270)	(0.458)	(0.174)	(2.392)	(0.233)	(1.396)
High Penalty (HP)	-0.337	-1.416*	-0.250	-0.455	-0.516***	-6.984***	-0.724***	-2.305*
	(0.207)	(0.742)	(0.270)	(0.458)	(0.174)	(2.392)	(0.233)	(1.396)
Low Reward (LR)	0.029	-0.334	-0.196	-0.347	-0.095	-1.442	-0.256	0.061
	(0.207)	(0.742)	(0.270)	(0.458)	(0.174)	(2.392)	(0.233)	(1.396)
High Reward (HR)	-0.042	-0.421	-0.491*	-1.029**	-0.410**	-3.969*	-0.354	-0.724
	(0.197)	(0.706)	(0.257)	(0.435)	(0.166)	(2.275)	(0.222)	(1.328)
constant	1.408***	5.709***	0.628***	2.733***	2.367***	14.654***	0.565***	2.737***
	(0.072)	(0.259)	(0.094)	(0.160)	(0.061)	(0.835)	(0.081)	(0.487)
Observations (N*T)	156	156	156	156	156	156	156	156
Participants (N)	78	78	78	78	78	78	78	78

Beta coefficients with standard errors in parentheses. * p < .10, ** p < .05, *** p < .01
Participant fixed effects included but results omitted for brevity.

Table 6 reports the effect of incentives on *levels* of braking and swerving. For braking, absolute levels at baseline were very low (Braking −0.4g: mean=5.71 seconds, sd=4.61; Braking −0.5g: mean=2.73 seconds, sd=3.07) despite the very high levels of participation reported above. That is to say, most participants had non-zero levels of braking at the relevant threshold but typically only for a handful of seconds each. Pre-post reductions were smaller still and of no practical significance despite reaching or approaching statistical significance for some thresholds in some models (see results for models 11, 12 and 13 in Table 6). Results for the braking models are therefore reported only for completeness and we make no attempt to interpret results.

For swerving outcomes, LP and HP provoked significant reductions at the lower severity threshold (Swerving±0.05g). Reductions in swerving were also observed for LP and HP at the higher severity threshold (Swerving±0.10g) and for the HR condition at the lower severity threshold (Swerving±0.05g) but treatment effects were smaller in magnitude and failed to reach statistical significance at the .05 level in some models. LR had an approximately zero impact on swerving in all models and at both thresholds.

To test if penalties were significantly more effective than rewards in reducing levels of swerving, we again contrasted coefficients on low-value penalties and low-value rewards and high-value penalties and high-value rewards. At the lower severity threshold (Swerving±0.05g), LP was significantly more effective than LR for model 14 ($F=6.78$, $p=.01$) and model 15 ($F=4.55$, $p=.04$). At the higher severity threshold (Swerving±0.10g), the larger reduction in response to LP approached significance for untransformed swerving data (model 17: $F=2.97$, $p=.09$) but not for log-transformed data (model 16: $F=0.29$, $p=.59$). Despite achieving consistently larger reductions in swerving under the HP condition than under the HR condition, differences between HP and HR coefficients were small in magnitude and failed to reach significance for all models and at both thresholds ($p\geq.25$).

4. DISCUSSION

Results suggest that incentives can influence driving behaviour in simulated driving scenarios. The high-value penalty provoked potentially important reductions in speeding and swerving across a range of different severity thresholds and model specifications, though these reductions fell short of statistical significance in some models. Moreover, results suggest that 'smart' but lower-cost financial incentives have the potential to reduce risky driving behaviour. Our low-value penalties provoked statistically significant reductions in swerving and reductions in speeding outcomes (including the primary outcome) that were similar in magnitude to those observed for high-value incentives.

With respect to motivational crowding out, we hypothesised that financial incentives would *increase* participation in risky behaviours and that this increase would be larger for low-value incentives than for high-value incentives. Results were inconsistent with both of these hypotheses suggesting that, at least for novice drivers in a simulated driving scenario, motivational crowding out is unlikely to negate the positive effects of low-value incentives on safe driving.

