RISK PREFERENCES, TIME PREFERENCES AND SMOKING BEHAVIOUR

by

Glenn W. Harrison, Andre Hofmeyr, Don Ross and J. Todd Swarthout†

January 2017

ABSTRACT

There is a rich theoretical literature in economics which models habit-forming behaviours, of which addiction is the exemplar, but there is a paucity of experimental economic studies eliciting and comparing the preferences that economic theory suggests may differ between addicts and non-addicts. We evaluate an incentive-compatible risk and time preference experiment conducted on a sample of student smokers and non-smokers at the University of Cape Town in 2012. We adopt a full information maximum likelihood statistical framework, which is consistent with the data generating processes proposed by structural theories and accounts for subject errors in decision making, to explore the relationship between risk preferences, time preferences and addiction. Across different theories and econometric specifications we find no differences in the risk preferences of smokers and non-smokers but do find that smokers discount significantly more heavily than non-smokers. We also identify a nonlinear effect of smoking intensity on discounting behaviour and find that smokers may be more likely to discount hyperbolically than non-smokers, which means they may be more prone to time inconsistency. These results highlight the importance of the theory-experimental design-econometric trinity and have important implications for theories of addiction.

Keywords: smoking, discount rates, risk aversion, time inconsistency, addiction

JEL codes: I1, D81, D91

† Department of Risk Management & Insurance and Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, USA (Harrison); School of Economics, University of Cape Town, South Africa (Hofmeyr); Department of Philosophy, University College Cork, Ireland, School of Economics, University of Cape Town, South Africa, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, USA (Ross); and Department of Economics, Andrew Young School of Policy Studies, Georgia State University, USA (Swarthout). Harrison is also affiliated with the School of Economics, University of Cape Town. E-mail contacts: gharrison@gsu.edu, andre.hofmeyr@uct.ac.za, don.ross931@gmail.com and swarthout@gsu.edu. We are grateful to the University of Cape Town and the American Cancer Society for funding this research. These funders had no involvement in any aspect of the research project nor in the preparation of this manuscript.

I. INTRODUCTION

Addiction is a puzzle for economic theory: how can rational-agent modelling accommodate the fact that most addicts expend resources to acquire their targets of addiction but simultaneously incur real costs to try and reduce or limit their consumption of these goods? Furthermore, why is the typical course of addiction characterised by repeated unsuccessful attempts to quit prior to final abstention? From the standpoint of standard consumer theory in economics these patterns of behaviour are difficult to rationalise.

A number of economists over the years have risen to the challenge. In Section II we review these efforts, and conclude that making further progress, especially in critically bringing economic modelling of addiction to bear on psychological and clinical studies, requires as a first step more rigorous specification and identification of the relationships between structural risk and time preferences, on the one hand, and statistical vulnerability to addiction, on the other. Such progress requires careful experimentation to calibrate parametric relationships among preference structures and choices that generate, sustain, and mitigate addiction. We undertake such experimentation, using regular smokers as the representative addicts.

We evaluate a risk and time preference experiment conducted on a sample of student smokers and non-smokers at the University of Cape Town (UCT) in 2012. An incentive-compatible experimental design allows us to explore potential differences in the risk and time preferences of smokers and non-smokers and jointly estimate utility function curvature and discounting functions. We find no significant differences in the risk preferences of smokers and non-smokers but do *find that smokers discount the future significantly more heavily than non-smokers*. These results are robust to different assumptions about the way people evaluate lotteries and the way they discount utility flows. In addition, we identify *a nonlinear effect of smoking intensity on discounting behaviour* and find a "total effect" that suggests *smokers may be more likely to discount hyperbolically than non-smokers*, which, under the assumption of an additively-separable intertemporal utility function, means they may be more prone to time inconsistency.

This research makes a number of contributions to the literature. Instead of adopting the standard two-step approach to data analysis (see Section III), which is statistically invalid, we estimate risk and time preference parameters as a linear function of observable characteristics (e.g., age, gender, and smoking status) so that the uncertainty of the parameter estimates propagates into the inferences which are drawn from the data.

In addition, when analysing risk preferences and smoking behaviour, we allow risk attitudes to be determined both by utility function curvature and probability weighting. Prior studies in the literature either focus on utility function curvature or probability weighting, but not on both together.

This is only the second study in the smoking-discounting literature to incorporate utility function curvature in the estimation of time preference models, and it is the first which allows rank-dependent utility theory to characterise choice under risk. In addition, this is the first study to identify a nonlinear relationship between smoking intensity and discounting behaviour. Smoking more cigarettes is associated with increased discounting but only up to a point, after which each additional cigarette is associated with lower discounting.

The design and analysis are sensitive to recognition that multiple decision processes characterise the discounting of delayed rewards. It is crucial for researchers to be cognisant of this fact when exploring the addiction-discounting relationship. Smokers may be more likely than non-smokers to discount hyperbolically and this may be a factor in their addiction.

Following the review of economic models of addiction in Section II, Section III reviews previous research on the relationship between risk preferences, time preferences and smoking behaviour. Section IV discusses our experimental design and presents summary statistics for the sample. Section V formulates our statistical approach to data analysis. Section VI presents the results and Section VII concludes.

II. ECONOMIC MODELS OF ADDICTION

Existing work by economists in modelling addictive consumption may be grouped into two broad approaches.[1]

The first approach, often referred to in the literature as *rational addict* modelling, was pioneered by Becker and Murphy (1988). It attributes addiction to unusual properties of certain goods, which causes flows of utility from their consumption to accumulate as capital that incentivises further consumption and reduces marginal utility from non-addictive substitutes. On this kind of account, agents fall into addiction without at any point behaving contrary to their consistent preferences, and it is not even necessary to posit uncertainty about outcomes or forecasts of utility.

Rational addiction models have been widely criticized for systematically mis-predicting the patterns of temporary cessation and relapse, followed by eventual success in achieving control, that characterises the typical life course of an addiction (e.g., Ross (2010)). The natural prediction of the basic Becker and Murphy (1988) model is that an addict will simply keep consuming their addictive target unless and until its price rises beyond the point where its consumption is optimal at the moment of choice. The model does, however, offer a prediction, which psychologists have

[1] Outside of the two general approaches we review, some economists have favoured models in which the dynamics of addictive processes occur outside the logical space of economic agency, even if within the brain and nervous system of the person (e.g., Laibson (2001), Loewenstein, O'Donoghue and Rabin (2003), Gul and Pesendorfer (2007)). In such models, addictive temptations are exogenous sources of costs to maintenance of consistent or welfare-maximising choice that under some circumstances overwhelm the agent's budget of resources for resistance. A charitable way of interpreting such models is that they treat addiction as a purely psychological phenomenon with which economic agents must cope, in the same sense that they must deal with barriers to optimisation arising exogenously from the physical and social environments. These models raise complicated methodological and philosophical issues. If exogenous addiction risks arising in the interaction between nervous systems and environments can be identified and specified independently of the economic model, then we can think of the model as simply declining to model addiction as an economic phenomenon in the first place. In that case, however, we need to be able to recover latent preference structures, including temporal discount functions and risk attitudes, by reference to direct observation of neural or computational processes rather than inference from choices. Some research programmes in neuroeconomics aspire to such discoveries (see Glimcher (2011)). For reasons given elsewhere (see Ross (2011, 2014a, 2014b)), we regard the prospects for success of such programmes as limited. We here set aside the models of addiction that comport with them as not, strictly speaking, economic models *of addiction*, even though they remain economic models of latent processes that *respond* to addiction. For reasons presented in Andersen et al. (2010), these remarks should not be taken as expressing scepticism about prospects for integrating economic and psychological modelling in general; we merely do not think that such an integrated model of addiction has yet been specified.

generally considered reasonable, about the characteristics of people who are likely to be most vulnerable to addiction: those who discount future utility most steeply.

Orphanides and Zervos (1995) added an additional dimension to rational addict modelling by incorporating uncertainty on the part of potential consumers about the extent of their vulnerability to addiction when they first sample potentially addictive goods. This model yields the further prediction, which has again been regarded by many addiction scientists and clinicians as intuitive, that risk aversion, both instantaneous[2] and intertemporal, should be a protective factor against addiction.

The second broad approach to economic modelling of addiction responds to criticisms of rational addict models for failing to capture the observed synchronic and diachronic preference ambivalence of most addicts that is reflected in their apparent efforts to resist and modify their own revealed preferences for addictive goods. Economists have attempted to deal with this by complicating the agency of addicts in one or both of two ways: with either diachronic or synchronic dual self models.

Diachonic dual self models (Winston (1980), Thaler and Shefrin (1981), Schelling (1984), Gruber and Köszegi (2001), Bénabou and Tirole (2004)) divide the addicted agent into temporal successions of sub-agents that implement divergent temporal discounting functions. Both Gruber and Köszegi (2001) and Bénabou and Tirole (2004) incorporate the quasi-hyperbolic (QH) intertemporal discounting model of Laibson (1997) to explain why addicts choose to consume addictive targets at a present moment while simultaneously preferring to refrain from such consumption in the future. Such a pattern implies inconsistent choice over time by the succession of sub-agents considered as a group. Diachronic dual self models can then capture varying levels of success in resolving such ambivalence by allowing for variation in the extent to which addicts accurately recall or predict their own preference histories and courses. Thus, these models also involve choice under uncertainty.

[2] The prefix "instantaneous" is used to differentiate instantaneous risk and time preferences from intertemporal risk preferences. Intertemporal risk preferences refer to preferences over intertemporal lotteries, the outcomes of which may be temporally correlated. By contrast, instantaneous risk and time preferences define, respectively, atemporal attitudes to risk and uncertainty, and the valuation of goods which are available at different points in time. We only empirically examine instantaneous risk preferences so all subsequent references to "risk preferences" refer to the instantaneous or atemporal variety.

By contrast, *synchronic dual self* models incorporate sub-agents that compete for control of the agent's choices at a given point in time (Benhabib and Bisin (2004), Bernheim and Rangel (2004)). In these models, the competing agents again differ from one another in the intertemporal discounting behaviour that they implement when they respectively gain control, and also face varying degrees of uncertainty concerning the implications of addictive consumption for present welfare, future welfare, or both. Fudenberg and Levine (2006, 2011, 2012) develop models that combine diachronic and synchronic complexity of agency. While varying in their details and the specific behavioural phenomena they are designed to capture, the three Fudenberg and Levine models share as their core strategic interaction and partial conflict between short-run sub-agents ("selves") that are relatively less patient than, and relatively more risk averse than, long-run sub-agents ("selves").

As we document in Section III with specific reference to addictive smoking, psychological studies of addiction have also focused recurrently on steep temporal discounting and relative indifference to risk as factors that may contribute to the formation and persistence of addiction; for a review of psychological literature of this kind going beyond smoking, see Ross et al. (2008) chapters 3 and 4. There is, furthermore, increasing consensus among psychologists that addictions are learned, and modifiable by incentivising interventions (Redish, Jensen and Johnson (2008), Heyman (2009)). Psychologists might therefore be expected to welcome efforts by economists to contribute improved specification precision and technical rigour with respect to the empirical identification of risk and time preference idiosyncrasies that distinguish addicts.

It thus constitutes a significant gap in the literature that economists have not yet directly empirically estimated differences in risk and time preferences, specified with full theoretical precision, between addicts and non-addicts. An important aspect of such precision is to respect the need for joint estimation of risk and time preferences established by Andersen et al. (2008). Unsurprisingly, none of the many empirical studies by psychologists of temporal discounting differences between addicts and non-addicts attempt, or indeed recognise the importance of, such joint estimation. In its absence, as Andersen et al. (2008) demonstrate theoretically and empirically, discount rate estimates are significantly biased upward for risk averse agents, which is also likely to result in mis-estimation of whether their structure is exponential, hyperbolic, or quasi-hyperbolic. Nor have structural interactions between risk and time preferences been explicitly specified in existing economic models of addiction. Such specification as it might feature in the distinguishing characterisation of addicts cannot be based on a priori theorising, but depends on empirical data.

Our empirical comparison of temporally indexed and risky choice behaviour in a sample of smokers and a sample of non-smokers is motivated by this concern with improved economic modelling of addiction in general. We chose to study smokers for three reasons: nicotine is the most readily available addictive drug in general populations; there is widespread agreement among addiction scientists and clinicians that almost all regular, daily smokers meet the criteria for addiction (West (2006)); and the relative non-interference of nicotine with basic cognition and judgment makes nicotine addicts a natural starting point population for any new laboratory paradigm.

A noteworthy feature of the limited existing empirical literature on addiction and risk and time preferences is that the latter are invariably measured in the domain of responses to monetary rewards, despite the fact that the most directly relevant arguments of utility functions where addiction is concerned refer to social and health status. While it is possible that most people's risk and time preferences are closely related across domains, this cannot be assumed, especially in a population that is already atypical in being characterised by addiction. It is practically challenging to address the question of cross-domain preference structure consistency in the laboratory using non-hypothetical rewards. Arguably, the best long-run methodology for handling this difficulty will be to use laboratory work on choices over money as a baseline for extensions into the field where participants' choices affect their real health and social well-being. In that case the first stage research involving monetary rewards is the immediate priority.

In our view, improved unification of economic and psychological approaches to addiction is most likely to the extent that research in both disciplines is alert to a self-conscious philosophical orientation. We are guided by the approach that Ainslie (1992, 2001) has dubbed 'picoeconomics' (see also Ross et al. 2008). This approach emphasises, as does Heyman (2009), that addiction is in large part learned behaviour, expressed through choices that are 'voluntary' in the non-metaphysical sense of being responsive to incentives.

The framework of Ainslie (1992, 2001) recognises that both exogenous and endogenous neurophysiological and neurochemical states and processes give rise to vulnerabilities and barriers to controlling addiction that an economic model will represent as variable costs. The picoeconomic model emphasises the role of inconsistent intertemporal discounting in generating and maintaining addictive choice patterns, but it does *not* predict, counterfactually, that most human choice over time reflects hyperbolic discounting. Rather, it applies a philosophical thesis that consistent valuation of rewards over time requires explanation and should not simply be assumed as a natural default disposition. Ainslie himself emphasises the importance of "personal rules," that is, self-enforcing linkages between discrete choices that should be reflected in agents' revealed preferences, but he is also alert to the importance of institutional and other environmental "scaffolding" (Clark (1997)) as providing support for intertemporally consistent valuation and choice.

Economists and psychologists, notwithstanding their different practical priorities, can join in seeking explanation of addiction in breakdowns and loopholes in personal rules, in challenges to their implementation resulting from errors in risk perception and estimation, and in strategic complications in the relationships between individuals and their social environments.

III. REVIEW OF LITERATURE ON RISK PREFERENCES, TIME PREFERENCES AND SMOKING BEHAVIOUR

Smoking is known to be one of the primary behavioural risk factors for the additional utilisation of health resources and expenditures on health. For just over 50 years the U.S. Surgeon General has been collating careful epidemiological evaluations of the causal effect of smoking on a large number of diseases (U.S. Department of Health and Human Services (2014)). And major litigation efforts have generated estimates of additional health expenditures running into the hundreds of billions of dollars (Coller, Harrison and McInnes (2002)). Evidently, a better understanding of the determinants of smoking behaviour continues to have great significance for health policy.

Smoking involves an intertemporal trade off that should be apparent: any short-term benefits from smoking are coupled with the potential for large long-term costs. In addition, the decision to smoke involves risks that should be apparent, such as the possibility of negative health consequences, and is made under conditions of uncertainty, without the agent knowing his or her susceptibility to these risks.

Table 1 provides a detailed summary of experimental studies investigating the relationship between smoking and time preferences. Online searches of PubMed and Econlit, employing the search criteria "smoking" and "discounting" and their variants (e.g., "smoke", "discount", and "time preference"), were used to locate these studies. An initial list of over 50 studies was trimmed according to the following criteria: the study had to include a clear smoker, non-smoker comparison[3]; study participants had to make choices between amounts of real money, rather than cigarettes or quality-adjusted life years, available at different points in time[4]; and the instrument used to assess discounting had to include at least 20 questions.[5] The 31 studies satisfying our inclusion criteria are listed in Table 1; a detailed discussion of this table is provided in Appendix A.

