Reinvestigating Adolescent Smoking Decisions: The Importance of Genetic Markers, Impulsivity and Their Interactions

Steven F. Lehrer
Queen's University and NBER

Janet Audrain-McGovern
University of Pennsylvania

Weili Ding
Queen's University

J. Niels Rosenquist
Harvard Medical School

November 2009

SPECIFIC MOTIVATION

Tobacco Policy – Rhetoric has long dominated as one of leading cause of disease and death worldwide.

Many in this room are familiar with how economic analyses that introduce objective and increasingly sophisticated knowledge into these debates (gender and race)

Several theoretical papers have started to close this gap and discuss the likely effects of policies as the decision to smoke represents an inter-temporal tradeoff (i.e. Gruber and Köszegi, 2001) "unrealized intentions to quit at some future date are a common feature of stated smoker preferences" -> Implications for policy design

In parallel there is a large literature in psychology and behavioural genetics that looks at time discounting and molecular genetic factors.

Goal of this paper is a first step to bridge the gap between disciplines from an empirical perspective

Can be viewed as a start at developing richer empirical evidence to inform both research and policy debates on addictive substances.

More generally this paper shows that there are gains from importing data that was initially collected by researchers in other disciplines

GENERAL MOTIVATION

With data on genetic markers at the molecular level increasingly available, I view this paper as a continuation on how empirical microeconomists can incorporate it into their analyses

- ➔ Entering the black-box of unobserved heterogeneity
- ➔ Easy accessible evidence on how the relative allele frequencies vary across different ethnic populations
- ➔ Also the prevalence of many traits correlates with relative allele frequencies -> Why we don't know
- ➔ Taken together this may be a case of OVB on variables that can't be identified using within individual variation
    - ➔ To illustrate these issues a simple application to a *bad* model "Traditional Model of Smoking Demand"

LINKS TO THREE LITERATURES

1) Psychological ingredients have drawn increasing attention in several domains of empirical economics particularly in regards to financial decisions

➔ Since risk attitudes affect economic choices important to understand the source of their variation

2) Results from the seminal Fuchs (1982) study suggest that time preferences can be proxied with measures of smoking behavior.

➔ Khwaja, Silverman and Sloan (2007) find the opposite

Source of differences could be age of the subjects not simply the metrics

3) Child Development (Critical and Sensitive Periods)

-> Only Audrain-McGovern et al. (2009) examine if it reflects a stable disposition

Possibly consistent with Gruber and Köszegi (2001) who provide evidence that, among high school seniors, 56% claimed that they would quit in 5 years, but only 31% quit in that time.

STRUCTURE OF THE TALK

Mini Digression on Definitions

- ➔ What is the appropriate endophenotype for smoking?

Data

Genetics and Science Primer

- ➔ Impulsivity is a complex psychological construct that has been shown to have a genetic element (45-50%)

Empirical Strategy

Evidence from reduced form relationships

Discussion -> Suggestive support for alternative theories

Conclusions and Directions for Future Research

DIGRESSION ON ENDOPHENOTYPES

An issue that makes this multi-disciplinary literature so confusing is that the same terminology is used for different concepts

Related to an important topic in the genetics literature- what is the correct endophenotype for a particular trait such as smoking.

Questions on time discounting, subadditive discounting, self-control and impulsivity are very similar in structure.

Most often they are designed to reveal preferences under commitment and, absent consideration of arbitrage opportunities with respect to the real interest rate, answers are thus thought to reflect the standard notion of subjective time discounting.

Are survey questions reliable?

We will calculate a discount rate over a *delay*–the difference between the time when the outcome is realized and the time when the choice is made – is independent of the number of intervals in the delay over which the discount rate is calculated, where an *interval* is the difference in time between two outcomes for which there is an intertemporal tradeoff.

DATA

Our primary data source is the Georgetown Adolescent TObacco Research Study

Also Adult Longitudinal Outcomes Health Assessment (ALOHA)

Information on Genetic Markers:

-Buccal swabs with standard techniques. Multiple markers have been collected and double-checked at 20%.

Four follow-up surveys with the students were conducted.