With respect to gain/loss asymmetry, penalties provoked larger reductions in speeding than equal value rewards but there remained a relatively high probability that this difference was due to chance ($p \geq .19$). We found stronger evidence for gain / loss asymmetry for swerving where low-value penalties were more effective than low-value rewards, though this difference between penalties and rewards was no longer statistically significant for high-severity swerving in log-transformed models. These results suggest that gain / loss asymmetry may usefully inform the design of incentive structures that do more with less (Kahneman & Tversky, 1979); making payments (or penalties) for safe driving more economically feasible for insurance companies or government agencies (Bolderdijk & Steg, 2011).

The present study offers a number of lessons for researchers and policy makers seeking to build an evidence-base to inform the design of smarter financial incentives. First, the present study evaluated the impact of incentives with respect to several risky driving behaviours. While the *level* of speeding and swerving varied between individuals and between baseline and experimental simulator runs (suggesting sensitivity to differences in driving behaviour), braking outcomes were subject to floor effects that limited their usefulness for evaluating treatment effects. Put simply, the very low levels of braking observed during the baseline run left little room for response to treatment and little opportunity to test our hypotheses regarding motivational crowding out and gain/loss asymmetry. For *participation*, both braking and swerving were subject to ceiling effects; with a very high proportion of participants braking or swerving over the relevant threshold (but for just a handful of seconds in each run). This meant that we were unable to test for the presence / absence of motivational crowding out using either braking or swerving data.

For swerving, ceiling effects in participation data may partly reflect severity thresholds that are more indicative of lane changing than of 'red' or 'yellow' evasive manoeuvres. In such circumstances, higher *levels* of swerving would reflect frequent lane changing and interpretation of *levels* of swerving as a measure of risky driving remains appropriate (Dula & Geller, 2003). For braking, floor and ceiling effects were unlikely to have been introduced as a consequence of our choice of severity thresholds or instrumentation / measurement of forces in the simulator environment. The thresholds against which our braking outcomes were defined are consistent with the US Department of Transportation threshold for a red event at our upper severity threshold and equate to a mid-way point between a yellow and red event for our low severity threshold (US Department of Transportation, 2014). The very high rate of participation in hard braking within our study sample is surprising given that a yellow event would "involv(e) sufficient forces to cause passenger discomfort" and a red event "could result in injury or cause vehicle passengers or cargo that are not securely restrained to be shifted within the vehicle" (US Department of Transportation, 2014 p9). It may be that our sample of novice drivers and an unfamiliar simulator environment combined to increase the probability of braking miscalculations but that the potential for multiple red or yellow events within a short and relatively 'simple' simulator run, remained low. Future studies may wish to provide novice drivers with the opportunity for practice runs of a longer duration than the 5 minutes afforded to participants in the present study and / or employ more 'difficult' or 'complex' driving scenarios for baseline and experimental runs.

Second, while details of the assigned treatment were withheld from participants until just before the experimental run and *after* completion of the baseline run; the initial call for participants and participant information sheet made it clear that the study offered an opportunity to receive additional money for safe-driving behaviour. For this reason, all groups may have driven more safely than usual in the baseline *and* experimental runs and the estimated treatment effects may therefore be an underestimate of the behavioural response to incentives. Along similar lines, estimated treatment effects may also be underestimated simply because the "knowledge of being monitored can alter driving behaviour" (Bolderdijk & Steg, 2011). Previous studies have demonstrated the potential significance of Hawthorne effects (see Dixit et al, 2017 for a review) and provide some guide to the likely direction and magnitude of any associated bias in estimated treatment effects. Of particular relevance, any Hawthorne effects present at baseline are likely to persist for the duration of our relatively short experimental protocol (Agerholm et al, 2008; Hultkrantz and Lindberg, 2011) and it is unlikely that a weakening of Hawthorne effects between baseline and experimental runs could account for the observed reductions in risky driving behaviours. For comparison between our four experimental conditions, any Hawthorne effects should be equivalent between groups but there remains the possibility that safer than usual driving during the baseline run (*and* experimental runs) may have limited scope for a behavioural response to incentives in the experimental run.

Third, our tests of motivational crowding out relied on pre-post and between-group comparisons with respect to participation in the relevant behaviour. However, estimates of pre-post and between-group differences were the net effect of some participants moving from zero to positive levels of the relevant behaviour and a second group of participants moving from positive to zero levels of the relevant behaviour. In short, motivational crowding out in initially 'safe' drivers may have been wholly or partly offset (and disguised in the net effect on participation) by cessation of unsafe driving behaviours in others. We disaggregate this net effect by reporting numbers of participants moving in and out of participation for each outcome but sample size considerations precluded estimation of treatment effects in the sub-sample of initially 'safe' drivers (or estimation of interactions between treatment group and ATRD scores in our full sample). Repeating our experiment in separate samples of 'safe' and 'unsafe' drivers (e.g. with higher ATRD scores) may shed further light on the potential for motivational crowding out and response to treatment in different population sub-groups.