[Table 1 here]

The last column of Table 1 reports whether a significant statistical relationship was found between smoking and discounting behaviour. A "positive" relationship between smoking and discounting means that smokers discount more heavily than non-smokers, consistent with expectations before the reported observations. Some of the entries in Table 1 report findings from several studies or from different treatments in the same study. For example, Baker, Johnson and Bickel (2003) report results from real and hypothetical experimental treatments whereas Chabris et al. (2008) report

[3] A number of studies (e.g., Field et al. (2006), Dallery and Raiff (2007), Epstein et al. (2003)) focus purely on discounting among smokers and were excluded due to the lack of non-smokers in the sample.
[4] Odum, Madden and Bickel (2002) and van der Pol and Ruggeri (2008) focus on the discounting of health outcomes and Field et al. (2006) and Odum and Baumann (2007) focus on the discounting of hypothetical cigarette rewards.
[5] Some panel studies, such as the Health and Retirement Study (HRS), include a module to assess discounting behaviour but the limited number of questions (e.g., three questions in the HRS, see Bradford (2010)) makes precise estimation and inference difficult. Hence these studies were excluded.

findings from multiple studies. In some cases (e.g., Baker, Johnson and Bickel (2003)) results were the same across studies and treatments, while in others (e.g., Chabris et al. (2008), Heyman and Gibb (2006)) they differed. The last column of Table 1 therefore summarises the set of 37 reported findings from the 31 studies.

Of the 37 reported findings in Table 1, 29 were positive and significant while the remaining 8 were null results.[6] Thus, the bulk of findings in this literature point to a positive relationship between smoking and greater discounting behaviour, irrespective of whether real or hypothetical rewards, long or short temporal horizons, choice or titration elicitation mechanisms, small or large samples, and simple or complex statistical procedures were used.

From a statistical perspective, the most striking feature of Table 1 is the near-universal two-step approach to data analysis. This approach entails using nonlinear least squares (NLLS), or some similar technique, to estimate discounting parameters at the level of the individual, and then using the, typically log-transformed, point estimates as data in subsequent statistical models. Harrison, Lau and Rutström (2010) (HLR) and Hofmeyr et al. (2016) are the only studies in Table 1 which do not use this method. Their reasons for avoiding it are sound. The problem with the two-step approach, aside from typically relying on tiny samples to estimate discounting parameters at the level of the individual, is that estimated discounting parameters are estimates, not data. Such estimates comprise both a point estimate (of the mean) and a standard error, and to use only the point estimate is to throw away information on the sampling variability of that estimate.

Moreover, using an estimated discounting parameter as data violates one of the statistical assumptions of the second-stage models: that the covariates are measured

[6] Some studies classified smokers using more than one category (e.g., heavy and light smokers in Stillwell and Tunney (2012)), others classified non-smokers using more than one category (e.g., never-smokers and ex-smokers in Bickel, Odum and Madden (1999)), and still others separated male and female smokers and non-smokers (e.g., Jones et al. (2009) and HLR). In a few of these cases, comparisons between some of the groups were significant while others were not, which makes coding the study problematic. Studies were therefore coded as having found a significant result if at least one smoker, non-smoker comparison was statistically significant. This procedure is preferable to coding a study as having found no statistically significant results just because one comparison (between, say, light smokers and non-smokers) was not significant even though another comparison (between, say, heavy smokers and non-smokers) was significant.

without error. Thus, statistical inferences drawn from this approach are simply invalid. HLR and Hofmeyr et al. (2016) estimate discounting parameters as a linear function of observable characteristics (e.g., age, gender, and smoking status) so that the uncertainty of the discounting parameter estimates propagates into the inferences which are drawn from the data. This valid statistical approach will be used here.

Table 2 provides a detailed summary of studies investigating the relationship between smoking and risk preferences. Unlike the literature on time preferences and smoking behaviour, there is a dearth of studies analysing the risk preferences of smokers and non-smokers. Online searches of PubMed and Econlit, employing the search criteria "smoking" and "risk preferences" and their variants (e.g., "smoke," "risk", and "probability discounting"), were used to locate these papers. An initial list of studies was trimmed according to the following rules: the study had to include a clear smoker, non-smoker comparison[7]; and study participants had to have made choices between lotteries[8] involving amounts of money, rather than cigarettes or quality-adjusted life years.[9] The 11 studies satisfying our inclusion criteria are listed in Table 2; a detailed discussion of this table is deferred to Appendix B.

Table 2 shows that a majority of the studies (8 out of 11) adopted the probability discounting (PD) approach to risk preferences, which defines risk aversion solely in terms of the shape of the probability weighting function (PWF).[10] The PD model is just Yaari's (1987) dual theory of choice under risk limited to a circumscribed class of lotteries and with a specific probability weighting function: $\pi(p) = p / [p + \gamma(1 - p)]$. If $\gamma > 1$ this specification represents probability pessimism and risk aversion. As subjective probability distortions drive risk preferences in the PD framework, it is surprising that 6 out of these 8 studies only used 5 probabilities in the elicitation task;

[7] Lawyer et al. (2011) investigate whether the risk preferences of smokers and non-smokers differ when they make choices over hypothetical or real rewards. However, they do not compare the risk preferences of smokers and non-smokers.
[8] A number of studies (e.g., Bradford (2010), Jusot and Khlat (2013)) use survey questions which try to elicit general attitudes toward risk and were excluded for this reason.
[9] van der Pol and Ruggeri (2008) investigate risk preferences over hypothetical health outcomes.
[10] Of these studies, 3 also employed the area under the curve (AUC) method of Myerson, Green and Warusawitharana (2001). When using the AUC method, one calculates the area under a subject's derived certainty equivalents and normalizes this to lie in the closed unit interval. Larger AUCs imply less risk aversion and, thus, the AUCs of smokers and non-smokers can be compared to determine whether the groups differ in their risk preferences.

the remaining two studies (Mitchell (1999) and Yi, Chase and Bickel (2007)) only used 6 and 7 probabilities, respectively.

[Table 2 here]

The final column of Table 2 shows whether the studies found a significant statistical relationship between risk preferences and smoking behaviour: the results are equivocal and, other than HLR, the statistical analyses are not valid. A positive relationship between smoking and risk preferences means that smokers are more risk averse than non-smokers, whereas a negative relationship means that smokers are less risk averse than non-smokers. Null results were reported in 3 studies, positive results were reported in 5 studies, and negative results were reported in 3 studies.[11] These conflicting results cut across different elicitation mechanisms, real and hypothetical rewards, different frameworks for choice under risk, and different methods of analysis. Thus Table 2 shows that the relationship between risk preferences and smoking behaviour, or lack thereof, differs markedly across studies.

Table 2 also shows that every study except HLR again adopted a two-step approach to statistical analysis: NLLS is used to estimate risk preference parameters at the level of the individual and then these point estimates are used as data in subsequent statistical models. For the reasons outlined above, this approach is statistically invalid.

We add to the extant literature by simultaneously investigating the relationship between risk preferences, time preferences and smoking behaviour using an incentive-compatible experimental design, a relatively large sample of South African university students, and a statistical framework which allows one to draw robust inferences about smokers and non-smokers.

[11] Some studies classified smokers using more than one category (e.g., heavy smokers and light smokers in Poltavski and Weatherly (2013), and smokers and "triers" in Reynolds et al. (2003)), and HLR separated male and female smokers and non-smokers. We again adopt the classification scheme that codes a study as having found a statistically significant result if at least one smoker, non-smoker comparison was significant, even if all comparisons were not.

IV. EXPERIMENTAL DESIGN AND SUMMARY STATISTICS

We recruited 175 subjects from undergraduate classes at UCT. Given the focus on smoking behaviour, sign-up sheets included a simple screening question: "Do you smoke cigarettes (Yes / No)." A pool of over 900 people applied to take part in the study and individuals from the smoking and non-smoking groups were randomly selected for inclusion in the project. Those who were selected were added to a website which allowed them to sign up for an experimental session that did not conflict with their academic timetable.

The experiment took place in a computer lab at UCT which had been set up to run the risk and time preference software developed by us. Subjects were separated by partitions and not allowed to talk to each other during the session. The experiment was conducted in August 2012 across 10 sessions. The median group size was 17 participants and one of us assumed the role of experimenter for every session; two research assistants (RAs) were also employed to help administer subject payments and answer questions.

Upon arrival at the lab, subjects were randomly allocated to computer terminals and given an overview of the tasks that they would complete. Subjects then signed informed consent before being taken through a detailed presentation of the risk or time preference task.[12] The order of these tasks was counter-balanced across sessions so subjects either performed the risk or time preference task first. Participants were given the opportunity to ask questions at any stage of the presentations or during the tasks. After questions had been addressed, subjects completed the first task.

Once all participants had completed the first task, the experimenter went through a detailed presentation of the other task. Subjects then completed this task before filling out a questionnaire which collected standard demographic characteristics and information on smoking behaviour. The experimenter or RAs then determined their

---

[12] The introductory presentation, the risk preference task presentation, and the time preference task presentation are included in Appendix C. The presentations were designed to make the tasks transparent and easy to understand. The payment system was also discussed in detail so that subjects understood how their final earnings were determined. This attention to detail, coupled with salient rewards, promotes incentive compatibility and the truthful revelation of preferences.

earnings for the tasks. All subjects received a show-up fee of R20. Earnings for the risk preference task were paid out immediately in cash and earnings for the time preference task were paid out on the date corresponding to the subject's choice on the randomly selected discounting question. Delayed payments were effected via electronic transfer and subjects received a payment notification on their cell phones as soon as the transfer took place. Such transfers are a common means of payment in South Africa and were used to reduce the transaction costs which subjects would have had to incur by coming to collect their delayed payments from us. Experimental sessions lasted approximately an hour and subjects earned R370 (roughly $66 at purchasing power parity (PPP) at the time) on average.

A. Risk Preference Task

The risk preference interface was based on Hey and Orme (1994). It presented subjects with a choice between two lotteries on a screen, displayed as pie charts with accompanying text that listed the probabilities of the prizes. Figure 1 shows a screenshot of the risk preference task. The display seen by subjects used colours, allowing for greater discrimination than might be apparent from a monochrome presentation (e.g., for the Right lottery).

[Figure 1 here]

The task used prize magnitudes between R0 and R280 (roughly $0 to $50 at PPP at the time) and probabilities which varied in increments of 0.05 between 0 and 1. Thus, other than HLR, this study used larger lottery prizes than any of the studies in Table 2 which have incentive-compatible experimental designs. In addition, this study had more variation in the probability domain than every other study in Table 2. This variation provides for enhanced sensitivity to any probability weighting that might be present.

The lottery pairs in the task were based on the set developed by Loomes and Sugden (1998) (LS) to test different stochastic specifications of choice under risk. LS designed the lottery pairs to accommodate a wide range of risk preferences, to provide good coverage of the probability space, and to generate common-ratio tests of expected utility (EU) theory. However, all the lotteries over which each subject made choices had the same context (i.e., the same set of prizes).[13] By contrast, we used four prize contexts in the experiment: (R0, R140, R280), (R40, R80, R240), (R20, R100, R220), and (R60, R120, R180). Incorporating a number of different prizes and probabilities is helpful for the separate identification of the utility function and the PWF in models which admit both sources of risk preferences (e.g., rank-dependent utility theory).

[Figure 2 here]

Figure 2 shows the set of Marschak-Machina (MM) triangles representing the lotteries, and lottery pairs, which were used in the risk preference task. The top of each diagram lists the context of the lotteries (e.g., (R0, R140, R280)) and the gradient of the lines connecting lottery pairs. Each point in the MM triangle represents a lottery and the line connecting two, or more, points represents a lottery pair, or set of lottery pairs, on offer in the choice task. Figure 2 shows that the risk preference task provided thorough coverage of the MM triangle, in the sense of including a combination of interior and boundary choices, and that it captures the full range of risk preferences, under the null hypothesis of EU theory: risk-loving (gradients less than 1), risk neutral (gradients equal to 1), and risk averse (gradients greater than 1). Subjects made 40 choices in the risk preference task and one choice was selected at random at the end of the experimental session for payment.

B. Time Preference Task

The time preference task presented subjects with choices between smaller, sooner (SS) and larger, later (LL) monetary rewards. Figure 3 shows a screenshot of the time preference task. On each screen subjects had to make 4 choices before proceeding to the next screen. The principal (i.e., SS reward) and time horizon were fixed on each screen but varied across screens. A calendar was displayed on every screen to show the subjects when they would receive the amounts of money they chose.

[Figure 3 here]

[13] LS used two experimental treatments: one where subjects made choices over lotteries defined on the context ($0, $10, $20) and one where subjects made choices over lotteries defined on the context ($0, $10, $30). The probability distributions over these contexts were identical across the two groups except for 8 out of the 45 lotteries in the task.

Following Coller and Williams (1999), three front end delays (FEDs) to the SS rewards were used: zero days, 7 days, and 14 days. This design allows one to hold subjective transaction costs constant for the SS and LL rewards at positive FEDs. It also facilitates estimation of the parameters of a quasi-hyperbolic or β-δ discounting function because the zero day FED allows one to pin down the estimate of β, which captures a "passion for the present" or "present-bias" in decision making, whereas the positive FEDs allow one to estimate the long-term discounting parameter δ. Subjects in an experimental session were only exposed to one of these FED treatments.

Two principals (R150 and R250: roughly \$27 and \$45 at PPP at the time), 14 time horizons between the SS and LL rewards (7 to 98 days, in 7-day increments), and nominal annual interest rates between 5% and 250% were used in the time preference task. These parameters define a battery of 224 possible choice pairs. Each subject made 60 choices in the task which were drawn randomly, without replacement, from this battery. At the end of the experimental session, one of these choices was randomly selected for payment.

C. Summary Statistics

Table 3 presents summary statistics for the sample of 175 students. The average age in the sample is approximately 20 years old, 42% of the sample is White[14], two-thirds are enrolled in the Commerce faculty at UCT, and approximately one-third receives financial aid. Smokers were defined as those people who answered "yes" to the question: "Do you currently smoke cigarettes?" Current smokers make up 62% of the sample[15] and this is the largest number of smokers (i.e., 108 smokers) ever recruited for a study exploring risk preferences and smoking behaviour. Smokers were deliberately oversampled to investigate whether intensity of smoking is related to risk

[14] Designation of population groups or 'races' follows the traditional categorization in South Africa that is still employed in affirmative action and related policies in South Africa, notwithstanding recognition that it involves cultural and historical discriminations that are without biological significance. Approximately 24% of the sample is Black. 14% is Coloured, a culturally salient population group in South Africa composed of individuals of mainly Indonesian descent who speak Afrikaans as a first language. 17% is Indian. The remaining 3% preferred not to classify their race.

[15] The remaining 38% of the sample comprises both former-smokers and never-smokers who will be referred to collectively as non-smokers.

and time preferences. The mean number of cigarettes smoked per day is 8.67 with a standard deviation of 5.81 and a range of 1 to 25.[16,17]

[Table 3 here]

Smokers also completed the Fagerström Test for Nicotine Dependence (FTND) due to Heatherton et al. (1991). The FTND is a measure of smoking severity that scores people on a scale of 0 to 10, where higher numbers indicate greater severity. The average FTND score among smokers is 2.22 with a standard deviation of 2.08. Thus, on average, the smokers in this sample are relatively light smokers. In addition, given the young age of the sample, the smokers' lifetime exposure to cigarettes is relatively low. In the literature on risk preferences, time preferences and smoking behaviour, researchers often try to maximise the difference between smokers and non-smokers by selecting heavy smokers to take part in the study. We recruited smokers across the entire spectrum of severity to determine whether being a smoker, irrespective of intensity, is associated with risk and time preferences. This also allows us to explore the relationship between smoking severity and risk and time preferences.