Each survey conducted at school contains standard demographic information and detailed smoking information.

Academic performance and health measures collected a subset of the surveys.

Very little attrition in GATOR -- Follow up is 95%, 96%, 93% and 89%.

Aloha follows cohort of 1000 adolescents into young adulthood (to age 22).

Psychological Variables of Interest

Delay discounting was measured from the pattern of choices across 27 questions on a monetary choice questionnaire (Kirby et al., 1999). The 27 choices define 10 ranges of discount rates with delays ranging from 7 days to 186 days. Delay discounting is measured by fitting a hyperbolic function to bivariate data on indifference points between choices of small, medium, and large delayed rewards and the time delay. Delay discounting was measured at three time points (10th grade spring and in the first 2 years after high school ~ages 15 – 21).

Self Control (Good and Bad) In the 10th grade, seventeen items that measured positive aspects of self-control were derived from an inventory of general control in daily situations. At the same interview 24 items that measured aspects of poor self-control were derived from inventories of general poor control, impulsive behavior, and anger coping.

Risk of Smoking Five items requesting information on adolescents' perceptions of immediate harm of smoking in general ("... someone who starts smoking a pack of cigarettes a day at age 16"), and the personal risk of smoking (e.g., "... is smoking very risky to your health ..."). The smoking risk belief items were measured in the 11th grade.

Novelty-Seeking Personality A 20-item version of the Temperament and Character Inventory was used to measure novelty seeking

Table 1: Summary Characteristics of the Sample

Time Invariant Variables N=893						
Variable		Mean		Standard Deviation		
Male		0.469		0.499		
African American		0.073		0.260		
Hispanic		0.093		0.291		
Asian		0.106		0.308		
Caucasian		0.667		0.471		
Biological Parent smoked		0.449		0.498		
Body Mass Index		23.426		4.410		
Obese (BMI>=30)		0.081		0.272		
AD diagnosis		0.043		0.202		
HD diagnosis		0.040		0.197		
ADHD diagnosis		0.063		0.243		
Time Varying Variables						
	Grade 10 Mean	Grade 10 Standard Deviation	Grade 11 Mean	Grade 11 Standard Deviation	Grade 12 Mean	Grade 12 Standard Deviation
Tried Smoking	0.433	0.495	0.483	0.500	0.533	0.499
Current Smoker	0.091	0.288	0.152	0.359	0.178	0.382
Years as a Regular Smoker	0.116	0.398	0.245	0.680	0.399	0.968
Smoker in Household	0.241	0.428	0.246	0.431	0.231	0.422
Grade Point Average (GPA)	3.184	0.567	3.148	0.598	3.176	0.571
Age	16.03	0.399	17.03	0.396	18.03	0.400
Depressed last period	0.168	0.374	0.169	0.375	0.122	0.327
N	834		863		879	

SCIENCE

Lots of evidence suggests that the role of genetic factors leading to poor health outcomes is substantial

As a result, pharmaceutical companies regularly develop rugs that target specific genetic markers.

Each person inherits an allele of a gene from each parent at conception. The two alleles combine to form a marker.

Alleles differ by the building blocks that make up DNA. (SNPs)

Any difference in the coding of a specific marker between children of the same parents is an “experiment in nature"

The path from genotype to phenotype has a number of steps that are influenced by environment

HOW DO WE KNOW WHAT SPECIFIC GENETIC VARIANTS DO

Different regions of the brain engage in different processes.

Historically evidence came first from stroke victims and other anatomical deficits

Ethical constraints place limits

Genetics for addiction initially focus easily on genes that are believe to be located within the brain's reward region

