Self-control in food purchases

Rachel Griffith

joint work with Laurens Cherchye, Bram De Rock, Martin O’Connell, Kate Smith and Frederic Vermeulen

ECARES Brussels, April 2018
Self-control problems

- A large theoretical literature posits that people suffer from self-control problems
  - predicts that people over-consume some tempting goods relative to what their long-run self would choose
- An extensive psychological literature suggests that individuals use self-regulation and behavior modification in an attempt to mitigate self-control problems
  - New Years’ resolutions to eat a more healthy diet are one of the most common forms of self-regulation and behavior modification that people use
- Poor decisions over food choices have potentially important consequences for well-being and welfare
Richard H. Thaler - Facts

Richard H. Thaler

Born: 1945, East Orange, NJ, USA

Affiliation at the time of the award: University of Chicago, Chicago, IL, USA

Prize motivation: “for his contributions to behavioural economics"

Prize share: 1/1

© Nobel Media III: N. Elmehed

To cite this page
Richard H. Thaler - Facts

Richard H. Thaler

Born: 1945, East Orange, NJ, USA

Affiliation at the time of the award: University of Chicago, Chicago, IL, USA

Prize motivation: “for his contributions to behavioural economics”

Prize share: 1/1

© Nobel Media III: N. Elmehed

To cite this page

Share this: 55

Answers: 2563

Nobel Prize Lessons
“A new year, a new you? Heterogeneity and self-control in food purchases”

Laurens Cherchye, Bram De Rock, Rachel Griffith, Martin O’Connell, Kate Smith and Frederic Vermeulen

“New Year’s Day ... now is the accepted time to make your regular annual good resolutions. Next week you can begin paving hell with them as usual.”

– Mark Twain
Self-control problems

- Food consumption is an obvious setting in which self-control problems may play an important role:
  - the theoretical literature (ODonoghue and Rabin, 2000; Gul and Pesendorfer, 2001; ...)
  - evidence from the experimental literature (Read and Van Leeuwen, 1998; Gilbert et al., 2002; Downs et al.; 2009)
  - the existence of a multi-billion dollar diet industry

- Of particular concern because poor decisions over food choices have potentially important consequences
  - a lack of self-control is possibly causally linked to poverty (Bernheim, Ray and Yeltekin, 2015; Mani, Mullainathan, Shafir and Zhao, 2013; Haushofer and Fehr, 2014)
  - and seemingly worse in the young (Ameriks, Caplin, Leahy and Tyler, 2007; Bucciol, 2012)
Child obesity is more common in deprived areas.
Why is self-control related to poverty?

The poor often behave in less capable ways, which can further perpetuate poverty. **THEORY**: Poverty reduces cognitive capacity because poverty-related concerns consume mental resources, leaving less for other tasks.
Why is self-control related to poverty?

POVERTY AND THE BRAIN

FINDING 1  Human mental bandwidth is finite

THE RULE OF 7
The largest number of discrete pieces of information the average brain can manage is around seven.

CHUNKING
Way to improve memory
Like phone numbers
123-456-7890 = groups of 3 + 3+ 4 numbers
Why is self-control related to poverty?

FINDING 2
Poverty imposes such a massive cognitive load on the poor that they have little bandwidth left over to do many of the things that might lift them out of poverty like:

- Going to night school
- Searching for a new job
- Remembering to pay bills on time
Why is self-control related to poverty?

**WOW** The condition of poverty’s mental burden equals losing 13 IQ points.

**OR** of losing a night’s sleep.

**I.Q. FACTS**

- Poverty is a bigger risk factor for mental illness than being exposed to warfare.

- 68% of the U.S. population has an average IQ of 90-110.
- 95% of the population has an IQ of 90-110.
- 1% of the population has an IQ of 136 or higher.
Why is self-control related to poverty?

“Poverty is the equivalent of pulling an all-nighter. Picture yourself after an all-nighter. Being poor is like that every day.”