Table 3 shows that randomisation across experimental treatments ensured that approximately 50% of the sample completed the risk preference task prior to the time preference task. FED treatments were split evenly across the sample and 50% of choices in the time preference task involved the high principal of R250.

[16] Estimates from the *South African National Health and Nutrition Examination Survey (SANHANES-1)* of the mean number of cigarettes smoked per day for people aged 15-24 is 5.9 (Shisana et al. (2013, p. 111)). For the population as a whole, the mean number of cigarettes smoked per day is 8.5. Thus, our sample, at least in terms of the mean number of cigarettes smoked per day, is very similar to the general population.

[17] According to *The Tobacco Atlas* (see www.tobaccoatlas.org and Eriksen et al. (2015)), 22.2% of men and 9% of women smoke tobacco daily in South Africa. The prevalence rate for men is lower than in other middle-income countries but the prevalence rate for women is higher than in other middle-income countries. Prevalence rates for selected high-income countries are: US – men: 17.2%, women: 14.2%; UK – men: 23.2%, women: 20.3%; Australia – men: 15.1%, women: 11.6% ; Germany – men: 28%, women: 22.2%.

V. STATISTICAL SPECIFICATION

The statistical method we employ is direct estimation by maximum likelihood of structural models of latent choice processes. The latent choice processes in question are captured by models of risk and time preferences. These models provide the structure necessary to estimate risk and time preferences using the observed choice data. One of the benefits of the maximum likelihood approach is that it uses all of the available information to estimate discounting and risk preference parameters and the precision of these estimates. This estimation strategy closely follows Andersen et al. (2008) and HLR so we provide a brief explanation of the method below, focussing on the canonical cases of EU theory and exponential (E) discounting. Further details are provided in Appendix D. We also discuss the extension to other risk and time preference models.

Assume that utility of income is defined by a power utility function which displays constant relative risk aversion (CRRA):

$$U(y) = y^r, \tag{1}$$

where y is a lottery prize in the risk preference task and r is a parameter to be estimated.

To estimate the parameter r we formed a latent index, based on latent preferences, that captured the difference in the expected utility of the Right and Left lotteries presented to subjects. The value of this index, for each observation, was determined by the lottery prizes, their associated probabilities, and an initial estimate of r. This latent index was linked to the subjects' binary choices (i.e., the Left or Right lottery) using the cumulative normal distribution function. This "probit" link function determined the likelihood of selecting the Left lottery, and hence the likelihood of selecting the Right lottery, for each observation in the dataset given the value of the latent index. Maximum likelihood estimation was then used to determine the value of r that maximised the likelihood of observing all of the data from the experiment.

It is a straightforward extension to make the parameter r a linear function of individual characteristics in order to draw robust inferences about potential differences in the risk preferences of participants. In addition, every estimate of r

includes a standard error which reflects our uncertainty as to the "true" value of r. This stands in sharp contrast to the bulk of studies in Table 2 which use risk preference point estimates as data in subsequent statistical models. We also extended the model by adopting the "contextual utility" (CU) behavioural error specification of Wilcox (2011) to allow mistakes on the part of subjects from the perspective of the deterministic EU model and to draw robust inferences about the primitive "stochastically more risk averse than" relation.[18]

It is a simple matter to incorporate other theories of choice under risk in this statistical framework. Quiggin (1982) developed the rank-dependent utility (RDU) model, which assumes that a decision maker transforms objective probabilities into subjective decision weights which are then used to evaluate lotteries. In this context, we estimate the parameters of a utility function and PWF which maximise the likelihood of observing the data from the experiment on the basis of a latent index which captures the difference in the rank-dependent utility of the lotteries.

We estimate EU and RDU models to compare the risk preferences of smokers and non-smokers. In addition, we estimate the parameters of a variety of PWFs to ensure that the results are robust across different specifications.

Shifting to time preferences, under the E model, δ is the discounting parameter which equalises the *utility* of income received at time t (i.e., the utility of the SS reward) with the *utility* of income received at time $t + \tau$ (i.e., the utility of the LL reward):

$$[1 / (1 + \delta)^t]U(y_t) = [1 / (1 + \delta)^{t+\tau}]U(y_{t+\tau}), \quad (2)$$

for some utility function $U(\cdot)$.

Under the assumptions that EU characterises choices over risky prospects and that subjects employ the power utility function, we can add more structure to this indifference condition. Specifically, (2) becomes:

$$[1 / (1 + \delta)^t](y_t)^r = [1 / (1 + \delta)^{t+\tau}](y_{t+\tau})^r, \quad (3)$$

where the general form of the utility function $U(\cdot)$ in (2) has been replaced with the specific power utility function $U(y) = y^r$ in (3).

[18] The "stochastically more risk averse than" relation is the stochastic choice counterpart to the "more risk averse than" relation (see Pratt (1964)) which is defined for the deterministic EU model.

To estimate the parameters of our time preference model, conditional on EU theory, power utility, and the E model, we form a latent index that captures the difference in the present value of the utility of the SS and LL rewards, and we incorporate the behavioural error term originally due to Fechner (1966/1860).

This "joint estimation" approach, developed by Andersen et al. (2008), uses subjects' choices in the risk preference task to pin down the parameters of the utility function, and subjects' choices in the time preference task to pin down the parameters of the E discounting model, conditional on the shape of the utility function. This approach ensures that we estimate time preferences defined over utility flows, and not flows of money.

It is straightforward to incorporate other discounting models in this statistical framework. In the case of Weibull discounting, for instance, (3) becomes:

$$[\exp(-\delta t^{(1/\beta)})](y_t)^r = [\exp(-\delta(t+\tau)^{(1/\beta)})](y_{t+\tau})^r \quad (4)$$

We then form the latent index that captures the difference in the present value of the utility of the SS and LL rewards and proceed as before.

VI. RESULTS

We present the results from a set of risk and time preference models so as to explore potential differences in the risk and time preferences of smokers and non-smokers. We begin with the risk preference results because they provide a natural segue to the time preference results which are conditional on the utility function curvature identified by the risk preference task.

A. Risk Preferences

We estimate an EU model employing a power utility function and the CU behavioural error specification; see Appendix E for more details. We find a relatively high level of risk aversion in the sample; a statistically significant estimate of the behavioural error parameter, implying that subjects make behavioural errors in the risk preference task; and no substantive differences in the risk preferences of smokers and non-smokers. We also estimate a model which allows risk preferences to vary as a quadratic function of smoking intensity as measured by the number of cigarettes smoked per day: risk preferences are not significantly related to smoking intensity. These results are robust to the assumption that Saha's (1993) expo-power utility function – which admits increasing relative risk aversion, decreasing relative risk aversion, and CRRA – characterises choice under risk.

The EU results suggest that there are no significant differences in the risk preferences of smokers and non-smokers. However, this analysis, by assumption, ignored the role of probability weighting and it may be the case that smokers perceive probabilities differently to non-smokers. For example, smokers may underweight moderate to high probabilities more so than non-smokers, and may, therefore, underestimate the likelihood of the negative consequences associated with smoking. To explore this possibility, we estimate RDU models.

One of the key components of a RDU model is the specification of the PWF. We estimate the power PWF, the PWF popularised by Tversky and Kahneman (1992) (TK), and the Prelec (1998) two-parameter PWF which exhibits considerable flexibility; see Appendix D for more details. The functional form for this PWF is:

$$\pi(p) = \exp[-\eta(-\ln p)^{\gamma}], \tag{5}$$

which is defined for $1 > p > 0$, $\eta > 0$, and $\gamma > 0$.[19] This function allows independent specification of location and curvature in probability weighting. It also nests the power PWF when $\gamma = 1$, and nests a one-parameter function when $\eta = 1$, which is similar to the TK function and admits linear, inverse S-shaped, and S-shaped forms.

We find statistically significant evidence of inverse S-shaped probability weighting. This will need to be taken into account when adopting the joint estimation approach to discounting behaviour because the extent of utility function curvature identified by the risk preference task propagates into estimates of discounting parameters. Thus, if one ignores probability weighting when it is present, this would lead to biased estimates of utility function curvature and, hence, biased estimates of discounting

[19] Prelec (1998, proposition 1, part C, p. 503) provides these parameter restrictions. Prelec (1998, proposition 1, part B, p. 503) constrains $1 > \gamma > 0$, but this constraint can be quite restrictive in practice because it limits the PWF to be inverse-S shaped. When estimating the models we impose these constraints using nonlinear transformations of the parameters. To recover the core parameters we use the inverse of these nonlinear transformations, and then apply the "delta method" to derive standard errors and *p*-values for the estimates (see Oehlert (1992)).

parameters. In effect, when probability weighting is present, one ought to apportion risk preferences into their concave utility and probability weighting components so that accurate inferences about discounting behaviour can be drawn.

To investigate the possibility that smokers perceive probabilities differently to non-smokers, even if their utility functions do not differ, we estimate the TK and Prelec (1998) functions, which admit inverse S-shaped PWFs, and allow the parameters to vary as a function of observable characteristics and task parameters. Results are presented in Table 4.[20]

[Table 4 here]

In both of the models in Table 4, smokers do not differ significantly from non-smokers in the shape of their utility functions (i.e., in the estimate of r) nor in the way they perceive probabilities (i.e., in the estimates of γ and η). In addition, tests of the joint hypothesis that the coefficients for smokers across r, γ, and η are equal to zero, cannot be rejected under either model ($p = 0.771$ for the TK model and $p = 0.823$ for the Prelec model).[21] Thus, at least in this sample, there are no significant differences in the risk preferences of smokers and non-smokers. This result is robust to different theories of choice under risk, different PWFs, and a utility function that admits varying relative risk aversion.

B. Time Preferences: Baseline Results

We estimate four time preference models: the E model, the quasi-hyperbolic (QH) model, Mazur's (1984) hyperbolic (H) model, and the Weibull (WB) model; see

[20] The experimental design of the risk preference task lends itself to common-ratio tests of EU theory. To complement the analyses in this section, we conduct a set of common-ratio tests for the lotteries represented in the MM triangles in Figure 2 to determine whether the choice patterns of smokers are more or less EU-consistent than non-smokers. We adopt the non-parametric Cochran Q test and find that both smokers and non-smokers violate EU theory in every MM triangle in Figure 2 ($p < 0.001$ in every test) except the MM triangle with a gradient of 3. In this latter MM triangle, we cannot reject the hypothesis that non-smokers satisfy EU theory ($p = 0.111$) but we can reject this hypothesis for smokers ($p = 0.027$). Thus, in only 1 of the 8 MM triangles of Figure 2 are non-smokers more EU-consistent than smokers. The bulk of the evidence, therefore, suggests little difference in the extent to which smokers and non-smokers violate EU theory; one reaches the same conclusion from the estimates in Table 4.

[21] We also estimate a RDU model with the expo-power utility function, the Prelec PWF, and the full set of covariates from Table 4. The smoker variable is not significantly different from zero for any of the parameters in the model. In addition, a test of the joint hypothesis that the coefficients for smokers across r, α, γ, and η are equal to zero, cannot be rejected ($p = 0.967$).

Appendix F for more details and Andersen et al. (2014) for a review all of the major discounting models. We jointly estimate the parameters of these models with the curvature of the utility function to focus on the discounting of utility flows. In the context of addiction, the crucial difference between these specifications is that, under the assumption of an additively-separable intertemporal utility function, the E model implies time-consistent preferences whereas the other models can yield time-inconsistent preferences.

Table 5 presents results from the four discounting models employing a Fechner error term, assuming RDU[22] and the Prelec (1998) PWF characterise choice under risk, and using years, rather days, as the unit of measurement for the estimation of the parameters.

The estimate of the E discount rate in Table 5 is 0.493, which is far below the estimate of the E discount rate (i.e., 3.234) under the assumption of linear utility (see Appendix F). Similar declines are evident across all discounting models, which highlights the point, now familiar from Andersen et al. (2008), that incorporating concavity of the utility function leads to substantial declines in inferred discount rates.

[Table 5 here]

In the QH model, the estimate of $\beta = 0.988$, which captures a "present-bias" or a "passion for the present" in discounting behaviour, is statistically significantly less than 1 ($p = 0.002$), which provides evidence of quasi-hyperbolic discounting and declining discount rates. The same is true in the WB results: the estimate of $\beta = 1.611$, which "expands" or "contracts" time, is statistically significantly greater than 1 ($p < 0.001$) which leads us to infer that people perceive time as "slowing down," generating declining discount rates. Thus, both the QH and WB results suggest that discount rates decline over time, which, when coupled with an additively-separable intertemporal utility function, raises the spectre of time-inconsistent choices.

[22] Given the presence of probability weighting in this dataset, we employ RDU theory to apportion risk preferences into their concave utility and probability weighting components so as to draw accurate inferences about discounting behaviour. In Appendix F we test the robustness of our results by estimating these models assuming EU theory characterises choice under risk; the results are qualitatively identical to those reported in the main text.

However, the two discounting functions provide competing explanations for this result: a present-bias in the case of the QH model and subjective time perception in the case of the WB model.

C. Smoking and Discounting Behaviour

As a descriptive prelude to the formal statistical results, Figure 4 shows a kernel-weighted local polynomial regression, with a 95% confidence interval, of the fraction of LL choices by smokers and non-smokers for the nominal annual interest rates on offer in the time preference task. At each interest rate, the point estimate of the fraction of LL choices by smokers is less than the point estimate of the fraction of LL choices by non-smokers, and the 95% confidence intervals do not overlap. This suggests that smokers discount more heavily than non-smokers, but clearly this result must be subjected to closer statistical scrutiny before any definitive conclusions are reached.

[Figure 4 here]

Table 6 presents results from the four time preference models, assuming RDU and the Prelec PWF, where risk and discounting parameters are allowed to vary by smoking status. These models, therefore, capture the "total effect" of smoking on discounting behaviour without controlling for any potential differences between smokers and non-smokers such as age, gender, etc. Across all specifications, the effect of smoking on the estimate of δ is positive and statistically significant, implying that smokers tend to discount the future more heavily than non-smokers. The magnitude of this difference in discounting behaviour is economically significant. In the E model, for example, smokers have an annual discount rate which is 20 percentage points higher than non-smokers. Thus, the positive relationship between smoking and discounting identified in Table 1 has been replicated using a joint estimation approach to time preferences which controls for utility function curvature and probability weighting.[23]

The estimates of β in the QH and WB models, by contrast, do not vary according to smoking status. Thus, smokers are no more present-biased than non-smokers in the

[23] Appendix F presents results from the four time preference models where EU theory is assumed to characterise choice under risk: the results are virtually identical to the models in Table 6.

QH model nor are they more likely to perceive time as slowing down in the WB model. It is only the long-term discount rate δ which differs between smokers and non-smokers in these models.

[Table 6 here]

To investigate whether smoking intensity and discounting behaviour are related, we estimate the four time preference models and allow the parameters of interest to vary as a quadratic function of the number of cigarettes smoked per day. In all models, both the linear and quadratic terms are statistically significant in the estimate of δ: the linear term is positive and significant whereas the quadratic term is negative and significant. Thus, there is a concave relationship between discounting behaviour and number of cigarettes smoked per day: every additional cigarette is associated with an increase in discounting, but at a decreasing rate until a maximum is reached, after which every additional cigarette is associated with a decrease in discounting.[24]

Table 7 maps out the response surface for estimates of δ in the four time preference models evaluated at different values of number of cigarettes smoked per day. At low values of number of cigarettes, the conditional marginal effect of additional cigarettes is positive. By 15 cigarettes, though, the conditional marginal effect of additional cigarettes is negative. Table 7 highlights the nonlinear effect of smoking intensity on discounting behaviour. To our knowledge, this is the first study of time preferences and smoking behaviour which has identified this effect.