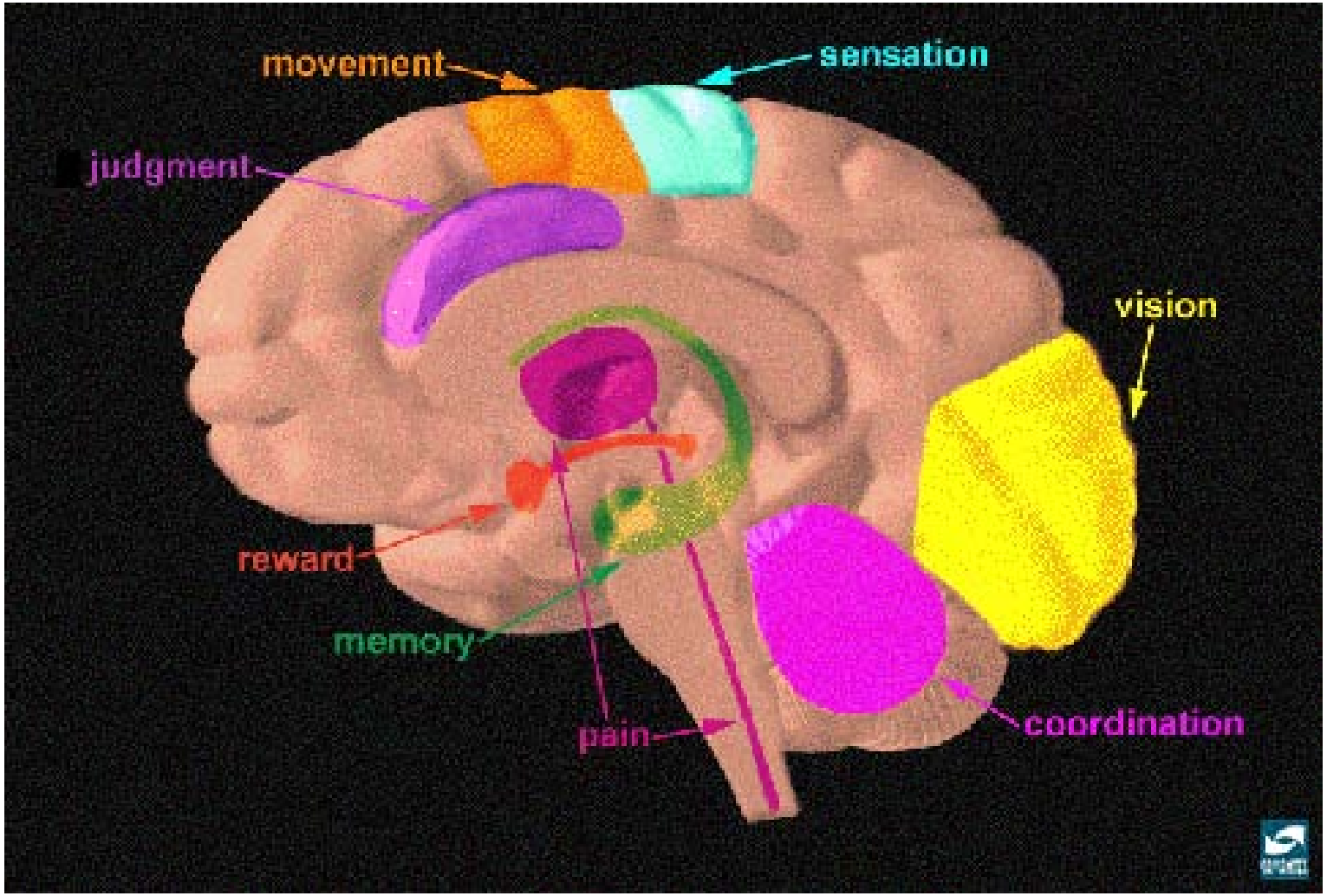
Guidance now often comes from well-replicated GWAS or biological pathways
-> Leamer would have a clear preference

Knockout mice (animal models)