Fourth, our tests of gain / loss asymmetry relied on comparisons between penalties and 'otherwise equivalent' rewards. While every effort was made to ensure that the penalties and rewards were described in neutral language and framed in a similar manner, some differences in wording between penalties and rewards were required in order to accurately describe the relevant incentive structure. It is possible that these minor differences in wording were partly responsible for our finding that penalties were more effective than rewards. If so, it is also possible that use of a different framing /

wording to describe penalties or rewards may have produced a smaller or larger effect than observed in the present study. While our results are consistent with the weight of evidence regarding loss aversion and gain / loss asymmetry (Knetsch & Wong, 2009), further research evaluating the impact of wording / framing effects may help to optimise presentation of both penalties and rewards.

Fifth, while our results regarding motivational crowding out provide reassurance that we can avoid the perverse result of paying good money to achieve worse outcomes, we still need to set the incentive at a dollar-value sufficient to change behaviour but not so high as to compromise cost-effectiveness. Optimising the design of smart but lower-cost incentives therefore requires detailed information regarding the relationship between price and outcomes. Results suggest that, at least over the range considered in the present study, low-value and high-value incentives have much the same effect on speeding and swerving. Hultkrantz and Lindberg (2011) report similar results, finding no difference in effectiveness between a 'low' penalty of up to 1 SEK per minute of speeding and a 'high' penalty of up to 2 SEK per minute of speeding. The dollar-value of our low-value incentive was similar to (Dijksterhuis et al, 2015; Hultkrantz and Lindberg, 2011) or *lower* than (Mullen et al, 2015) the dollar-value of incentives offered in previous simulator studies. Despite this fact, even our low-value incentive would be unaffordable if converted to a dollar-value per minute of safe-driving and directly applied to policy-holders in PAYD insurance schemes. Along similar lines, Dijksterhuis et al (2015) note that "even very small rates of gain" in simulator studies would translate into "an unrealistically large amount of money… under a real PAYD system" (p103). Further research may therefore be required if our aim is to achieve the best balance between cost and effectiveness for *very* low value incentives in the simulator environment and for low value incentives in PAYD insurance.

Sixth, findings from a number of previous studies suggest rewards are (unsurprisingly) preferred to penalties (e.g. Wit & Wilke, 1990; see also Dijksterhuis et al, 2015 for a brief review). In the present study, 'penalties' were structured as the loss of an upfront payment deposited into a 'safe driving account'; with the upfront payment designed to redress differences in acceptability between rewards and penalties. A similar approach has been employed in previous studies, with penalties framed as deductions from an upfront discount on the PAYD insurance premium (Bolderdijk et al, 2011). Under this type of incentive structure, recruitment into an on-road trial of pay-as-you-speed (PAYS) insurance proved disappointing and the authors concluded that a 30% premium discount was insufficient to sell PAYS to young drivers (Lahrmann et al, 2012). The magnitude and framing of the upfront payment or discount is therefore crucial in ensuring acceptability of penalties embedded within PAYD insurance. In the present study, we found no differences in discontinuations from the experiment between penalty and reward conditions (see Figure 1). Nonetheless, further work will be required to adapt our penalties (deductions from a 'safe-driving account') for application in commercial PAYD insurance and to test acceptability of the resulting product in novice drivers.

Finally, post-hoc power calculations suggested that the present study may have been underpowered to identify treatment effects relative to baseline for some outcomes. Replication of the present study using a larger sample size may provide more definitive conclusions regarding the design of 'smart' and potentially more cost-effective incentives for safe driving.