[Table 7 here]

The preceding results are only preliminary because we must control for a number of factors which may mediate the relationship between smoking and discounting behaviour. Thus, we estimate the four time preference models while taking into account observed individual heterogeneity by conditioning the discounting and risk preference parameter estimates on the set of demographic characteristics and task

[24] In the QH model, smoking intensity is not significantly related to the extent of present-bias. In the WB model, though, the number of cigarettes' linear term is negative and significant in the estimate of β, albeit at the 10% level. Thus, the more cigarettes smoked per day, the less likely people are to perceive time as slowing down.

parameters from Table 3. These models then capture the "marginal effect" of smoking status, and the results are presented in Appendix G. Across all of the models, the estimate of δ for smokers is positive and statistically significant at the 1% level. By contrast, in the QH and WB models the estimate of β for smokers is not statistically significant.

In sum, we observe a positive relationship between smoking and discounting behaviour which holds across all of the time preference models and all of the model specifications we estimate. This result is also robust to the assumption that EU characterises choice under risk (see Appendix F). However, smokers do not differ from non-smokers with regard to present-bias in the QH model nor in terms of time perception in the WB model.

D. Mixture Models of Discounting Behaviour

The analyses conducted thus far have been based on the implicit assumption that the observations are produced by only one discounting data generating process (DGP): either E, H, QH, or WB. However, the data may be a result of more than one DGP. For example, the E model may explain some discounting choices better than the H model whereas the H model may explain other choices better than the E model. The assumption that only one DGP characterises all of the data precludes such a possibility.[25]

Finite mixture models[26] allow two or more DGPs to account for the data and also provide a measure of the proportion of the data which is best explained by each

[25] One can use Vuong (1989) and Clarke (2007) non-nested model selection tests to formally determine which discounting function, in a pairwise comparison, finds more support in the dataset as a whole. The choice between these tests is based on the distribution of the individual log-ratios of the models. When the distribution of these log-ratios is leptokurtic, as we find in our data, the Clarke (2007) test is superior, from both statistical efficiency and power perspectives, to the Vuong (1989) test. Clarke (2007) tests of the four discounting models provided the following transitive ranking: WB > H > E > QH. In words, the WB model finds the most support in the data, the QH model finds the least support, and the E and H models are intermediate to these. Thus, if one had to select a single model that best characterises the time preferences of this sample, the WB model would be the choice. However, when multiple time preference processes are present in a dataset, one should be cognisant of this fact and estimate a mixture model to determine the proportion of choices best explained by each process. If one discounting model is truly superior to another, in the sense that it better explains all the data, then this will be reflected in a mixture probability estimate of zero for the inferior model.

[26] For detailed discussions of mixture models see McLachlan and Peel (2000), Harrison and Rutström (2009), and Conte, Hey and Moffatt (2011). Mixture models have been applied to discounting behaviour by Andersen et al. (2008), Coller, Harrison and Rutström (2012) and Andersen et al. (2014).

process. In the current context, one can estimate a mixture model of, say, the E and H discounting functions and then ask the data to determine each function's level of support. To do so one specifies a "grand likelihood" function which is just a probability-weighted average of the likelihoods of the two models.

Letting π^E represent the probability that the E model is correct, and $\pi^H = (1 - \pi^E)$ the probability that the H model is correct, the grand likelihood is the probability-weighted average of the two conditional likelihoods L^E and L^H for the E and H models, respectively. Thus, the likelihood for the mixture model is given by:

$$\ln L_i(r, \gamma, \eta, \delta_E, \delta_H, \mu, \nu, \kappa; z, X) = \Sigma_i \ln [(\pi^E \times L^E) + (\pi^H \times L^H)], \quad (6)$$

where κ is a parameter which defines the log odds of the probability of the E model: $\pi^E = 1 / (1 + \exp(\kappa))$. This transformation allows the parameter κ to take on any value during the maximisation process but constrains the probabilities π^E and π^H to lie within the unit interval. The grand likelihood in (6) is maximised to estimate the parameters of each model and the weight accorded to each model in the data, under the assumptions that RDU and the Prelec PWF characterise choice under risk.

[Table 8 here]

Table 8 presents estimates of the mixture model of the E and H discounting functions assuming homogenous preferences.[27] The estimate of $\pi^E = 0.347$ implies that the E model accounts for approximately 35% of the choices in the data; the H model therefore accounts for roughly 65% of the choices. A hypothesis test that $\pi^E = 0.5$ is easily rejected ($p < 0.001$) but so too is the hypothesis that $\pi^E = 0$ ($p < 0.001$).[28] Thus the E and H discounting models both find significant support in the data, even though the H model finds more support. Consequently it is a mistake to assume that only one DGP characterises the data.

[27] Appendix H contains the results from all of the two-process mixture models that can be estimated from the four discounting specifications. We only present the results from the E and H mixture model in this section because these are the most commonly used discounting functions in the addiction literature and they are representative of the results from the other mixture models.

[28] In all of the mixture models in Appendix H, one of the discounting functions explains significantly more of the choices than the other discounting function. However, all discounting functions find significant support in the data which reinforces the point that it is a mistake to assume only one DGP characterises all discounting choices all of the time. The mixture probability estimates for the other mixture models are: E-QH model - $\pi^E = 0.636$; E-WB model - $\pi^E = 0.406$; H-QH model - $\pi^H = 0.634$; H-WB model - $\pi^H = 0.611$; QH-WB model - $\pi^{QH} = 0.609$.

The mixture model in Table 8 also shows how discounting parameter estimates are distorted when the E or H models have to account for all the data. Figure 5 plots the discount factors from the E and H models in Table 5 (i.e., when they are assumed to account for all the data) and the discount factors from the mixture model in Table 8.

[Figure 5 here]

In Model 1 of Table 5, where the E model was assumed to be the sole DGP, the estimate of $\delta_E = 0.493$. In the mixture model, the estimate of δ_E, which we refer to as δ_E^{mix} in Table 8 and Figure 5, is statistically significantly lower at 0.137 ($p < 0.001$). This implies that when one tries to make all the data fit the E model, one inflates the estimate of the discounting parameter since 65% of the data "wants" to be modelled as H. Similarly, in Model 2 of Table 5, where the H model was assumed to be the sole DGP, the estimate of $\delta_H = 0.502$. In the mixture model in Table 8, the estimate of δ_H^{mix} is statistically significantly higher at 0.730 ($p < 0.001$). Thus, by assuming one DGP we are averaging the estimates that we derive when allowing multiple DGPs to characterise the data.[29]

Finally, the estimate of the Fechner error term $\nu = 0.051$ in the mixture model in Table 8 is statistically significantly lower than the estimates of ν for the E ($p < 0.001$) and H models ($p < 0.001$) in Table 5. Thus, what was being captured as subject errors in decision making when estimating the E and H models separately is partly the product of forcing the data to fit one DGP.

Mixture models also allow one to explore the hypothesis that smokers are more likely to discount hyperbolically than non-smokers by making the mixture probability a function of smoking status. Recall that with an additively-separable intertemporal utility function the E model implies time-consistent preferences whereas the H model

[29] Harrison and Rutström (2009) reach a similar conclusion in the context of choice under risk. Specifically, under the assumption that prospect theory (PT) is the sole DGP, Harrison and Rutström (2009, p. 146) find limited evidence of loss aversion, no significant evidence of probability weighting, and utility function estimates for the gain and loss frames which are not significantly different from one another. In a mixture model of EU and PT, by contrast, they find substantial evidence of loss aversion, significant probability weighting, and utility function estimates which differ significantly across the gain and loss frames.

may yield time-inconsistent choices. If one finds that current smokers are more likely to discount according to the H model as opposed to the E model, this suggests that they may be more prone to time inconsistency.

Table 9 presents estimates of the mixture model of the E and H discounting functions where the parameters of interest are allowed to vary by smoking status. Of particular interest is that smokers are significantly less likely to discount exponentially, which means here that they are significantly more likely to discount hyperbolically, than non-smokers. The magnitude of this result is economically significant since smokers are 15% more likely to discount hyperbolically than non-smokers. Consequently, under the assumption of an additively-separable intertemporal utility function, smokers may be more likely to make time-inconsistent choices than non-smokers. However, when we include the full set of demographic characteristics and task parameters in the mixture probability equation, the smoker covariate's point estimate of -0.141 does not quite reach conventional levels of statistical significance (p = 0.116).

[Table 9 here]

In sum, the results from the mixture model analyses suggest that it is a mistake to assume that only one model of discounting behaviour accurately characterises all discounting choices by all subjects. Recognising that some choices are better explained by one specification while other choices are better explained by another specification allows one to draw more accurate inferences about the type and extent of discounting behaviour. Finally, our results suggest that smokers may be more likely to discount hyperbolically than exponentially, but conclusive evidence of this assertion is still lacking.

VII. DISCUSSION AND CONCLUSIONS

We analyse the relationship between risk preferences, time preferences and smoking behaviour using an incentive-compatible experimental design and a joint estimation approach to data analysis. We find that both probability weighting and utility function curvature affect attitudes to risk in this sample but we find no statistically significant relationship between risk preferences and smoking status. This result is robust to different theories of choice under risk, different PWFs, and different utility functions which admit varying relative risk aversion.

To analyse the time preferences of smokers and non-smokers we adopt the methodology of HLR which jointly estimates utility function curvature and discounting functions so as to characterise time preferences over utility flows, not flows of money. We find that controlling for the concavity of the utility function leads to a dramatic decline in estimates of δ, replicating the result of Andersen et al. (2008). We also allow RDU to characterise choice under risk so as to apportion risk preferences into their utility curvature and probability weighting components.

We explore the relationship between time preferences and smoking behaviour in three ways. First, a smoking status covariate is added to the discounting models to capture the "total effect" of smoking on discounting behaviour without controlling for other factors such as age, gender, etc. Across every discounting model, the estimate of δ for smokers is positive and statistically significant, implying smokers discount more heavily than non-smokers. However, smoking status is not related to the extent of present-bias in the QH model nor in the perception of time in the WB model.

Second, to investigate whether smoking intensity is related to discounting behaviour, we estimate four time preference models and allow the parameters of interest to vary as a quadratic function of number of cigarettes smoked per day. These analyses reveal a concave relationship between smoking intensity and estimates of the discounting parameter δ. Specifically, every additional cigarette is associated with an increase in discounting, but at a decreasing rate until a maximum is reached, after which every additional cigarette is associated with a decrease in discounting.

Finally, to explore the marginal effect of smoking status on time preferences, we estimate the discounting models and make the parameters of interest a linear function of observable characteristics and task parameters. Across all specifications, the estimate of δ for smokers is positive and statistically significant, which replicates the earlier result while controlling for other variables which may mediate the relationship between smoking and discounting. In Appendix F we also test to see whether these results are robust to the assumption that EU characterises choice under risk: the results are qualitatively identical to those in Section VI.

In the final set of statistical analyses, we estimate mixture models of the different discounting specifications, assuming that RDU and the Prelec PWF capture choice under risk. These mixture models show that multiple decision processes characterise the discounting of delayed rewards and that it is a mistake to force all discounting choices to fit one particular model. In addition, the mixture models allow us to explore the hypothesis that smokers are more likely to discount hyperbolically than non-smokers by making the mixture probability a function of smoking status. We find that smokers are significantly more likely than non-smokers to discount hyperbolically, but this result does not attain statistical significance when the full set of demographic characteristics and task parameters are added to the mixture probability.

This research makes a number of contributions to the literature. When analysing risk preferences and smoking behaviour, we allow risk attitudes to be determined both by utility function curvature and probability weighting. Prior studies in the literature either focus on utility function curvature or probability weighting, not both. Consequently they are always open to the critique that the other source of risk attitudes, the one not explored in the study, differs between smokers and non-smokers. Incorporating both utility function curvature and probability weighting in estimates of risk attitudes provides us with immunity to this critique and allows us to make stronger claims about differences in the risk preferences of smokers and non-smokers.

This is only the second study in the smoking-discounting literature to incorporate utility function curvature in the estimation of time preference models, and it is the first which allows RDU to characterise choice under risk. Although the qualitative discounting estimates do not differ significantly across the EU and RDU specifications, it is nevertheless theoretically appropriate to quantitatively apportion risk preferences into their utility curvature and probability weighting components.

This is the first study to identify a nonlinear effect of smoking intensity on discounting behaviour. Smoking more cigarettes is associated with an increase discounting but only up to a point, after which each additional cigarette is associated with lower discounting. This nonlinear effect may explain why some studies, which only recruited heavy smokers and never-smokers, fail to find a difference in discounting behaviour between these groups.

In addition, this nonlinear effect of smoking intensity may provide an explanation for patterns of cigarette consumption. It has long been assumed that the marked modal clustering around 20 cigarettes per day in mature smokers simply reflects the fact that cigarettes are typically sold in packs of 20. It may be the case, though, that cigarette companies learned to sell cigarettes in packs of 20 because that is where the psychofunctional, and not merely the homeostatic, equilibrium lies for the majority of mature smokers.

This research also reiterates the point that multiple decision processes characterise the discounting of delayed rewards. It is crucial for researchers to be cognisant of this fact when exploring the smoking-discounting relationship. Smokers may be more likely than non-smokers to discount hyperbolically and this may be a factor in addiction. To our knowledge, this is the first study in the literature to identify this difference between a set of smokers and non-smokers.

This research also suffers from a number of limitations. Clearly a young, university sample of smokers is not representative of smokers in general, so the external validity of these results for the South African population as a whole is questionable. This study is an important first step, however, towards population-based studies of risk preferences, time preferences and smoking behaviour. The tasks and instructions developed for this study could be used to elicit a representative sample of South Africans' attitudes toward risk and time, which can then be analysed using the tools developed here.

Another important issue is whether risk and time preferences are domain- or context-specific. The finding that subjects' choices over intertemporal payments are associated with smoking behaviour suggests that a time preference experiment over money could be used as a screening tool for people who may be at risk for addiction. But the lack of a relationship between risk preferences over lottery choices and smoking behaviour may indicate the need to elicit these preferences (and time preferences) in the health domain too.

A perennial question one can ask about economic experiments is whether the rewards on offer are sufficiently salient to incentivise truthful revelation of preferences over an income domain which is relevant to policy analysis. The rewards in this study are large in comparison to those typically paid out in the literature (see appendices A and B), but still larger incentives would help to alleviate concerns about the extent to which these results scale up as the magnitude of the prizes increases.

Another potential issue with the sample is the extent of possible selection bias. As discussed earlier, a large number of people applied to take part in the study, so people in the smoking and non-smoking groups were randomly selected to form the study pool. It may be the case that those who were selected were not representative of their group. Ideally we would use information on the population of smokers and non-smokers at UCT to correct for any sample selection issues present in the data.[30] Unfortunately, we do not have any additional information on the population of smokers and non-smokers at UCT.

These issues notwithstanding, we provide a rigorous framework within which to analyse risk and time preferences and smoking behaviour. Future experimental research should seek to replicate the nonlinear effect of smoking intensity on discounting behaviour identified in this study and the presumptive evidence of differences between smokers and non-smokers in the likelihood of making time-inconsistent choices. If these results hold in other samples, our understanding of smoking specifically, and addiction generally, will be sharpened.

[30] Harrison, Lau and Rutström (2009) and Harrison and Lau (2014) analyse the effect of sample selection bias on estimated risk preference parameters. They used the Danish Registry to gather information on people who were invited to participate in their experiment but who did not take part and this allowed them to make sample selection corrections for the sample of people who were invited and who did participate in the experiment. Harrison, Lau and Rutström (2009) find that correcting for sample selection bias leads to attenuated risk aversion estimates, implying that their sample was more risk averse than the population from which it was drawn. Similarly, Harrison and Lau (2014) find that sample selection corrections lead to lower estimates of risk aversion.