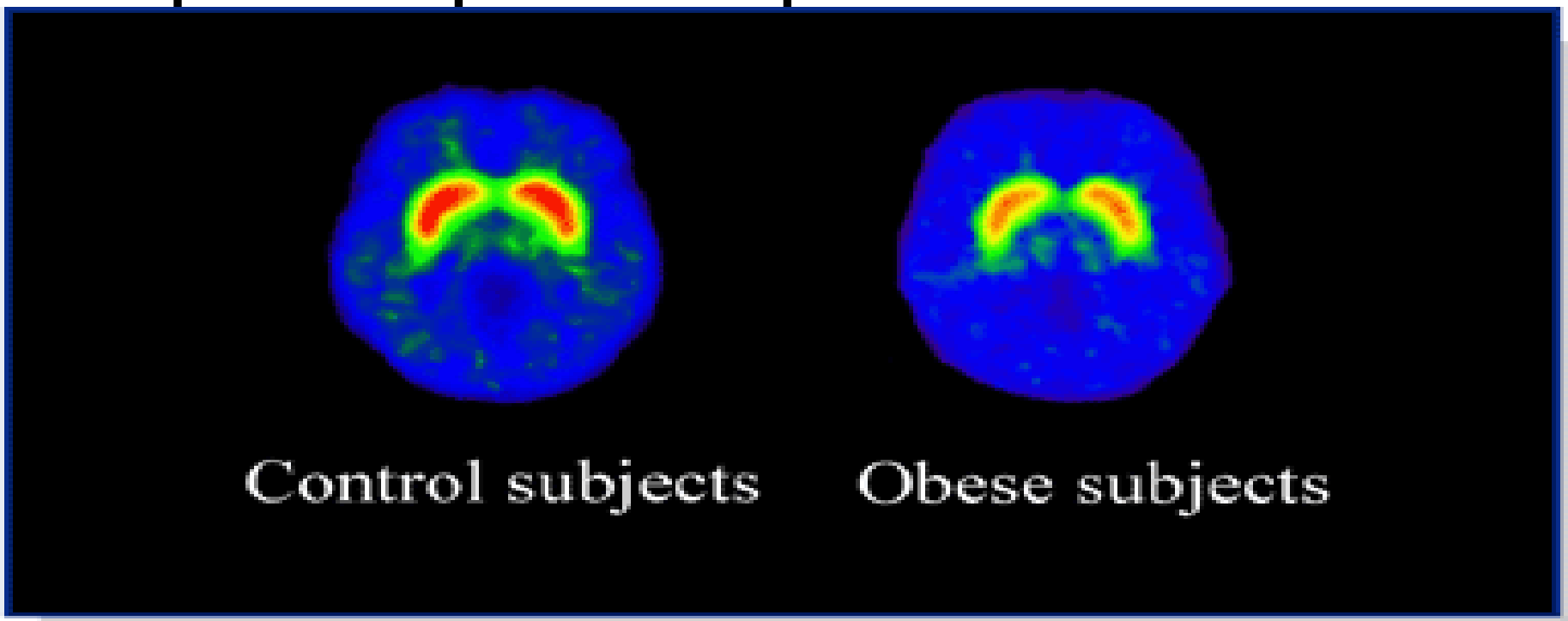
Imaging studies (comparing function)