Harvard Economist Sandhil Mullainathan
Governments around the world are struggling with how to address rising rates of obesity. The presence of self-control problems opens the possibility that public policy may be able to improve welfare by increasing the cost of consumption today to reflect unanticipated future costs, or by facilitating the use of commitment devices (ODonoghue and Rabin (2003, 2006) and Gruber and Koszegi (2004)).
Self-control problems

- Attention has focused on variation in diet quality across individuals
  - impact on rising obesity and diet-related diseases (Cutler et al. (2003), Finkelstein and Zuckerman (2008), Bleich et al. (2008), Baum and Ruhm (2009) and Cutler and Lleras-Muney (2010))

- Cross-sectional heterogeneity in diet quality has been shown to be associated with important inequalities in health outcomes
  - e.g. Cutler and Lleras-Muney (2010) show that obesity rates are twice as high among the less well educated

- We focus on variation in diet quality within-person over time
  - we show that this within-individual variation in diet quality is substantial from a nutritional perspective
Our results build on a literature that finds empirical evidence of similar effects in other settings:

- financial products – e.g. Ashraf et al. (2006), DellaVigna and Malmendier (2006), Bucciol (2012)
- magazine subscriptions – Oster and Morton (2005)
- alcohol – Hinnosaar (2016)

as well as in grocery purchases using alternative identification strategies:

- food stamps – Shapiro (2005)
- on-line grocery ordering – Milkman et al. (2010), Sadoff et al. (2015).
Google searches for “Diet”

[Graph showing Google search intensity for "Diet" from 01 Jan 2013 to 01 Jan 2017, with data for the US and UK indicated]
Google searches for “Healthy food”
Share of calories in grocery basket from healthy food
Magnitude of decline in share of calories from healthy foods

- Previous figure shows a decline in the calories purchased from healthy foods of over 10%, on average, over the calendar year.
- This is a sizeable change, it is similar in magnitude to variation across individuals.
- If we considered the average shopping basket, this change would be approximately equivalent to cutting the calories from chocolate, sweets, and cake in half, and doubling calories from fruit.
Empirical challenge in studying self-control

- A key empirical challenge is to disentangle self-control problems from heterogeneity in preferences
  - if I eat more chocolate cake than you it does not necessarily mean I lack self-control, I might just really like chocolate cake more
  - by looking within individuals over time we are able to control for (fixed) preference heterogeneity

- Individual might have self-control and plan to vary her diet
  - will predominantly be day to day e.g. eat cake today and eat less tomorrow
  - e.g. the 5:2 diet advocate this approach

- We study variation in individuals monthly shopping baskets and therefore we believe it is much less plausible that individuals make and execute plans over this longer time horizon
Empirical challenge in studying self-control

- Preference heterogeneity an important driver of both:
  - cross-sectional variation in choice behaviour
  - differences in how people respond to changes in their economic environment (i.e. price and income)

- The challenge is to separate evidence that is indicative of self-control problems from this preference heterogeneity
  - develop an empirically tractable approach that
    - encompasses standard economic models of consumer choice (prices and income), in addition to psychological factors related to temptation and self-control
    - allows us to quantify the extent of individual’s self-control problems
  - apply to longitudinal data on individuals’ grocery purchases
    - long time dimension help to allow for rich preference heterogeneity
Heterogeneity in healthy share, mean across people
Heterogeneity in healthy share, within person

(a) over months

(b) from Q1 to Q4
Two-selves model

- We model food purchase behaviour as a compromise between two-selves each with stable preferences over separate food and drink baskets
  - healthy self – preferences over fruit, vegetable, wholegrains
  - unhealthy self – preferences over soft drinks, crisps, confectionery
- Allocation of products between selves is endogenous and individual specific
Two-selves model

- Multi-selves models have a long history in the theoretical literature
  - Many of these multi-selves models deliver observationally equivalent predictions about choice behavior
  - but have importantly different implications for welfare
  - The main intuition that we take away from these multi-selves models is that consumer decision making can be thought of as a process in which two selves each have an influence on decision-making
Two-selves model

- Broadly this multi-selves literature models choice as a two-stage process:
  - First self (healthy self) chooses a menu from a set of available menus.
  - Second self (visceral self) chooses from that menu.