REFERENCES

AINSLIE, G. (1992): *Picoeconomics: The Interaction of Successive Motivational States within the Individual*. Cambridge: Cambridge University Press.

AINSLIE, G. (2001): *Breakdown of Will*. Cambridge: Cambridge University Press.

ANDERSEN, S., G. W. HARRISON, M. I. LAU, AND E. E. RUTSTRÖM (2008): "Eliciting Risk and Time Preferences," *Econometrica*, 76, 583-618.

ANDERSEN, S., G. W. HARRISON, M. I. LAU, AND E. E. RUTSTRÖM (2010): "Behavioural Econometrics for Psychologists," *Journal of Economic Psychology*, 31, 553-576.

ANDERSEN, S., G. W. HARRISON, M. I. LAU, AND E. E. RUTSTRÖM (2014): "Discounting Behavior: A Reconsideration," *European Economic Review*, 71, 15-33.

BECKER, G. S., AND K. M. MURPHY (1988): "A Theory of Rational Addiction," *Journal of Political Economy*, 96, 675-700.

BÉNABOU, R., AND J. TIROLE (2004): "Willpower and Personal Rules," *Journal of Political Economy*, 112, 848-886.

BENHABIB, J., AND A. BISIN (2004): "Modeling Internal Commitment Mechanisms and Self-Control: A Neuroeconomics Approach to Consumption-Saving Decisions," *Games and Economic Behavior*, 52, 460-492.

BERNHEIM, B. D., AND A. RANGEL (2004): "Addiction and Cue-Triggered Decision Processes," *American Economic Review*, 94, 1558 - 1590.

BICKEL, W. K., A. L. ODUM, AND G. J. MADDEN (1999): "Impulsivity and Cigarette Smoking: Delay Discounting in Current, Never, and Ex-Smokers," *Psychopharmacology*, 146, 447-54.

BRADFORD, W. D. (2010): "The Association between Individual Time Preferences and Health Maintenance Habits," *Medical Decision Making*, 30, 99-112.

CLARK, A. (1997): *Being There*. Cambridge, MA: MIT Press.

CLARKE, K. A. (2007): "A Simple Distribution-Free Test for Nonnested Model Selection," *Political Analysis*, 15, 347-363.

COLLER, M., G. W. HARRISON, AND M. M. MCINNES (2002): "Evaluating the Tobacco Settlement Damage Awards: Too Much or Not Enough?," *American Journal of Public Health*, 92, 1-6.

COLLER, M., G. W. HARRISON, AND E. E. RUTSTRÖM (2012): "Latent Process Heterogeneity in Discounting Behavior," *Oxford Economic Papers*, 64, 375-391.

COLLER, M., AND M. B. WILLIAMS (1999): "Eliciting Individual Discount Rates," *Experimental Economics*, 2, 107-127.

CONTE, A., J. D. HEY, AND P. G. MOFFAT (2011): "Mixture Models of Choice under Risk," *Journal of Econometrics*, 162, 79-88.

DALLERY, J., AND B. R. RAIFF (2007): "Delay Discounting Predicts Cigarette Smoking in a Laboratory Model of Abstinence Reinforcement," *Psychopharmacology*, 190, 485-96.

EPSTEIN, L. H., J. B. RICHARDS, F. G. SAAD, R. A. PALUCH, J. N. ROEMMICH, AND C. LERMAN (2003): "Comparison between Two Measures of Delay Discounting in Smokers," *Experimental and Clinical Psychopharmacology*, 11, 131-8.

ERIKSEN, M., J. MACKAY, N. SCHLUGER, F. I. GOMESHTAPEH, AND J. DROPE (2015): *The Tobacco Atlas*. Atlanta, GA: American Cancer Society.

FECHNER, G. T. (1966/1860): *Elements of Psychophysics*. New York, NY: Holt, Rinehart and Winston.

FIELD, M., M. SANTARCANGELO, H. SUMNALL, A. GOUDIE, AND J. COLE (2006): "Delay Discounting and the Behavioural Economics of Cigarette Purchases in Smokers: The Effects of Nicotine Deprivation," *Psychopharmacology*, 186, 255-63.

FUDENBERG, D., AND D. K. LEVINE (2006): "A Dual-Self Model of Impulse Control," *American Economic Review*, 96, 1449-1476.

FUDENBERG, D., AND D. K. LEVINE (2011): "Risk, Delay and Convex Self-Control Costs," *American Economic Journal: Microeconomics*, 3, 34-68.

FUDENBERG, D., AND D. K. LEVINE (2012): "Timing and Self-Control," *Econometrica*, 80, 1-42.

GLIMCHER, P. W. (2011): *Foundations of Neuroeconomic Analysis*. Oxford: Oxford University Press.

GRUBER, J., AND B. KÖSZEGI (2001): "Is Addiction "Rational"? Theory and Evidence," *Quarterly Journal of Economics*, 116, 1261 - 1303.

GUL, F., AND W. PESENDORFER (2007): "Harmful Addiction," *The Review of Economic Studies*, 74, 147-172.

HARRISON, G. W., AND M. I. LAU (2014): "Risk Attitudes, Sample Selection and Attrition in a Longitudinal Field Experiment," *Working Paper* 2014-04, Center for the Economic Analysis of Risk (CEAR), Georgia State University.

HARRISON, G. W., M. I. LAU, AND E. E. RUTSTRÖM (2009): "Risk Attitudes, Randomization to Treatment, and Self-Selection into Experiments," *Journal of Economic Behavior and Organization*, 70, 498-507.

HARRISON, G. W., M. I. LAU, AND E. E. RUTSTRÖM (2010): "Individual Discount Rates and Smoking: Evidence from a Field Experiment in Denmark," *Journal of Health Economics*, 29, 708-717.

HARRISON, G. W., AND E. E. RUTSTRÖM (2009): "Expected Utility Theory and Prospect Theory: One Wedding and a Decent Funeral," *Experimental Economics*, 12, 133-158.

HEATHERTON, T. F., L. T. KOZLOWSKI, R. C. FRECKER, AND K. FAGERSTRÖM (1991): "The Fagerström Test for Nicotine Dependence: A Revision of the Fagerström Tolerance Questionnaire," *British Journal of Addiction*, 86, 1119-1127.

HEY, J. D., AND C. ORME (1994): "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica*, 62, 1291-1326.

HEYMAN, G. M. (2009): *Addiction: A Disorder of Choice*. Cambridge, MA: Harvard University Press.

HOFMEYR, A., J. MONTEROSSO, A. C. DEAN, A. M. MORALES, R. M. BILDER, F. W. SABB, AND E. D. LONDON (2016): "Mixture Models of Delay Discounting and Smoking Behavior," *The American Journal of Drug and Alcohol Abuse, forthcoming*.

JONES, B. A., R. D. LANDES, R. YI, AND W. K. BICKEL (2009): "Temporal Horizon: Modulation by Smoking Status and Gender," *Drug and Alcohol Dependence*, 104 Supplement 1, S87-93.

JUSOT, F., AND M. KHLAT (2013): "The Role of Time and Risk Preferences in Smoking Inequalities: A Population-Based Study," *Addictive Behaviors*, 38, 2167-73.

LAIBSON, D. (1997): "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112, 443-477.

LAIBSON, D. (2001): "A Cue-Theory of Consumption," *Quarterly Journal of Economics*, 116, 81-119.

LAWYER, S. R., F. SCHOEPFLIN, R. GREEN, AND C. JENKS (2011): "Discounting of Hypothetical and Potentially Real Outcomes in Nicotine-Dependent and Nondependent Samples," *Experimental and Clinical Psychopharmacology*, 19, 263-74.

LOEWENSTEIN, G., T. O'DONOGHUE, AND M. RABIN (2003): "Projection Bias in Predicting Future Utility," *Quarterly Journal of Economics*, 118, 1209-1248.

LOOMES, G., AND R. SUGDEN (1998): "Testing Different Stochastic Specifications of Risky Choice," *Economica*, 65, 581-598.

MAZUR, J. E. (1984): "Tests of an Equivalence Rule for Fixed and Variable Reinforcer Delays," *Journal of Experimental Psychology: Animal Behavior Processes*, 10, 426-437.

MCLACHLAN, G., AND D. PEEL (2000): *Finite Mixture Models*. New York, NY: Wiley.

MYERSON, J., L. GREEN, AND M. WARUSAWITHARANA (2001): "Area under the Curve as a Measure of Discounting," *Journal of the Experimental Analysis of Behavior*, 76, 235-43.

ODUM, A. L., AND A. A. BAUMANN (2007): "Cigarette Smokers Show Steeper Discounting of Both Food and Cigarettes Than Money," *Drug and Alcohol Dependence*, 91, 293-6.

ODUM, A. L., G. J. MADDEN, AND W. K. BICKEL (2002): "Discounting of Delayed Health Gains and Losses by Current, Never- and Ex-Smokers of Cigarettes," *Nicotine & Tobacco Research*, 4, 295-303.

OEHLERT, G. W. (1992): "A Note on the Delta Method," *The American Statistician*, 46, 27-29.

ORPHANIDES, A., AND D. ZERVOS (1995): "Rational Addiction with Learning and Regret," *Journal of Political Economy*, 103, 739-758.

POLTAVSKI, D. V., AND J. N. WEATHERLY (2013): "Delay and Probability Discounting of Multiple Commodities in Smokers and Never-Smokers Using Multiple-Choice Tasks," *Behavioural Pharmacology*, 24, 659-67.

PRATT, J. W. (1964): "Risk Aversion in the Small and in the Large," *Econometrica*, 32, 122-136.

PRELEC, D. (1998): "The Probability Weighting Function," *Econometrica*, 66, 497-527.

QUIGGIN, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3, 323-343.

REDISH, A. D., S. JENSEN, AND A. JOHNSON (2008): "A Unified Framework for Addiction: Vulnerabilities in the Decision Process," *Behavioral and Brain Sciences*, 31, 415-487.

REYNOLDS, B., K. KARRAKER, K. HORN, AND J. B. RICHARDS (2003): "Delay and Probability Discounting as Related to Different Stages of Adolescent Smoking and Non-Smoking," *Behavioural Processes*, 64, 333-344.

ROSS, D. (2010): "Economic Models of Pathological Gambling," in *What Is Addiction?*, ed. by D. Ross, H. Kincaid, D. Spurrett, and P. Collins. Cambridge, MA: MIT Press, 131-158.

ROSS, D. (2011): "Estranged Parents and a Schizophrenic Child: Choice in Economics, Psychology and Neuroeconomics," *Journal of Economic Methodology*, 18, 215-229.

ROSS, D. (2014a): *Philosophy of Economics*. Basingstoke: Palgrave Macmillan.

ROSS, D. (2014b): "Psychological Versus Economic Models of Bounded Rationality," *Journal of Economic Methodology*, 21, 411-427.

ROSS, D., C. SHARP, R. VUCHINICH, AND D. SPURRETT (2008): *Midbrain Mutiny: The Behavioural Economics and Neuroeconomics of Disordered Gambling*. Cambridge, MA: MIT Press.

SAHA, A. (1993): "Expo-Power Utility: A 'Flexible' Form for Absolute and Relative Risk Aversion," *American Journal of Agricultural Economics*, 75, 905-913.

SCHELLING, T. C. (1984): *Choice and Consequence*. Cambridge, MA: Harvard Univerity Press.

SHISANA, O., D. LABADARIOS, T. REHLE, L. SIMBAYI, K. ZUMA, A. DHANSAY, P. REDDY, W. PARKER, E. HOOSAIN, P. NAIDOO, C. HONGORO, Z. MCHIZA, N. P. STEYN, N. DWANE, M. MAKOAE, T. MALULEKE, S. RAMLAGAN, S. ZUNGU, M. G. EVANS, L. JACOBS, M. FABER, AND SANHANES-1 TEAM (2013): *South African National Health and Nutrition Examination Survey (Sanhanes-1)*. Cape Town: HSRC Press.

STILLWELL, D. J., AND R. J. TUNNEY (2012): "Effects of Measurement Methods on the Relationship between Smoking and Delay Reward Discounting," *Addiction*, 107, 1003-12.

THALER, R. H., AND H. M. SHEFRIN (1981): "An Economic Theory of Self-Control," *Journal of Political Economy*, 89, 392-406.

TVERSKY, A., AND D. KAHNEMAN (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297-323.

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES (2014): *The Health Consequences of Smoking: 50 Years of Progress. A Report of the Surgeon General*. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

VAN DER POL, M., AND M. RUGGERI (2008): "Is Risk Attitude Outcome Specific within the Health Domain?," *Journal of Health Economics*, 27, 706-17.

VUONG, Q. H. (1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57, 307-333.

WEST, R. (2006): *Theory of Addiction*. Oxford: Blackwell Publishing.

WILCOX, N. T. (2011): "'Stochastically More Risk Averse:' a Contextual Theory of Stochastic Discrete Choice under Risk," *Journal of Econometrics*, 162, 89-104.

WINSTON, G. C. (1980): "Addiction and Backsliding: A Theory of Compulsive Consumption," *Journal of Economic Behavior and Organization*, 1, 295-324.

YAARI, M. E. (1987): "The Dual Theory of Choice under Risk," *Econometrica*, 55, 95-115.

TABLE 1: REVIEW OF EXPERIMENTAL LITERATURE ON SMOKING AND DISCOUNTING BEHAVIOUR

Study	Sample (size)	Elicitation method	Task-related incentives (max LL)	Horizon	Front end delay (FED)	Correct for non-linear utility	Models (estimated rates)	Statistical method	Hyperbolicky discounting?	Significant relationship with smoking?
Bickel, Odum and Madden (1999)	Adults in Burlington, VT, USA (N_S = 23, N_{NS} = 22, N_{ES} = 21)	Choice (ordered)	No ($1000)	7 - 9131 days	No	No	H and E (δ^H_S = 0.054) (δ^H_{NS} = 0.007) (δ^H_{ES} = 0.007)	NLLS for discounting, ANOVA and non-parametric tests for analysis	Yes (compared to E) based on R^2 comparisons	Yes, positive for S relative to NS ($p<0.01$) and ES ($p<0.01$); No for NS relative to ES.
Mitchell (1999)	Adults in Durham, NH, USA (N_S = 20, N_{NS} = 20)	Choice (random)	Yes ($10)	0 - 365 days	No	No	H (δ_S = 0.012) (δ_{NS} = 0.006)	NLLS for discounting, non-parametric tests for analysis	By assumption	Yes ($p<0.06$), positive.
Baker, Johnson and Bickel (2003)	Adults in Burlington, VT, USA (N_S = 30, N_{NS} = 30)	Titration (random - Richards et al. (1999))	Yes ($100)	Real: 1 - 183 days	No	No	H NRD but from Figure 2: ($10: δ_S = 0.008, δ_{NS} = 0.001) ($100: δ_S = 0.005, δ_{NS} = 0.001)	NLLS for discounting, ANOVA for analysis	By assumption	Real: Yes ($p<0.01$), positive.
			No ($1000)	Hypothetical: 1 - 9131 days	No	No	H NRD but from Figure 2: ($10: δ_S = 0.008, δ_{NS} = 0.003) ($100: δ_S = 0.006, δ_{NS} = 0.0008) ($1000: δ_S = 0.004, δ_{NS} = 0.0005)			Hypothetical: Yes ($p<0.01$), positive.
Reynolds, Karraker, Horn and Richards (2003)	Adolescents in Morgantown, WV, USA (N_S = 19, N_{NS} = 19, N_T = 17)	Titration (random - Richards et al. (1999))	1-out-of-2-tasks ($10)	1 - 365 days	No	No	H (δ_S = 0.010) (δ_{NS} = 0.007) (δ_T = 0.016)	NLLS for discounting, ANOVA for analysis	By assumption	No.
Reynolds (2004)	Adolescents and young adults in Morgantown, WV, USA ($N_{S(adolescent)}$ = 19, $N_{S(adult)}$ = 25, N_{NS} = 29)	Titration (random - Richards et al. (1999))	1-out-of-2-tasks ($10)	1 - 365 days	No	No	H ($\delta_{S(adolescent)}$ = 0.016) ($\delta_{S(adult)}$ = 0.075) ($\delta_{NS(adult)}$ = 0.012)	NLLS for discounting, ANOVA, correlations and post hoc tests for analysis	By assumption	Yes, positive for $S_{(adult)}$ relative to $S_{(adolescent)}$ ($p<0.05$) and $NS_{(adult)}$ ($p<0.05$). No for $S_{(adolescent)}$ relative to $NS_{(adult)}$.
Reynolds, Richards, Horn and Karraker (2004)	Mostly students in Morgantown, WV, USA (N_S = 25, N_{NS} = 29)	Titration (random - Richards et al. (1999))	1-out-of-2-tasks ($10)	1 - 365 days	No	No	H (δ_S = 0.066) (δ_{NS} = 0.015)	NLLS for discounting, ANOVA for analysis	By assumption	Yes ($p<0.05$), positive.
Ohmura, Takahashi and Kitamura (2005)	Students in Sapporo, Japan (N_S = 27, N_{NS} = 23)	Titration (random - Richards et al. (1999))	No (¥100,000 = $1000)	7 - 1826 days)	No	No	H, E and AUC. (AUC_S = 0.54) (AUC_{NS} = 0.58)	AUC and NLLS for discounting, correlations and t tests for analysis	Yes (compared to E) based on R^2 comparisons	No.