5. CONCLUSION

While it is well-known that incentives can influence behaviour, maximising the impact of incentives on specific behaviours, in specific populations requires detailed and context-specific evidence. This study provides fine-grained evidence to inform the design of 'smart' and potentially more cost-effective incentives for safe driving in novice drivers. Our findings suggest that penalties may be more effective than rewards of equal value, such that it may be possible to exploit gain/loss asymmetry to amplify the effectiveness of financial incentives. Our findings also suggest that even low-value incentives can deliver net reductions in risky driving behaviours and that, at least over the range considered in the present study, low-value incentives have much the same effect on speeding and swerving as high-value incentives. Collectively, these findings would suggest that low-value penalties are likely to offer a more cost-effective means of reducing risky driving behaviour than high-value rewards. These design principles are currently being used to optimise the design of financial incentives embedded within PAYD insurance, with their impact on the driving behaviour of novice drivers to be evaluated in on-road trials (Stevenson et al, 2018).

ACKNOWLEDGMENTS This work is supported by Australian Research Council Grant Number: LP150100680. MS is supported by a National Health and Medical Research Council Fellowship. The authors thank Seraina Agramunt and Lynn Meuleners for assistance with administering the experimental protocol.

COMPETING INTERESTS None.

ETHICS APPROVAL Australian Research Council, Melbourne University Human Research Ethics Committee, Curtin University Human Research Ethics Committee.

REFERENCES

Agerholm N, Waagepetersen R, Tradisauskas N, Harms L, Lahrmann HI. (2008). Preliminary results from the Danish Intelligent Speed Adaptation project Pay as You Speed. *IET Intelligent Transport Systems,* 2: 143-153.

Australian Bureau of Statistics (ABS). *2033.0.55.001 - Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA)*, Australian Bureau of Statistics: Australia, 2011

Avineri E, Chatterjee K, Darnton A, Goodwin P, et al. (2010). Individual behaviour change: Evidence in transport and public health. The Department for Transport, London, UK. http://eprints.uwe.ac.uk/11211/2/

Bischoff, I. (2008), 'Endowment effect theory, prediction bias and publicly provided goods: an experimental study', *Environmental and Resource Economics*, **39**: 283-96.

Bolderdijk, J. W., Knockaert, J., Steg, E. M., & Verhoef, E. T. (2011). Effects of Pay-as-you-drive vehicle insurance on young drivers' speed choice: Results of a Dutch field experiment. *Accident Analysis & Prevention*, 43, 1181-1186.

Bolderdijk JW, Steg L (2011) Pay-as-you-drive vehicle insurance as a tool to reduce crash risk: Results so far and further potential. Discussion Paper No. 2011-23: ITF/OECD Joint Transport Research Centre.

Brooks J, Goodenough R, Crisler M, Klein N, Alley R, Koon B, Logan W, Ogle J, Tyrrell R, Wills R. (2010). Simulator sickness during driving simulation studies. *Accident Analysis & Prevention*, 42(3), 788-796.

Chang S, Fan C. (2016). Identification of the technology life cycle of telematics: A patent-based analytical perspective. *Technological Forecasting and Social Change*, 105:1-10.

Culyer AJ. Ethics and economics in blood supply. *The Lancet* 1971;**297**:602-3.

Dijksterhuis C, Lewis-Evans B, Jelijs B, de Waard D, Brookhuis K, Tucha O. (2015). The impact of immediate or delayed feedback on driving behaviour in a simulated Pay-As-You-Drive system. *Accident Analysis & Prevention*, 75, 93-104.

Dixit V, Ortmann A, Rutström E, Ukkusuri S. (2017). Experimental Economics and choice in transportation: Incentives and context. *Transportation Research Part C: Emerging Technologies*, 77: 161-184.

Dula C, Geller S. (2003). Risky, aggressive, or emotional driving: Addressing the need for consistent communication in research. *Journal of Safety Research*, 34: 559-566.

Fitzpatrick C, Samuel S, Knodler M. (2017). The use of a driving simulator to determine how time pressures impact driver aggressiveness. *Accident Analysis & Prevention*, 108: 131-138.

Frey B, Oberholzer-Gee F. The cost of price incentives: an empirical analysis of motivation crowding-out. *American Economic Review*1997;**87**:746-55.

Frey B, Jegen R. Motivation crowding theory. *Journal of Economic Surveys* 2001;**15**:589-611.

Gneezy U, Rustichini A. Pay enough or don't pay at all. *Quarterly Journal of Economics* 2000;**115**:791-810.

Gneezy U, Stephan Meier S, Rey-Biel P (2011). When and Why Incentives (Don’t) Work to Modify Behavior. Journal of Economic Perspectives 25: 191–210.