In vitro studies of protein expression

NEUROANATOMY



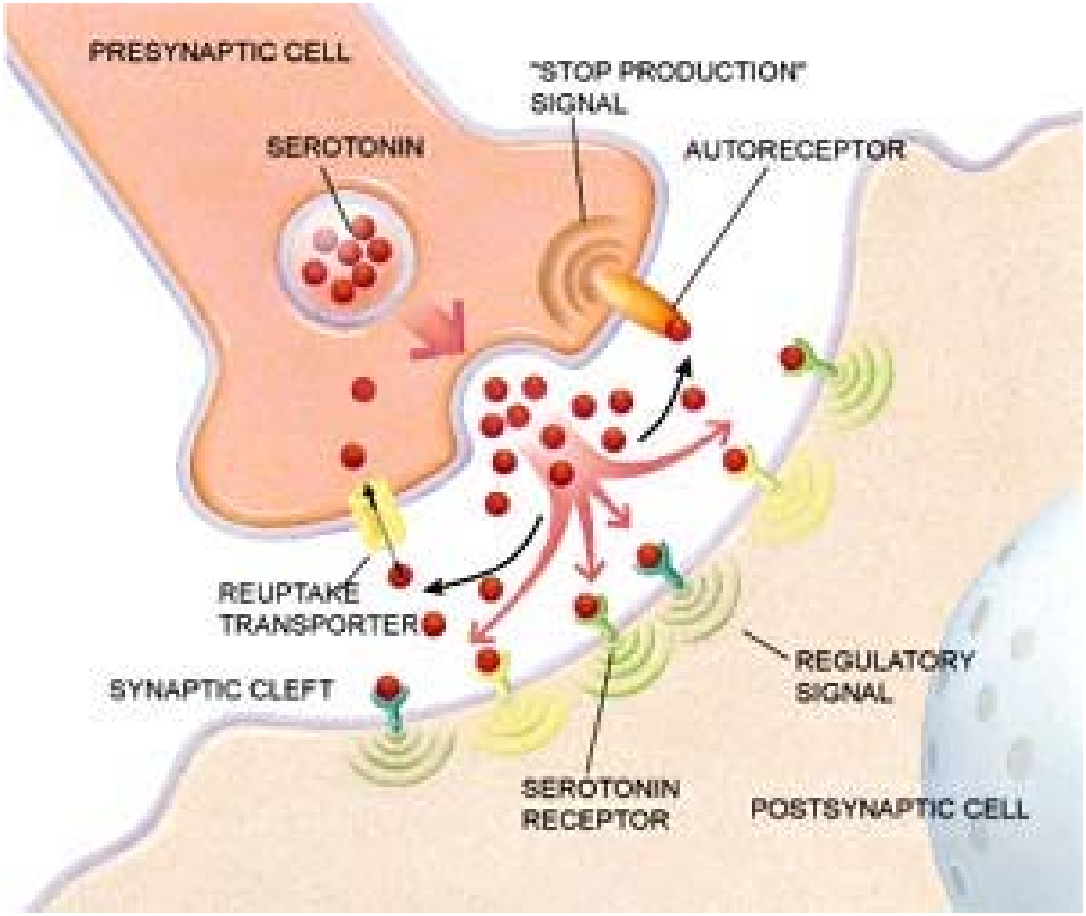
Example: the Dopamine Receptor D2 locus.



These composite brain scan images show that obese individuals have significantly fewer dopamine receptors in the outlined area than control subjects. These receptors transmit pleasurable feelings from basic activities such as eating and sex. Low levels of these receptors also have been found in people addicted to drugs of abuse. The reduced reward experienced by people with this deficiency may make them more likely to engage in addictive behaviors.

WHAT IS GOING ON

The activity and number of the serotonin transporter proteins determine the length of time that the chemical signal remains in the synapse



Genetic Markers in the GATOR / ALOHA Data

COMT (Catechol-*O*-methyl transferase) is one of several enzymes that degrade catecholamines such as dopamine and norepinephrine. The valine variant catabolizes dopamine at up to four times the rate of its methionine counterpart.

Genetic variants of the OPRM1 (opioid receptor, mu 1) and GSTM1 (glutathione S-transferase) change an individual's susceptibility to carcinogens and toxins as well as affect the toxicity and efficacy of certain drugs.

Cytochrome P450 2A6 is the primary enzyme responsible for the oxidation of nicotine and cotinine. CYB 2B6 big links with buproprin.

Dopamine active transporter, DAT, SLC6A3) and serotonin transporter (5HTT) code for proteins that lead to the reuptake of dopamine and serotonin respectively. These are the primary mechanisms through which dopamine and serotonin is cleared from synapses, transporting dopamine from the synapse into a neuron.

The dopamine D2 receptor (DRD2) gene associated with many addictive behaviors and sensation-seeking activities. Hypothesized to be linked to the density of dopamine receptors.

Evidence that *interactions* of markers could have powerful effects. (Gene-gene interactions)

EMPIRICAL FRAMEWORK

Essentially what we do is similar in spirit to Khwaja et al. (2007, JHE) who correlate different psychological measures with smoking. We will also control for parental smoking and genetic markers to show when they alter conclusions.

Today we define smoking as being a regular smoker in this survey.