Source: Authors' construction. See Appendix A for more details on these studies.

Notes: S = smoker; NS = non-smoker/never-smoker; ES = ex-smoker; LS = light smoker; T = trier; FS = fast smoking adopter; SS = slow smoking progressor.

H = hyperbolic; E = exponential; QH = quasi-hyperbolic; AUC = area under the curve; NRD = not reported directly; [a] = annual rate; [b] = weekly rate; NLLS = non-linear least squares; ML = maximum likelihood.

ANOVA = analysis of variance; ANCOVA = analysis of covariance

TABLE 1: REVIEW OF EXPERIMENTAL LITERATURE ON SMOKING AND DISCOUNTING BEHAVIOUR (CONTINUED)

Study	Sample (size)	Elicitation method	Task-related incentives (max LL)	Horizon	Front end delay (FED)	Correct for non-linear utility	Models (estimated rates)	Statistical method	Hyperbolicky discounting?	Significant relationship with smoking?
Heyman and Gibb (2006)	Students in Cambridge, MA, USA (N_S = 19, N_{NS} = 31, N_{LS} = 21)	Choice (ordered)	Yes ($29)	Real: 1 – 30 days	No	No	H Real: (δ_S = 0.074) (δ_{NS} = 0.036) (δ_{LS} = 0.045)	Algebra and averaging for discounting, F-test and post-hoc tests for analysis	By assumption	Real: Yes, positive for S relative to NS (p<0.01) and LS (p<0.05); No for LS relative to NS.
			No ($1000)	Hypothetical: 7 – 3650 days	No	No	H Hypothetical: (δ_S = 0.007) (δ_{NS} = 0.009) (δ_{LS} = 0.004)			Hypothetical: No.
Reynolds (2006)	Adults in Buffalo, NY, USA (N_S = 15, N_{NS} = 15)	Titration (random – Richards et al. (1999))	No ($10)	1 – 365 days	No	No	H (δ_S = 0.088) (δ_{NS} = 0.020)	NLLS for discounting, non-parametric tests for analysis	By assumption	Yes (p<0.01), positive.
Johnson, Bickel and Baker (2007)	Adults in Burlington, VT, USA (N_S = 30, N_{NS} = 30, N_{LS} = 30)	Titration (random – Richards et al. (1999))	Yes ($100)	Real: 1 – 183 days	No	No	H Real: ($10: δ_S = 0.006, δ_{LS} = 0.003, δ_{NS} = 0.0009) ($100: δ_S = 0.003, δ_{LS} = 0.001, δ_{NS} = 0.0008)	NLLS for discounting, ANOVA for analysis	By assumption	Real: Yes, positive for S (p<0.05) and LS (p<0.05) relative to NS; No for S relative to LS.
			No ($1000)	Hypothetical: 1 – 9131 days	No	No	H Hypothetical: ($10: δ_S = 0.006, δ_{LS} = 0.007, δ_{NS} = 0.002) ($100: δ_S = 0.004, δ_{LS} = 0.002, δ_{NS} = 0.0005) ($1000: δ_S = 0.002, δ_{LS} = 0.0008, δ_{NS} = 0.0003)			Hypothetical: Yes, positive for S (p<0.01) and LS (p<0.05) relative to NS; No for S relative to LS.
Reynolds et al. (2007)	Adolescents in Columbus, OH, USA (N_S = 25, N_{NS} = 26)	Titration (random – Richards et al. (1999))	Yes ($10)	1 – 365 days	No	No	AUC NRD but from Figure 1: (AUC_S = 0.129) AUC_{NS} = 0.234)	AUC for discounting, ANOVA and ANCOVA for analysis	AUC, but dropped subjects that had poor H fit	Yes (p<0.05), positive.
Bickel, Yi, Kowal and Gatchalian (2008)	Adults in Little Rock, AR, USA (N_S = 30, N_{NS} = 29)	Titration (random - Richards et al. (1999))	No ($1000)	1 - 9131 days	No	No	H and E (δ^H_S = 0.007) (δ^H_{NS} = 0.001)	NLLS for discounting, ANCOVA for analysis	Yes (compared to E) based on R^2 comparisons	Yes (p<0.05), positive.

Source: Authors' construction. See Appendix A for more details on these studies.

Notes: S = smoker; NS = non-smoker/never-smoker; ES = ex-smoker; LS = light smoker; T = trier; FS = fast smoking adopter; SS = slow smoking progressor.

H = hyperbolic; E = exponential; QH = quasi-hyperbolic; AUC = area under the curve; NRD = not reported directly; [a] = annual rate; [b] = weekly rate; NLLS = non-linear least squares; ML = maximum likelihood.

ANOVA = analysis of variance; ANCOVA = analysis of covariance

TABLE 1: REVIEW OF EXPERIMENTAL LITERATURE ON SMOKING AND DISCOUNTING BEHAVIOUR (CONTINUED)

Study	Sample (size)	Elicitation method	Task-related incentives (max LL)	Horizon	Front end delay (FED)	Correct for non-linear utility	Models (estimated rates)	Statistical method	Hyperbolicky discounting?	Significant relationship with smoking?
Chabris et al. (2008)	1. Adults in Boston, MA, USA ($N = 126$)	Choice (random – Kirby et al. (1999))	1-in-6-chance ($85)	7 – 186 days	No	No	H ($\delta = 0.015$, SD = 0.02)	ML for discounting, OLS, Tobit, Probit for analysis	By assumption	Yes ($p<0.05$), positive.
	2. Adults in the USA (recruited online) ($N = 326$)		1-in-6-chance ($85)	7 – 186 days	No	No	H ($\delta = 0.008$, SD = 0.009)	ML for discounting, OLS, Tobit, Probit for analysis	By assumption	No
Sweitzer et al. (2008)	Adults in Allegheny County, PA, USA ($N_S = 101$, $N_{NS} = 145$, $N_T = 279$, $N_{ES} = 185$)	Choice (random)	No ($100)	7 - 1825 days	No	No	H ($\delta_S = 0.120$) ($\delta_{NS} = 0.079$) ($\delta_T = 0.090$) ($\delta_{ES} = 0.086$)	NLLS for discounting, ANCOVA for analysis	By assumption	Yes, positive for S relative to NS ($p<0.01$), ES ($p<0.01$) and T ($p<0.01$); No for all other comparisons.
Adams and Nettle (2009)	Adults in 15 major urban areas in the USA (recruited online) ($N_S = 70$, $N_{NS} = 346$)	Choice (ordered)	No ($1000)	30 - 3652 days	No	No	H ($\delta = 1.3$)[a]	NLLS for discounting, logistic regression for analysis	By assumption	No.
Audrain-McGovern et al. (2009)	High school students in northern Virginia, USA ($N_{NS} = 556$, $N_{FS} = 112$, $N_{SS} = 241$)	Choice (random - Kirby et al. (1999))	Not reported ($85)	7 - 186 days	No	No	H Assuming ln transformation: ($\delta_{FS} = 0.023$) ($\delta_{SS} = 0.016$) ($\delta_{NS} = 0.010$)	Algebra and averaging for discounting, latent growth curve modeling (LGCM) and growth mixture modeling (GMM) for analysis	By assumption	LCGM: Yes ($p<0.05$), positive. GMM: Yes, positive for FS ($p<0.05$) and SS ($p<0.05$) relative to NS; No for FS relative to SS.
Jones, Landes, Yi and Bickel (2009)	Adults in Little Rock, AR, USA ($N_S = 86$, $N_{NS} = 141$)	Titration (ordered or random)	No ($1000)	1 - 9131 days	No	No	H NRD but from Figure 3: ($100: $\delta_{S(men)} = 0.012$, $\delta_{NS(men)} = 0.001$, $\delta_{S(women)} = 0.0015$, $\delta_{NS(women)} = 0.002$) ($1000: $\delta_{S(men)} = 0.0075$, $\delta_{NS(men)} = 0.0005$, $\delta_{S(women)} = 0.001$, $\delta_{NS(women)} = 0.001$)	NLLS for discounting, ANCOVA for analysis	By assumption	Yes, positive for $S_{(men)}$ ($p<0.01$) relative $NS_{(men)}$ at $100 and $1000 magnitudes; No for $S_{(women)}$ relative to $NS_{(women)}$ at both magnitudes.
Melanko et al. (2009)	Adolescents in central Ohio, USA ($N_S = 50$, $N_{NS} = 25$). Smokers were split into high and low pyschopathology groups.	Titration (random - Richards et al. (1999))	Yes ($10)	1 - 365 days	No	No	AUC NRD but from Figure 1: ($AUC_{S(low)} = 0.126$) ($AUC_{S(high)} = 0.214$) ($AUC_{NS} = 0.275$)	AUC for discounting, ANOVA for analysis	AUC, no assumption about form of discounting	Yes, positive for $S_{(low)}$ relative to NS ($p=0.01$); No for all other comparisons.
Businelle, McVay, Kendzor and Copeland (2010)	Adults in southern USA ($N_S = 20$, $N_{NS} = 34$)	Choice (ordered)	No ($1000)	0.25 - 9131 days	No	No	H and AUC ($\delta_S = 0.077$) ($\delta_{NS} = 0.039$)	NLLS and AUC for discounting, ANCOVA for analysis	By assumption (but also used AUC)	Yes ($p=0.01$), positive.

Source: Authors' construction. See Appendix A for more details on these studies.

Notes: S = smoker; NS = non-smoker/never-smoker; ES = ex-smoker; LS = light smoker; T = trier; FS = fast smoking adopter; SS = slow smoking progressor.

H = hyperbolic; E = exponential; QH = quasi-hyperbolic; AUC = area under the curve; NRD = not reported directly; [a] = annual rate; [b] = weekly rate; NLLS = non-linear least squares; ML = maximum likelihood.

ANOVA = analysis of variance; ANCOVA = analysis of covariance

TABLE 1: REVIEW OF EXPERIMENTAL LITERATURE ON SMOKING AND DISCOUNTING BEHAVIOUR (CONTINUED)

Study	**Sample (size)**	**Elicitation method**	**Task-related incentives (max LL)**	**Horizon**	**Front end delay (FED)**	**Correct for non-linear utility**	**Models (estimated rates)**	**Statistical method**	**Hyperbolicky discounting?**	**Significant relationship with smoking?**
Harrison, Lau and Rutström (2010)	Adults in Denmark (N_S = 71, N_{NS} = 181)	Choice (ordered)	1-in-10-chance ($1175)	30 - 730 days	Yes	Yes	H and E Linear utility: ($\delta^H_{S(men)} = 0.341$)[a] ($\delta^H_{NS(men)} = 0.240$)[a] ($\delta^H_{S(women)} = 0.329$)[a] ($\delta^H_{NS(women)} = 0.250$)[a]	ML for discounting and analysis	25% - 40% of choices by smokers and non-smokers best characterised by H	Linear utility: Men: Yes (p<0.05), positive; Women: Yes (p<0.10), positive.
							H and E Concave utility: ($\delta^H_{S(men)} = 0.127$)[a] ($\delta^H_{NS(men)} = 0.093$)[a] ($\delta^H_{S(women)} = 0.109$)[a] $\delta^H_{NS(women)} = 0.095$)[a]			Concave utility: Men: Yes (p<0.05), positive; Women: No.
Bickel et al. (2012)	Adults in the USA (recruited online) (N_S = 182, N_{NS} = 614)	Choice (random)	No ($85)	10 - 75 days	No	No	H (Not reported)	Algebra and averaging for discounting, ANCOVA for analysis	By assumption	Yes (p<0.01), positive.
Mitchell and Wilson (2012)	1. Adults in Portland, OR, USA (N_S = 20, N_{NS} = 20)	Choice (random)	Yes ($50)	14 - 154 days	Yes	No	H and QH (0 FED: $\delta^H_S = 0.230$, $\delta^H_{NS} = 0.020$) (+ FED: $\delta^H_S = 0.070$, $\delta^H_{NS} = 0.010$)	NLLS and ML for discounting, ANOVA for analysis	By assumption (but also estimated QH)	Yes (p<0.01), positive.
	2. Adults in Portland, OR, USA (N_S = 16, N_{NS} = 16)		No ($50)	14 - 154 days	Yes	No	H and QH (0 FED: $\delta^H_S = 0.120$, $\delta^H_{NS} = 0.020$) (+ FED: $\delta^H_S = 0.050$, $\delta^H_{NS} = 0.010$)			Yes (p<0.01), positive.
Reynolds and Fields (2012)	Adolescents in Columbus, OH, USA (N_S = 50, N_{NS} = 50, N_T = 41)	Titration (random - Richards et al. (1999))	Yes ($10)	1 - 365 days	No	No	AUC NRD but from Figure 1: ($AUC_S = 0.166$) ($AUC_T = 0.224$) ($AUC_{NS} = 0.347$)	AUC for discounting, ANOVA and ANCOVA for analysis	AUC, no assumption about form of discounting	Yes, positive for S (p<0.01) and T (p<0.05) relative to NS; No for S relative to T.
Stillwell and Tunney (2012)	International online study (N_S = 1592, N_{LS} = 669, N_{NS} = 6777)	Choice (ordered or random)	No ($1000)	7 - 1826 days	No	No	H NRD but from Figure 3: ($\delta_S = 0.437$) ($\delta_{LS} = 0.397$) ($\delta_{NS} = 0.369$)	NLLS for discounting, ANOVA for analysis	Yes (compared to E) based RSS comparisons	Yes, positive for S relative to LS (p<0.01) and NS (p<0.01) and positive for LS relative to NS (p<0.01).
Wing, Moss, Rabin and George (2012)	Adults in the greater Toronto area, Canada (N_S = 23, N_{NS} = 37)	Choice (random - Kirby et al. (1999))	No ($85)	7 - 186 days	No	No	H NRD but from Figure 1: ($\delta_S = 0.017$) ($\delta_{NS} = 0.011$)	Algebra and averaging for discounting, ANCOVA for analysis	By assumption	No.

Source: Authors' construction. See Appendix A for more details on these studies.

Notes: S = smoker; NS = non-smoker/never-smoker; ES = ex-smoker; LS = light smoker; T = trier; FS = fast smoking adopter; SS = slow smoking progressor.

H = hyperbolic; E = exponential; QH = quasi-hyperbolic; AUC = area under the curve; NRD = not reported directly; [a] = annual rate; [b] = weekly rate; NLLS = non-linear least squares; ML = maximum likelihood.