Godley S, Triggs T, Fildes B. (2002). Driving simulator validation for speed research. *Accident Analysis & Prevention*, 34(5), 589-600.

Greaves S, Fifer S. Analysis of a financial incentive to encourage safer driving practices. Proceedings of the 34th Australasian Transport Research Forum. Adelaide, 2011.

Heyman J, Ariely D. Effort for payment: a tale of two markets. *Psychological Science* 2004;15:787-93.

Horrey W, Lesch M, Dainoff M, Robertson M, Noy Y. (2012). On-Board Safety Monitoring Systems for Driving: Review, Knowledge Gaps, and Framework, *Journal of Safety Research*, 43: 49-58.

Hultkrantz L, Lindberg G. (2011). Pay-as-you-speed: An Economic Field Experiment. *Journal of Transport Economics and Policy*, 45: 415-436.

Institute for Health Metrics and Evaluation (IHME). *GBD Compare Data Visualization*. Seattle, WA: IHME, University of Washington, 2016. Available from http:// vizhub.healthdata.org/gbd-compare. (Accessed 8^{th} June 2017)

Iversen H. (2004). Risk-taking attitudes and risky driving behaviour. *Transportation Research Part F: Traffic Psychology and Behaviour*, 7(3), 135-150.

Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica* 1979;**47**(2):263-91.

Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology, 3*(3), 203-220.

Knetsch, J.L and W-K. Wong (2009), ‘The endowment effect and the reference state: evidence and manipulations’, *Journal of Economic Behaviour and Organization*, **71**: 407-413.

Lahrmann H, Agerholm N, Tradisauskas N, Berthelsen K, Harms L. (2012). Pay as You Speed, ISA with incentives for not speeding: Results and interpretation of speed data. *Accident Analysis & Prevention*, 48: 17-28.

Mazureck U, van Hattem J. (2006). Rewards for safe driving behaviour: influence on following distance and speed. *Transformation Research Record : Journal of the Transportation Research Board,* 1980: 31-38.

Mclaughlin S, Hankey J. (2015). *Naturalistic Driving Study: Linking The Study Data To The Roadway Information Database*. Report S2-S31-RW-3, Strategic Highway Research Program 2 (SHRP 2), Transportation Research Board, Washington D.C. http://Nap.Edu/22200

Mellstrom C, Johannesson M. Crowding out in blood donation: was Titmuss right? *Journal of the European Economics Association.* 2008;**6**:845-63.

Mullen N, Maxwell H, Bédard M. (2015). Decreasing driver speeding with feedback and a token economy. T*ransportation Research Part F: Traffic Psychology and Behaviour*, 28, 77-85.

NCTCG (North Central Texas Council of Governments). Pay as you drive (PAYD) insurance pilot program phase 2 final project report. Prepared by Progressive County Mutual Insurance & North Central Texas Council of Governments, November 2008.

Reagan IJ, Bliss JP, Van Houten R, Hilton BW. The effects of external motivation and real-time automated feedback on speeding behavior in a naturalistic setting. *Hum Factors* 2013;**55**:218-30.

Stevenson M, Harris A, Mortimer D, Wijnands J, Tapp A, Peppard F, Buckis S. (2018). The effects of feedback and incentive-based insurance on driving behaviours: study approach and protocols. *Injury Prevention*, 24: 89-93.

Stevenson M, Thompson J. (2014). On the road to prevention: road injury and health promotion. *Health Promot J Austr*, 25: 4-7.

Titmuss R. *The Gift Relationship: From Human Blood to Social Policy*. London: Allen and Unwin, 1970.

US Department of Transportation (2014). *Study of the Impact of a Telematics System on Safe and Fuel-efficient Driving in Trucks*. US Department of Transportation, Federal Motor Carrier Safety Administation, Washington DC. https://rosap.ntl.bts.gov/view/dot/177

Wit A, Wilke H. (1990). The presentation of rewards and punishments in a simulated social dilemma. *Social Behaviour*, 5, 231–245.

Yan X, Abdel-Aty M, Radwan E, Wang X, Chilakapati P. (2008). Validating a driving simulator using surrogate safety measures. *Accident Analysis & Prevention*, 40(1), 274-288.

SUPPLEMENTARY MATERIALS

S1. Recruitment Flyer

S2. Participant Information Sheet

S3. Participant Survey

S4. Schematic & screen-shots for simulated driving scenario