We will also document associations between these genetic markers and both levels of as well as changes in smoking and impulsivity measures

Traditional Model of Cigarette Demand

Standard, constrained, lifetime utility-maximizing framework of economics:

$$U_t = f[C_t, X_t)]$$

C_t - consumption of addictive substance at time t and X_t – consumption of composite good at time t

One then maximizes utility subject to a budget constraint which produce a demand function $C_t = g[P_t , Y_t , \mathbf{Z}_t]$

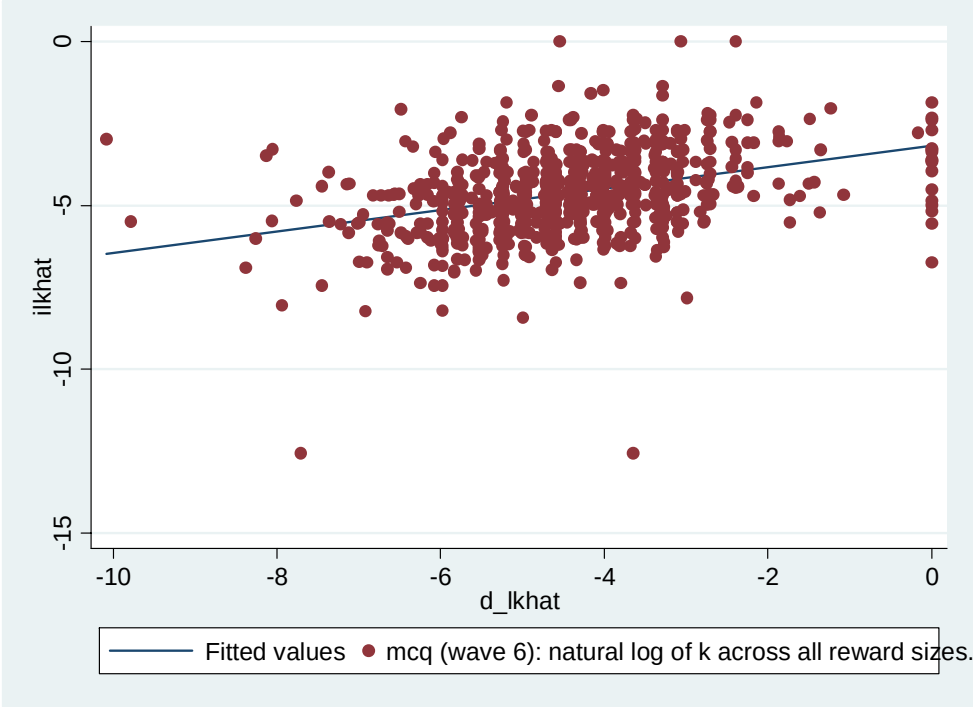
P_t - current price of addictive substance, Y_t – income, $\mathbf{Z}_t$ - vector of variables reflecting tastes

Key point is only contemporaneous inputs (prices) matter and no state dependence. (Note Myopic models Pollak et al. add lagged consumption. Also does not include future or "irrational"

EMPIRICAL EVIDENCE

1) Existing economic literature on time preferences and smoking implicitly rests on the assumption that time discounting is, within person, *well-correlated* over the life course.

Exact correlations in the data 0.3661between 5 year measures



Wave 6

Yet the correlation is 0.8995 between successive years.

2)Raw differences in impulsivity and smoking behavior is observed

Wave	1	2	3	4	5	6	7	8	9
Non-Smokers	-4.389 (1.411)	-4.417 (1.409)	-4.407 (1.408)	-4.407 (1.407)	-4.449 (1.384)	-4.556 (1.321)	-4.624 (1.300)	-4.636 (1.297)	-4.646 (1.326)
Smokers	-3.880 (1.249)	-3.904 (1.311)	-3.853 (1.340)	-4.062 (1.379)	-4.045 (1.478)	-4.021 (1.521)	-4.073 (1.517)	-4.055 (1.530)	-4.063 (1.479)
Test	-2.810 [0.0025]	-3.604 [0.0002]	-4.024 [0.0000]	-2.899 [0.0019]	-3.629 [0.0001]	-5.188 [0.0000]	-5.377 [0.0000]	-5.746 [0.0000]	-5.719 [0.0000]