ANOVA = analysis of variance; ANCOVA = analysis of covariance

TABLE 1: REVIEW OF EXPERIMENTAL LITERATURE ON SMOKING AND DISCOUNTING BEHAVIOUR (CONTINUED)

Study	Sample (size)	Elicitation method	Task-related incentives (max LL)	Horizon	Front end delay (FED)	Correct for non-linear utility	Models (estimated rates)	Statistical method	Hyperbolicky discounting?	Significant relationship with smoking?
Balevich, Wein and Flory (2013)	Students in Flushing, NY, USA ($N_S = 50$, $N_{NS} = 102$, $N_T = 91$)	Choice (random) or titration (random)	No ($100)	1 - 1825 days	No	No	H ($\delta_S = 0.126$) ($\delta_{NS} = 0.135$) ($\delta_T = 0.138$)	NLLS for discounting, ANOVA for analysis	By assumption	No.
Poltavski and Weatherly (2013)	Students in Grand Forks, ND, USA ($N_S = 16$, $N_{LS} = 74$, $N_{NS} = 92$)	Choice (random)	No ($100,000)	183 - 3652 days	No	No	H and AUC ($1000: $\delta_S = 0.010$, $\delta_{LS} = 0.010$, $\delta_{NS} = 0.007$) ($100,000: $\delta_S = 0.008$, $\delta_{LS} = 0.008$, $\delta_{NS} = 0.007$)	NLLS and AUC for discounting, ANOVA for analysis	By assumption (but also used AUC)	No.
Sheffer et al. (2013)	Adults in Little Rock, AR, USA ($N_S = 47$, $N_{NS} = 19$)	Titration (random - Richards et al. (1999))	No ($1000)	1 - 9131 days	No	No	H NRD but from Figure 1: ($\delta_S = 0.020$) ($\delta_{NS} = 0.004$	NLLS for discounting, ANCOVA for analysis	By assumption	Yes ($p<0.05$), positive.
Kang and Ikeda (2014)	Adults in Japan ($N_S \approx 862$, $N_{NS} \approx 2588$)	Choice (ordered)	No (¥1000,000 = $10000)	7 - 365 days	Yes	No	E and proxies for H (See Table III, the mean of δ^E ranges from 0.022 to 1.904)[a]	ML for discounting, hurdle model for analysis	Assumes E but constructs H proxies	Yes ($p<0.01$), positive.
Kobiella et al. (2014)	Adults in Mannheim, Germany ($N_S = 27$, $N_{NS} = 31$)	Choice (random)	Yes (€41.32)	14 - 28 days	Yes	No	H ($\delta_S = 0.055$)[b] ($\delta_{NS} = 0.038$)[b]	NLLS for discounting, t-tests for analysis	By assumption	Yes ($p<0.05$), positive.
Hofmeyr et al. (2016)	Adults in Los Angeles, CA, USA ($N_S = 163$, $N_{NS} = 834$, $N_{ES} = 208$)	Choice (random - Kirby et al. (1999))	UCLA: No ($85) USC: 1-in-2-chance ($85)	7-186 days	No	No	H, E and QH ($\delta^H{}_S = 0.021$) ($\delta^H{}_{NS} = 0.012$) ($\delta^H{}_{ES} = 0.013$)	ML for discounting and analysis	41% - 52% of choices best characterised by H	Yes, positive for S relative to NS ($p<0.01$) and ES ($p<0.01$); No for ES relative to NS.

Source: Authors' construction. See Appendix A for more details on these studies.

Notes: S = smoker; NS = non-smoker/never-smoker; ES = ex-smoker; LS = light smoker; T = trier; FS = fast smoking adopter; SS = slow smoking progressor.

H = hyperbolic; E = exponential; QH = quasi-hyperbolic; AUC = area under the curve; NRD = not reported directly; [a] = annual rate; [b] = weekly rate; NLLS = non-linear least squares; ML = maximum likelihood.

ANOVA = analysis of variance; ANCOVA = analysis of covariance

TABLE 2: REVIEW OF EXPERIMENTAL LITERATURE ON SMOKING AND RISK PREFERENCES

Study	Sample (size)	Elicitation method	Incentives (max prize)	Probabilities	Models (estimated rates)	Statistical method (valid?)	Significant relationship with smoking?
Mitchell (1999)	Adults in Durham, NH, USA ($N_S = 20$, $N_{NS} = 20$)	Choice (random)	Yes ($10)	0.1, 0.25, 0.5, 0.75, 0.9, 1	PD ($\gamma_S = 1.328$) ($\gamma_{NS} = 1.371$)	NLLS for risk aversion, non-parametric tests for analysis. (not valid)	No.
Reynolds, Karraker, Horn and Richards (2003)	Adolescents in Morgantown, WV, USA ($N_S = 19$, $N_{NS} = 19$, $N_T = 17$)	Titration (random - Richards et al. (1999))	1-out-of-2-tasks ($10)	0.25, 0.5, 0.75, 0.9, 1	PD NRD but from Figure 2: ($\gamma_S = 1.610$) ($\gamma_{NS} = 1.110$) ($\gamma_T = 3.820$)	NLLS for risk aversion, ANOVA for analysis. (not valid)	Yes, positive for T relative S (p<0.05) and NS (p<0.05); No for S relative to NS.
Reynolds, Richards, Horn and Karraker (2004)	Mostly students in Morgantown, WV, USA ($N_S = 25$, $N_{NS} = 29$)	Titration (random - Richards et al. (1999))	1-out-of-2-tasks ($10)	0.25, 0.5, 0.75, 0.9, 1	PD ($\gamma_S = 1.910$) ($\gamma_{NS} = 1.470$)	NLLS for risk aversion, ANOVA for analysis. (not valid)	Yes (p<0.05), positive (smokers are more risk averse)
Ohmura, Takahashi and Kitamura (2005)	Students in Sapporo, Japan ($N_S = 27$, $N_{NS} = 23$)	Titration (random - Richards et al. (1999))	No (¥100,000 = $1000)	0.1, 0.3, 0.5, 0.7, 0.9	PD and AUC $AUC_S = 0.230$ $AUC_{NS} = 0.180$	AUC and NLLS for risk aversion, correlations and t-tests for analysis. (not valid)	Yes (p=0.08), negative (smokers are less risk averse)
Reynolds (2006)	Adults in Buffalo, NY, USA ($N_S = 15$, $N_{NS} = 15$)	Titration (random - Richards et al. (1999))	No ($10)	0.25, 0.5, 0.75, 0.9, 1	PD ($\gamma_S = 3.908$) ($\gamma_{NS} = 1.574$)	NLLS for risk aversion, non-parametric tests for analysis. (not valid)	Yes (p<0.05), positive (smokers are more risk averse)
Reynolds et al. (2007)	Adolescents in Columbus, OH, USA ($N_S = 25$, $N_{NS} = 26$)	Titration (random - Richards et al. (1999))	Yes ($10)	0.25, 0.5, 0.75, 0.9, 1	AUC and PD (Not reported)	AUC for risk aversion, ANOVA for analysis. (not valid)	No.
Yi, Chase and Bickel (2007)	Adults in Little Rock, AR, USA ($N_S = 30$, $N_{NS} = 29$)	Titration (ordered)	No ($1000)	0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 0.95	PD and AUC (Not reported)	NLLS and AUC for risk aversion, ANOVA for analysis. (not valid)	No when analysing all the data; Yes (p<0.05), positive, when using only probabilities ≥ 0.5.
Anderson and Mellor (2008)	Adults subjects in Williamsburg, VA, USA ($N_S \approx 79$, $N_{NS} \approx 898$)	Choice (ordered - MPL)	Yes ($11.55)	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1	CRRA (1-*r*) ($r = 0.257$)	Algebra and averaging for risk aversion, probit model for analysis. (not valid)	Yes (p<0.1), negative (smokers are less risk averse).
Harrison, Lau and Rutström (2010)	Adults in Denmark ($N_S = 71$, $N_{NS} = 181$)	Choice (ordered - MPL)	1-in-10-chance ($687)	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1	CRRA (1-*r*) ($r_{S(men)} = 0.729$) ($r_{NS(men)} = 0.746$) ($r_{S(women)} = 0.811$) ($r_{NS(women)} = 0.755$)	ML for risk aversion and analysis. (valid)	Men: No; Women: Yes (p<0.06), positive (smokers are more risk averse)
Szrek, Chao, Ramlagan and Peltzer (2012)	Adults in Witbank, South Africa ($N_S \approx 59$, $N_{NS} \approx 292$)	Choice (ordered - MPL)	Yes (R48 ≈ $7)	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1	CRRA (1-*r*) ($r = 0.35$, SD = 0.62)	Algebra and averaging for risk aversion, logit model for analysis. (not valid)	No.
Poltavski and Weatherly (2013)	Students in Grand Forks, ND, USA ($N_S = 16$, $N_{LS} = 74$, $N_{NS} = 92$)	Choice (random)	No ($100,000)	0.01, 0.1, 0.5, 0.9, 0.99	PD and AUC ($1000: $\gamma_S = 0.118$, $\gamma_{LS} = 0.134$, $\gamma_{NS} = 0.307$) ($100,000: $\gamma_S = 0.031$, $\gamma_{LS} = 0.167$, $\gamma_{NS} = 0.181$)	NLLS and AUC for risk aversion, ANOVA for analysis. (not valid)	Yes, negative for S relative to NS (p<0.05); No for all other comparisons.

Source: Authors' construction. See Appendix B for more details on these studies.

Notes: S = smoker; NS = non-smoker/never-smoker; LS = light smoker; T = trier; PD = probability discounting; AUC = area under the curve; NRD = not reported directly; MPL = multiple price list.
NLLS = non-linear least squares; ML = maximum likelihood; ANOVA = analysis of variance.

TABLE 3
DESCRIPTIVE STATISTICS

Variable	Mean	Std Deviation
Demographics		
Age	19.789	1.815
White	0.417	0.495
Male	0.549	0.499
Commerce faculty	0.674	0.470
Financial aid	0.314	0.466
Smoke	0.617	0.487
Treatments		
Risk task first	0.514	0.501
FED: 0 days	0.343	0.475
FED: 1 week	0.326	0.469
FED: 2 weeks	0.331	0.471
High Principal	0.498	0.500

TABLE 4: RANK-DEPENDENT UTILITY THEORY ML ESTIMATES
HETEROGENOUS PREFERENCES

	Model 1		**Model 2**	
	TK		Prelec	
	Estimate	Std Error	Estimate	Std Error
Power function parameter (r)				
Age	0.005	0.013	-0.004	0.011
White	0.038	0.060	0.029	0.051
Male	0.114**	0.055	0.062	0.049
Commerce faculty	0.113*	0.060	0.030	0.062
Financial aid	-0.057	0.065	-0.051	0.058
Risk task first	-0.057	0.055	-0.015	0.050
Smoker	-0.048	0.068	-0.005	0.055
Constant	0.179	0.246	0.366	0.230
PWF parameter (γ)				
Age	-0.002	0.011	-0.003	0.006
White	0.021	0.054	0.001	0.047
Male	0.016	0.050	-0.009	0.044
Commerce faculty	-0.083	0.057	-0.084	0.120
Financial aid	0.061	0.059	0.034	0.056
Risk task first	0.055	0.051	0.054	0.080
Smoker	0.026	0.056	0.028	0.049
Constant	0.876***	0.228	0.871***	0.206
PWF parameter (η)				
Age			-0.027	0.046
White			-0.062	0.121
Male			-0.166	0.137
Commerce faculty			-0.216	0.184
Financial aid			-0.014	0.139
Risk task first			0.166	0.153
Smoker			0.146	0.153
Constant			1.425**	0.676
Error (μ)				
Constant	0.168***	0.008	0.166***	0.008
N	7000		7000	
log-likelihood	-4153.594		-4119.762	

Results account for clustering at the individual level
* p<0.10, ** p<0.05, *** p<0.01

TABLE 5: DISCOUNTING FUNCTION ML ESTIMATES
RANK-DEPENDENT UTILITY AND HOMOGENOUS PREFERENCES

	Model 1	**Model 2**	**Model 3**	**Model 4**
	Exponential Prelec	Hyperbolic Prelec	Quasi-Hyperbolic Prelec	Weibull Prelec
Power function parameter (r)	0.277***	0.327***	0.260***	0.238***
	(0.028)	(0.027)	(0.029)	(0.030)
PWF parameter (γ)	0.797***	0.797***	0.796***	0.795***
	(0.025)	(0.025)	(0.025)	(0.026)
PWF parameter (η)	0.838***	0.884***	0.823***	0.804***
	(0.032)	(0.034)	(0.032)	(0.031)
Discounting parameter (δ)	0.493***	0.502***	0.415***	0.204***
	(0.062)	(0.050)	(0.057)	(0.028)
Discounting parameter (β)			0.988***	1.611***
			(0.004)	(0.115)
Risk error (μ)	0.178***	0.169***	0.181***	0.186***
	(0.009)	(0.008)	(0.010)	(0.010)
Time error (ν)	0.151***	0.231***	0.128***	0.104***
	(0.041)	(0.055)	(0.036)	(0.031)
N	17500	17500	17500	17500
log-likelihood	-9471.828	-9441.151	-9383.297	-9234.32

Results account for clustering at the individual level
Standard errors in parentheses
* *p*<0.10, ** *p*<0.05, *** *p*<0.01

TABLE 6: DISCOUNTING FUNCTION ML ESTIMATES
RANK-DEPENDENT UTILITY AND SMOKING STATUS

	Model 1	**Model 2**	**Model 3**	**Model 4**
	Exponential	Hyperbolic	Quasi-Hyperbolic	Weibull
Power function parameter (r)				
Smoker	0.020	0.023	0.019	0.018
	(0.017)	(0.018)	(0.016)	(0.015)
Constant	0.259***	0.310***	0.245***	0.224***
	(0.028)	(0.027)	(0.030)	(0.030)
PWF parameter (γ)				
Smoker	0.041	0.042	0.040	0.039
	(0.055)	(0.052)	(0.055)	(0.056)
Constant	0.774***	0.774***	0.773***	0.772***
	(0.041)	(0.039)	(0.041)	(0.042)
PWF parameter (η)				
Smoker	0.099	0.105	0.096	0.094
	(0.079)	(0.079)	(0.079)	(0.078)
Constant	0.776***	0.820***	0.764***	0.747***
	(0.049)	(0.049)	(0.049)	(0.049)
Discounting parameter (δ)				
Smoker	0.200***	0.178***	0.176***	0.081**
	(0.060)	(0.051)	(0.059)	(0.032)
Constant	0.359***	0.385***	0.306***	0.156***
	(0.053)	(0.047)	(0.052)	(0.029)
Discounting parameter (β)				
Smoker			0.001	-0.149
			(0.007)	(0.268)
Constant			0.989***	1.670***
			(0.006)	(0.247)
Risk error (μ)				
Constant	0.179***	0.169***	0.182***	0.187***
	(0.009)	(0.008)	(0.010)	(0.010)
Time error (ν)				
Constant	0.136***	0.213***	0.118***	0.097***
	(0.035)	(0.048)	(0.032)	(0.028)
N	17500	17500	17500	17500
log-likelihood	-9338.422	-9306.809	-9263.355	-9118.667

Results account for clustering at the individual level
Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

TABLE 7: NUMBER OF CIGARETTES CONDITIONAL MARGINAL EFFECTS FOR δ

	Model 1	**Model 2**	**Model 3**	**Model 4**
	Exponential	Hyperbolic	Quasi-Hyperbolic	Weibull
Number of cigarettes				
0	0.041 (0.013)	0.036 (0.011)	0.038 (0.012)	0.026 (0.008)
5	0.023 (0.007)	0.020 (0.006)	0.022 (0.007)	0.016 (0.005)
10	0.005 (0.004)	0.004 (0.003)	0.005 (0.003)	0.006 (0.002)
15	-0.013 (0.006)	-0.013 (0.006)	-0.011 (0.005)	-0.004 (0.002)
20	-0.031 (0.011)	-0.029 (0.011)	-0.028 (0.009)	-0.014 (0.004)
25	-0.049 (0.017)	-0.045 (0.016)	-0.044 (0.014)	-0.025 (0.007)