- Sophisticated consumers make choices that account for the choice of the visceral self, either through intrinsic self-control or by using commitment devices, leading to a sub-game perfect equilibrium.

- Naive consumers assume preferences in the second stage are the same as in the first, making dynamically inconsistent choices.

- Our definition of self-control is closely related to this literature:
  - An increase in the influence of the visceral (short-run) self in decision making indicates a failure to exert self-control.
Two-selves model

- Many food and drinks products
  - a set are healthy \((h)\) and a set are unhealthy \((l)\), available at prices \((p^h_t, p^l_t)\)
  - each period, \(t\), individual \(i\) chooses a vector of quantities, \((q^h_i, q^l_i)\)
- An individual is characterized by a healthy and an unhealthy self, each with (different) rational preferences
  - the two selves enter into a bargaining process that is different for every individual and may not be stable over time
  - heterogeneity in preferences (and other characteristics) mean individuals differ in the weight of the healthy self in bargaining
- a resolute individual has a stable bargaining process over time
- an individual with self-control problems will, from time to time, be tempted by their unhealthy self, and so experiences greater fluctuations in the bargaining process
Collective approach to choice behaviour

- We borrow from the collective household economics literature
  - the two selves enter into a bargaining process that results in a Pareto optimal outcome
  - to study choice we are agnostic about the specific interaction between the selves (to study welfare we need to introduce more structure)

- We make use of the sharing rule concept used to quantify the bargaining power of individuals in collective models
  - Chiappori, 1988, 1992; Browning and Chiappori, 1998; Chiappori and Ekeland, 2009; Dunbar, Lewbel and Pendakur, 2013; Browning, Chiappori and Lewbel, 2013
  - Chiappori (1988) shows that the sharing rule is compatible with assumption that both selves choose Pareto efficient allocations, so the sharing rule is a direct indication of the bargaining power of the selves
Two-selves model

- Individual $i$ in period $t$ chooses a vector of healthy and unhealthy products, $(q_i^h, q_i^l)$ to solve:

$$\max_{q_i^h, q_i^l} \mu_{it} u^{ih}(q_i^h) + (1 - \mu_{it}) u^{il}(q_i^l)$$

subject to

$$p_t^h q_i^h + p_t^l q_i^l \leq x_{it}$$

where $u^{ih}(.)$, $u^{il}(.)$ are stable well-behaved utility functions

- $\mu_{it}$, the Pareto weight that represents the bargaining weight of the healthy self in $i$'s optimisation problem in $t$

  - under assumption of bargaining between two selves, $\mu_{it}$ generally depends on prices $(p_t^h, p_t^l)$, food budget ($x_{it}$) and possibly “non-standard” factors $z_{it}$

  - $\mu_{it} = \mu_i \in ]0, 1[$ corresponds to standard rational choice model with strong separability between healthy and unhealthy products
An important feature of our model is that it incorporates elements of non-standard decision making, without abandoning the assumption of rational choice behaviour altogether.

Provides a useful way to quantify the influence of both selves.

- The Pareto weight, $\mu_{it}$, is monotonically related to the observed healthy self expenditure share $\eta_{it} = x_{it}^h / (x_{it}^h + x_{it}^l)$

$$\eta_{it} = \eta_i(p^h_t, p^l_t, x_{it}, z_{it})$$
Two-selves model

- To summarize the two-selves model
  - individual has a New Years resolution to eat more healthy food
  - in the first few months of the year this individuals food purchasing behavior is driven by the preferences of the healthy self
    - i.e. the Pareto weight $\mu_{it}$ that is attached to the healthy self is reflected by a high share of the healthy food items in the food budget
  - Over time her resolution deteriorates and she is increasingly tempted, reflected by a change in one of the elements of $z_{it}$
    - i.e. the Pareto weight $\mu_{it}$ decreases reflected by an increase of the share of unhealthy food items in the food budget
    - variation in prices and budgets might also lead to variation in the share of the budget spent on healthy foods
  - In the empirical analysis we isolate variation in the sharing rule that is not driven by variation in prices, budgets, or other observable factors, and link this to self-control problems
Revealed preference methods