Wave	1	2	3	4	5	6	7	8	9
Non-Smokers	-4.668 (1.275)	-4.660 (1.290)	-4.653 (1.257)	-4.671 (1.253)	-4.681 (1.257)	-4.702 (1.271)	-4.802 (1.279)	-4.814 (1.271)	-4.817 (1.275)
Smokers	-4.327 (1.073)	-4.458 (1.117)	-4.529 (1.470)	-4.471 (1.397)	-4.461 (1.350)	-4.445 (1.261)	-4.437 (1.286)	-4.443 (1.265)	-4.428 (1.249)
One Sided Test	-1.643 [0.050]	-1.194 [0.117]	0.7602 [0.224]	1.5048 [0.066]	-1.829 [0.034]	-2.407 [0.008]	-3.453 [0.000]	-3.598 [0.000]	-3.726 [0.000]

3) Interesting temporal variation in impulsivity measure but early impulsivity is what seems to matter for smoking $C_{it} = \beta_0 + \beta_1 I_{1t} + \beta_2 Z_t + \beta_3 G_i + \beta_4 I_t * G_i + w_t + \varepsilon_t$

Wave 1 Ln of k across all reward sizes	0.024 (0.003)**	0.024 (0.003)**	0.049 (0.009)**
Asian	-0.055 (0.016)**	-0.061 (0.017)**	-0.060 (0.017)**
African American	0.012 (0.019)	-0.002 (0.020)	-0.015 (0.020)
Hispanic	-0.004 (0.016)	-0.007 (0.017)	-0.005 (0.017)
Male	0.019 (0.010)+	0.020 (0.010)*	0.022 (0.010)*
Biological Parent Smokes	0.105 (0.010)**	0.106 (0.010)**	0.105 (0.010)**
AA		0.045 (0.015)**	0.061 (0.049)
AC		0.012 (0.011)	-0.149 (0.035)**
CT		0.016 (0.012)	0.059 (0.039)
TT		-0.032 (0.027)	-0.132 (0.073)+
A1A1		0.032 (0.011)**	-0.062 (0.030)*
A1A2		0.006 (0.005)	0.046 (0.017)**
mp010		0.034 (0.018)+	-0.057 (0.061)
mp110		-0.007 (0.010)	0.040 (0.033)
Interactions			YES
R-squared	0.06	0.07	0.09

$$C_{it} = \beta_0 + \beta_1 I_{6t} + \beta_2 Z_t + \beta_3 G_i + \beta_4 I_t * G_i + w_t + \varepsilon_t$$

Wave 6 Ln of k across all reward sizes	0.044 (0.008)**	0.044 (0.008)**	0.070 (0.021)**
Asian	-0.073 (0.035)*	-0.067 (0.036)+	-0.053 (0.036)
African American	-0.167 (0.041)**	-0.170 (0.042)**	-0.195 (0.042)**
Hispanic	0.047 (0.036)	0.042 (0.037)	0.030 (0.036)
Male	0.088 (0.020)**	0.096 (0.020)**	0.090 (0.020)**
Biological Parent Smokes	0.160 (0.020)**	0.163 (0.020)**	0.151 (0.020)**
AA		0.032 (0.031)	-0.023 (0.129)
AC		0.025 (0.021)	-0.315 (0.087)**
CT		-0.010 (0.024)	0.340 (0.095)**
TT		-0.108 (0.060)+	-0.040 (0.235)
A1A1		0.055 (0.022)*	-0.070 (0.079)
A1A2		0.002 (0.011)	0.021 (0.042)
mp010		0.070 (0.037)+	-0.205 (0.161)
mp110		0.001 (0.021)	-0.199 (0.082)*
Interactions			YES
R-squared	0.15	0.15	0.17

$$C_{it} = \beta_0 + \beta_{16} I_{6t} + \beta_{11} I_{1t} + \beta_2 Z_t + \beta_3 G_i + \beta_4 I_t * G_i + w_t + \varepsilon_t$$