Standard errors in parentheses

TABLE 8: MIXTURE MODEL ML ESTIMATES
RANK-DEPENDENT UTILITY AND HOMOGENOUS PREFERENCES

	Estimate	Std Error	*p*-value	95% Confidence Interval	
			Rank-dependent utility theory		
Power function parameter (r)	0.336***	0.027	0.000	0.283	0.390
PWF parameter (γ)	0.797***	0.025	0.000	0.749	0.846
PWF parameter (η)	0.893***	0.035	0.000	0.825	0.961
			Exponential discounting model		
Discounting parameter (δ_E^{mix})	0.137***	0.017	0.000	0.104	0.169
Mixture probability (π^E)	0.347***	0.034	0.000	0.280	0.414
			Hyperbolic discounting model		
Discounting parameter (δ_H^{mix})	0.730***	0.069	0.000	0.596	0.865
Mixture probability (π^H)	0.653***	0.034	0.000	0.586	0.720
			Error terms		
Risk Error (μ)	0.167***	0.008	0.000	0.152	0.182
Time Error (ν)	0.051***	0.015	0.001	0.021	0.081
N	17500				
log-likelihood	-8808.992				

H_0: $\pi^E = 0.5$, p-value < 0.001

Results account for clustering at the individual level
* *p*<0.10, ** *p*<0.05, *** *p*<0.01

TABLE 9: MIXTURE MODEL ML ESTIMATES
RANK-DEPENDENT UTILITY AND HETEROGENOUS PREFERENCES

	Estimate	Std error	*p*-value	95% Confidence Interval	
Power function parameter (r)					
Smoker	0.055**	0.028	0.047	0.001	0.110
Constant	0.302***	0.028	0.000	0.247	0.358
PWF parameter (γ)					
Smoker	0.036	0.052	0.488	-0.066	0.138
Constant	0.778***	0.039	0.000	0.701	0.854
PWF parameter (η)					
Smoker	0.139*	0.082	0.091	-0.022	0.300
Constant	0.812***	0.048	0.000	0.718	0.905
Discounting parameter (δ_E^{mix})					
Smoker	0.051*	0.026	0.052	-0.000	0.102
Constant	0.111***	0.020	0.000	0.073	0.150
Discounting parameter (δ_H^{mix})					
Smoker	0.164**	0.081	0.044	0.004	0.323
Constant	0.622***	0.077	0.000	0.472	0.773
Mixture probability (π^E)					
Smoker	-0.149**	0.074	0.044	-0.293	-0.004
Constant	0.440***	0.064	0.000	0.315	0.565
Error terms					
Risk error (μ)	0.167***	0.008	0.000	0.152	0.182
Time error (ν)	0.048***	0.013	0.000	0.023	0.073
N	17500				
log-likelihood	-8657.459				

Results account for clustering at the individual level
* *p*<0.10, ** *p*<0.05, *** *p*<0.01

FIGURES

Figure 1

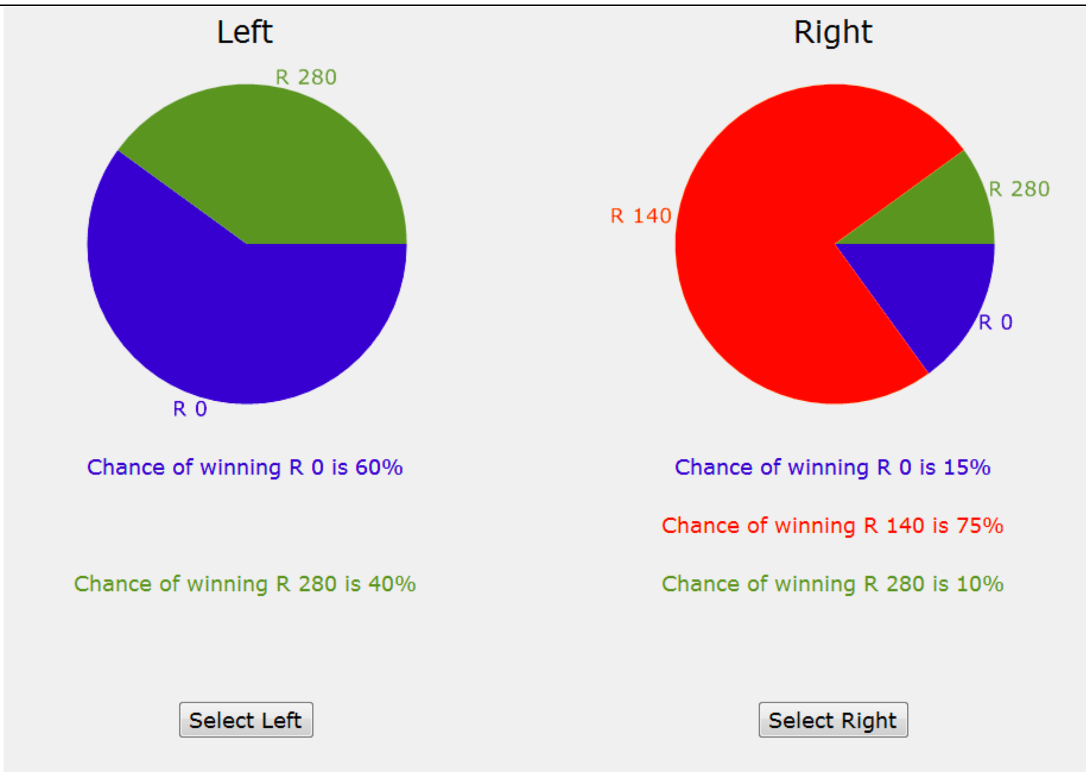

Figure 2: MM triangles of lotteries in the risk preference task

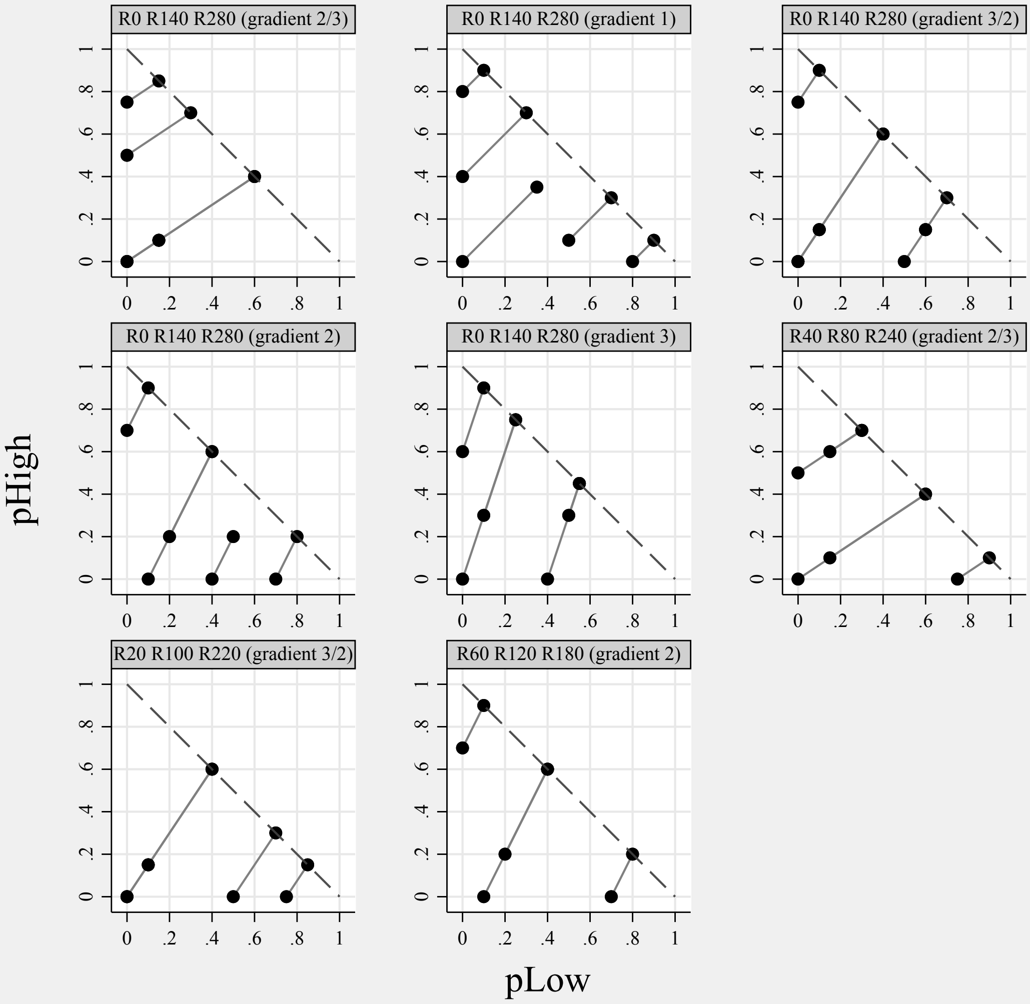

Figure 3

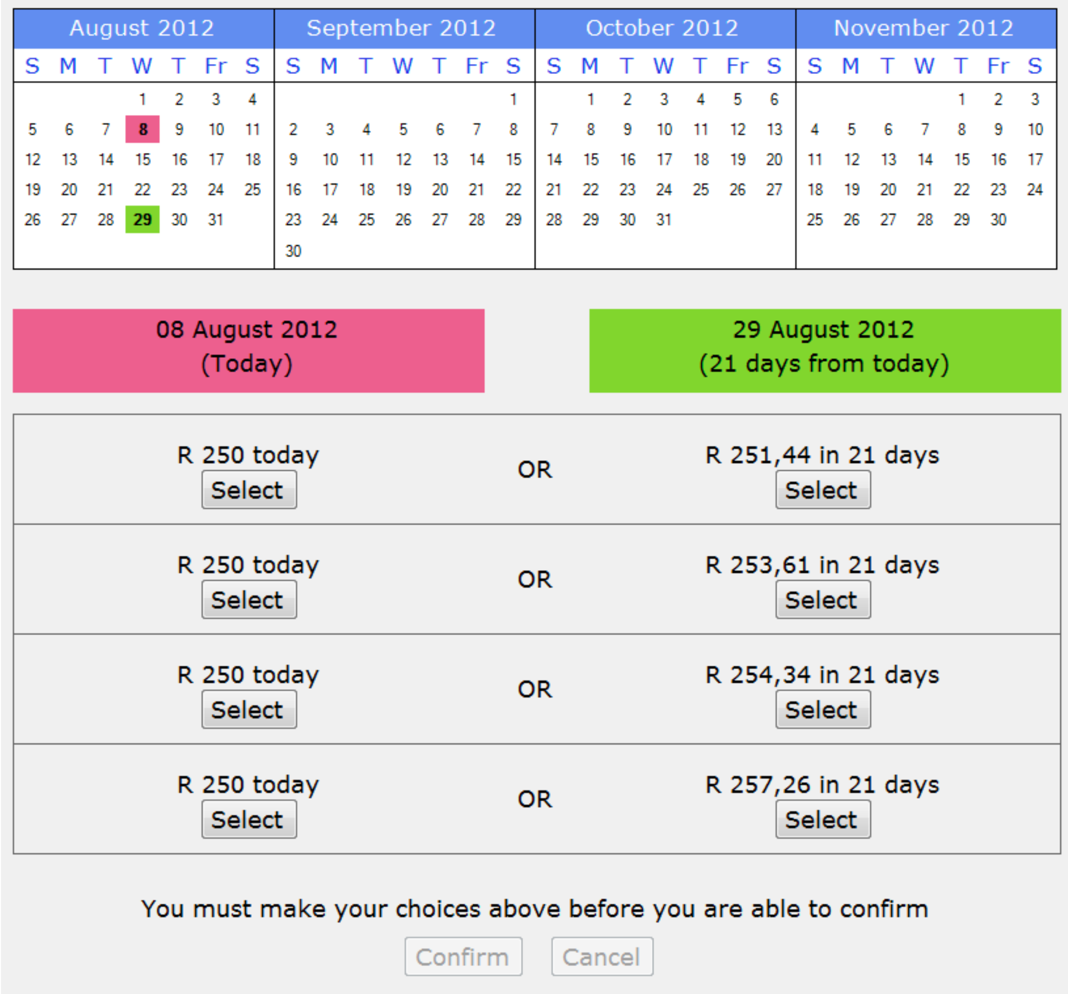
August 2012
September 2012
October 2012
November 2012
S M T W T Fr S
08 August 2012
(Today)
29 August 2012
(21 days from today)
R 250 today
Select
OR
R 251,44 in 21 days
Select
R 250 today
Select
OR
R 253,61 in 21 days
Select
R 250 today
Select
OR
R 254,34 in 21 days
Select
R 250 today
Select
OR
R 257,26 in 21 days
Select
You must make your choices above before you are able to confirm
Confirm
Cancel

Figure 4: Fraction of LL choices and interest rate offered

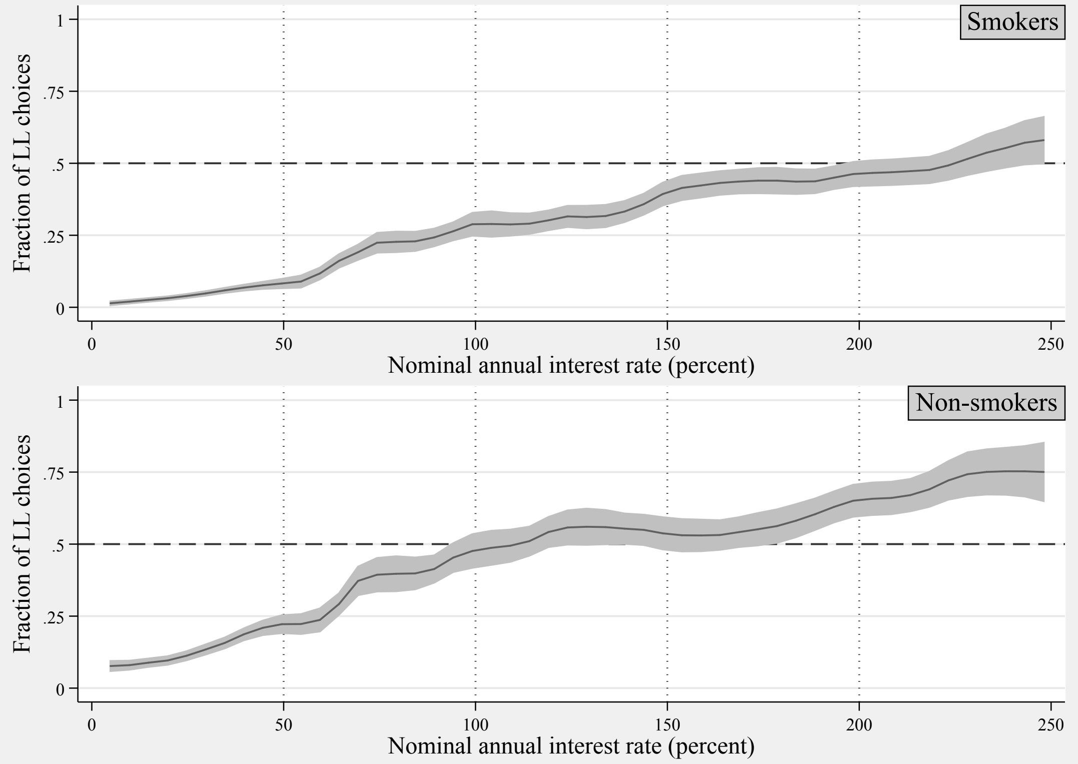

Figure 5: Exponential and hyperbolic discount factors