- We evaluate the empirical fit of the two-selves model using revealed preference tools


  - for given data the Afriat Inequalities define a utility level $U^i_j$ and a marginal utility of income $\lambda^i_j$ (associated with the observed budget $x_{it}$ for each observed bundle $q^j_{it}$)

- Data, $S^{ij} = \{(p^j_t; q^j_{it}), t = 1, \ldots, T\}$ (where $j = h, l$), are rationalizable if they satisfy the Afriat Condition:

  - The set $S^{ij}$ satisfies the **Afriat Condition** if there exist numbers $U^i_j, \lambda^i_j \in \mathbb{R}_{++}$ that meet, for all observations $s$ and $t$, the **Afriat Inequalities**:

    $$ U^i_j - U^i_j \leq \lambda^i_j p^j_t (q^j_{is} - q^j_{it}). $$
Revealed preference methods

- Checking behavioural consistency with the two-selves model requires verifying the Afriat Condition for each self separately
  - the Afriat Inequalities are linear in the unknowns $U_{ij}^t$ and $\lambda_{ij}^t$
  - we can use standard linear programming techniques to verify rationalisability of self j’s behaviour for a given individual i
  - if the two selves pass this check, then we conclude that the individual behaves in terms of two selves maximising their stable rational preferences subject to their budget constraints
Data - households

- Kantar Worldpanel
  - select 3,645 single individuals
  - all grocery purchases for 24+ months; quantities and prices

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>95% C.I</td>
<td>Mean</td>
<td>95% C.I</td>
</tr>
<tr>
<td>Male</td>
<td>41.6</td>
<td>[40.0, 43.2]</td>
<td>43.3</td>
<td>[42.4, 44.2]</td>
</tr>
<tr>
<td>Age</td>
<td>62.7</td>
<td>[62.3, 63.2]</td>
<td>59.6</td>
<td>[59.3, 59.9]</td>
</tr>
<tr>
<td>SES: Highly skilled</td>
<td>16.0</td>
<td>[14.9, 17.2]</td>
<td>14.9</td>
<td>[13.9, 15.9]</td>
</tr>
<tr>
<td>SES: Semi-skilled</td>
<td>50.1</td>
<td>[48.4, 51.7]</td>
<td>54.6</td>
<td>[53.2, 55.9]</td>
</tr>
<tr>
<td>SES: Unskilled</td>
<td>33.8</td>
<td>[32.2, 35.4]</td>
<td>30.5</td>
<td>[29.2, 31.8]</td>
</tr>
<tr>
<td>N</td>
<td>3645</td>
<td></td>
<td>11628</td>
<td></td>
</tr>
</tbody>
</table>
Data - products

- Data is at the transaction level
  - approximately 600,000 barcodes (UPCs)
  - over 100,000 distinct products (by brand)
  - we aggregate to 85 products (milk, bananas, fresh fish, biscuits,...)
- includes nutrients

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy (kj/100g)</td>
<td>843.8</td>
<td>1.5</td>
<td>3524.7</td>
</tr>
<tr>
<td>Saturated fat (g/100g)</td>
<td>3.6</td>
<td>0.0</td>
<td>40.6</td>
</tr>
<tr>
<td>Sugar (g/100g)</td>
<td>9.0</td>
<td>0.0</td>
<td>81.4</td>
</tr>
<tr>
<td>Sodium (mg/100g)</td>
<td>0.2</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Fibre (g/100g)</td>
<td>1.8</td>
<td>0.0</td>
<td>10.3</td>
</tr>
<tr>
<td>Protein (g/100g)</td>
<td>7.7</td>
<td>0.0</td>
<td>25.6</td>
</tr>
</tbody>
</table>
We use the Nutritional Profiling Score to rank products in terms of healthiness.