Wave 6 Ln of k across all reward sizes	0.023 (0.008)**	0.024 (0.008)**	0.034 (0.023)
Wave 1 Ln of k across all reward sizes	0.040 (0.007)**	0.041 (0.007)**	0.102 (0.019)**
Asian	-0.056 (0.036)	-0.049 (0.037)	-0.056 (0.037)
African American	-0.166 (0.042)**	-0.173 (0.043)**	-0.195 (0.044)**
Hispanic	0.043 (0.037)	0.037 (0.038)	0.030 (0.038)
Male	0.069 (0.020)**	0.077 (0.020)**	0.070 (0.020)**
Biological Parent Smokes	0.154 (0.020)**	0.156 (0.020)**	0.146 (0.020)**
AA		0.040 (0.032)	-0.231 (0.149)
AC		0.036 (0.022)+	-0.283 (0.097)**
CT		-0.008 (0.024)	0.305 (0.105)**
TT		-0.185 (0.065)**	-0.263 (0.323)
A1A1		0.058 (0.022)*	-0.009 (0.089)
A1A2		0.007 (0.011)	-0.020 (0.049)
Mp010		0.076 (0.037)*	-0.132 (0.174)
Mp110		-0.003 (0.021)	-0.216 (0.092)*
comt_new==HH		-0.054 (0.028)+	-0.254 (0.111)*
comt_new==HL		-0.041 (0.025)+	-0.088 (0.108)
Interactions R-squared	0.15	0.16	YES, 0.20

4) How risky is it to smoke?

	Non Smokers	Smokers
Very High	**205**	**10**
High	**382**	**29**
Low	**249**	**34**
Very Low	69	24

Ordinal Dependent variable

Any Links to Early Impulsivity With and Without Genes

$$Risk_{it} = \beta_0 + \beta_1 I_{1t} + \beta_2 Z_t + \beta_3 G_i + \varepsilon_t$$

Wave 1 Ln of k across all reward sizes	0.070 (0.045)	0.036 (0.047)
Asian	-0.299 (0.197)	-0.262 (0.212)
African American	-0.491 (0.241)*	-0.705 (0.276)*
Hispanic	-0.039 (0.199)	-0.081 (0.218)
Male	0.437 (0.119)**	0.415 (0.127)**
Biological Parent Smokes	0.442 (0.119)**	0.419 (0.127)**
AA		0.308 (0.194)
AC		-0.073 (0.134)
CT		0.337 (0.156)*
TT		-0.212 (0.332)
A1A1		0.246 (0.138)+
A1A2		0.026 (0.068)
mp010		0.036 (0.232)
mp110		0.120 (0.135)
comt_new==HH		-0.155 (0.175)
comt_new==HL		-0.098 (0.157)
Observations	981	866

5) Strong evidence of associations between genetic variants of the CYP gene, DAT gene and impulsive behaviour. On changes in impulsivity

Asian	0.054 (0.199)	-0.009 (0.204)
African American	-0.142 (0.230)	-0.187 (0.237)
Hispanic	0.084 (0.214)	0.038 (0.217)
Male	-0.152 (0.118)	-0.154 (0.118)
Age	14.205 (5.758)*	14.748 (5.766)*
Age Squared	-0.476 (0.196)*	-0.494 (0.197)*
AA		0.234 (0.185)
AC		0.280 (0.126)*
CT		0.050 (0.144)
TT		0.029 (0.337)
A1A1		0.126 (0.132)
A1A2		0.146 (0.064)*
mp010		-0.133 (0.224)
mp110		-0.144 (0.126)
comt_new==HH		0.059 (0.161)
comt_new==HL		0.189 (0.145)

Preliminary Conclusions

Substantial changes in Kirby Score (delay discounting) between 15 – 21.

Genetic factors are both highly correlated to the level and changes in these scores

The role of impulsivity on smoking behaviour is highly significant at many points in the adolescent lifecycle

Early measures seem to be more highly correlated with subsequent behaviour

Controlling for genetic factors alters the story a bit particularly for race

Adding Interactions really complicates the story but a subset of these are highly significant.

Interesting pattern in impacts of knowledge of smoking risks once we include genetic factors.

Lots of future directions in this research program that we can discuss over lunch