- Aggregates individual nutrients into a single index.
  - Is what is used by UK regulators.
  - Gives positive points for saturated fat, sodium, sugar and calories.
  - And negative points for fibers, proteins and fruit, vegetables and nuts.
  - Ranges from -15 to 40.
  - Lowest scores: pulses and vegetables, with scores of around -10.
  - Highest scores: solid fats, chocolates and biscuits, score over 20.

We need to allocate products to healthy or unhealthy baskets, for each individual we choose the cut-off that fits the data best.
Endogenous cutoff, healthy v unhealthy foods
Fit of the two-selves model

- A natural alternative to the two-selves model is a single-self model, in which an individual has a single stable utility function defined over all 85 goods
  - the pass rate of the two-selves model is about twice as high as the pass rate of the single-self model
  - the distribution of the Afriat index for the two-selves model stochastically dominates that for the single-self model
- Almost 20% of individuals have observed purchase behaviour that is exactly rationalisable by the two-selves model
  - for the remaining individuals a small perturbations (1% on average) of the budget ensures purchase behaviour is rationalised by the two-selves model
Distribution of Afriat indices for two-selves model
Estimating the sharing rule

- The two-selves model gives a structural interpretation to the share of spending on healthy food via the sharing rule:

\[ \eta_{it} = \eta_i(p_{rt}^h, p_{rt}^l, x_{it}, z_{it}) \]

- this is a consumer-specific nonparametric function of 85 prices, income and \( z_{it} \)

- to implement empirically we make restrictions

1. we can split \( z_{it} = \tilde{z}_{it} + \epsilon_{it} \) observable (\( \tilde{z}_{it} \)) and unobservable (\( \epsilon_{it} \)) components, and unobservable is one dimensional and separable:

\[ \eta_{it} = g_i(p_{rt}^h, p_{rt}^l, x_{it}, \tilde{z}_{it}) + \epsilon_{it} \]

2. 85 prices can be approximated by two price indices: one for healthy foods, \( \Pi_{irt}^h \), and one for unhealthy foods, \( \Pi_{irt}^l \):

\[ \eta_{it} = g_i(\Pi_{irt}^h, \Pi_{irt}^l, x_{it}) + z_{it}. \]

3. \( g_i \) can be approximated by individual specific log linear function
Recovering the sharing rule

- We approximate the sharing rule with

\[
\eta_{it} = \alpha_i + \beta_i \ln \left( \frac{\Pi^l_{i rt}}{\Pi^h_{i rt}} \right) + \gamma_i \left( \ln rX_{it} \right) + \theta'_i Z_{it} + \epsilon_{it}
\]

- $\ln x_{it}$: log of real food expenditure; ensures the sharing rule is homogeneous of degree zero in prices and expenditure

- $Z_{it}$ includes
  - monthly advertising expenditure on unhealthy foods (confectionery, snacks, soft drinks, prepared and convenience foods), from the AC Nielsen Advertising Digest
  - weather conditions in local area: minimum temperature in that month, maximum temperature and rainfall
Estimates from the sharing rule

- Estimate for each individual

  - partial $R^2$: variation in the sharing rule explained by variation in the economic environment (relative prices and food budgets)
    - mean across individuals is 0.12 – on average, around 12% of the within-person variation in the sharing rule is explained by variation in price and food budgets

  - total $R^2$: variation in the sharing rule explained by variation in the economic environment and other factors such as advertising and weather
    - mean across individuals is 0.24 – on average, around 24% of within-person variation in the sharing rule is explained by prices, food budgets, advertising and weather

- individuals’ response to variation in these variables does explain a portion of the variation in the sharing rule, but a considerable fraction of the fluctuations in the share of spending on healthy foods over the year remains unexplained
Preferences for healthy food, mean sharing rule, $\hat{\alpha}_i$
Deviations from the mean sharing rule over months

Percentage point difference in sharing rule from January

-8 -6 -4 -2 0

Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

-8 -6 -4

Sharing rule: % spending on healthy foods
Sharing rule: effect of prices, budgets, advertising, weather removed
Standard deviation of the sharing rule and the residual sharing rule

- We construct the standard deviation of the residual sharing rule for each individual $i$ over $t$
  \[ \hat{\sigma}_i = \text{sd}(\hat{\epsilon}_{it}) \]

- We also construct the standard deviation of the sharing rule for each individual $i$ over $t$
  \[ \tilde{\sigma}_i = \text{sd}(\eta_{it}) \]

- which measures the total variation in the individual’s sharing rule, driven by both changes in the economic environment, weather, and advertising.
Standard deviation of the sharing rule and the residual sharing rule

Standard deviation of:
- Sharing rule: % spending on healthy foods
- Sharing rule: effect of prices, budgets, advertising, weather removed

Graph showing the standard deviation of the sharing rule and the residual sharing rule with histograms for different values of standard deviation.
### Variation with stated preferences, mean sharing rule $\hat{\alpha}_i$

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2) Mean</th>
<th>(3) Difference</th>
<th>(4) 95% CI for diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above median preferences for healthy food</td>
<td>52.7</td>
<td>-4.5</td>
<td>[-5.4, -3.5]</td>
</tr>
<tr>
<td>Below median preferences for healthy food</td>
<td>48.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above median preferences for processed food</td>
<td>48.9</td>
<td>3.4</td>
<td>[2.5, 4.4]</td>
</tr>
<tr>
<td>Below median preferences for processed food</td>
<td>52.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above median tendency to buy on promotion</td>
<td>50.3</td>
<td>0.6</td>
<td>[-0.4, 1.6]</td>
</tr>
<tr>
<td>Below median tendency to buy on promotion</td>
<td>50.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above median shopping commitment</td>
<td>50.8</td>
<td>-0.5</td>
<td>[-1.5, 0.5]</td>
</tr>
<tr>
<td>Below median shopping commitment</td>
<td>50.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above median stated self-control</td>
<td>52.4</td>
<td>-2.4</td>
<td>[-3.7, -1.0]</td>
</tr>
<tr>
<td>Below median stated self-control</td>
<td>50.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Variation with stated preferences, $\hat{\sigma}_i$

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2) Mean</th>
<th>(3) Difference</th>
<th>(4) 95% CI for diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above median preferences</td>
<td>7.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>for processed food</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below median preferences</td>
<td>7.48</td>
<td>-0.32</td>
<td>[-0.46, -0.18]</td>
<td></td>
</tr>
<tr>
<td>for processed food</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above median tendency to buy</td>
<td>7.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>on promotion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below median tendency to buy</td>
<td>7.39</td>
<td>-0.46</td>
<td>[-0.61, -0.32]</td>
<td></td>
</tr>
<tr>
<td>on promotion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above median shopping</td>
<td>7.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>commitment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below median shopping</td>
<td>7.78</td>
<td>0.27</td>
<td>[0.13, 0.41]</td>
<td></td>
</tr>
<tr>
<td>commitment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above median stated self-control</td>
<td>7.40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below median stated self-control</td>
<td>7.84</td>
<td>0.45</td>
<td>[0.25, 0.65]</td>
<td></td>
</tr>
</tbody>
</table>
Variation with demographics, mean sharing rule, $\hat{\alpha}_i$
Variation with age, $\hat{\sigma}_i$

(a) Standard deviation of residual sharing rule, $\hat{\sigma}_i$

(b) Effect of responses to changes in prices and food expenditure, $\tilde{\sigma}_i - \hat{\sigma}_i$

![Graph 1](image1.png)

![Graph 2](image2.png)
Variation with income, $\hat{\sigma}_i$

(a) Standard deviation of residual sharing rule, $\hat{\sigma}_i$

(b) Effect of responses to changes in prices and food expenditure, $\tilde{\sigma}_i - \hat{\sigma}_i$

![Graph showing variation with income](image-url)
Robustness – by age group

<table>
<thead>
<tr>
<th>Specification</th>
<th>Mean $\hat{\sigma}_i$</th>
<th>$\hat{\sigma}_i$ by age group</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>&lt;40</td>
<td>70+</td>
</tr>
<tr>
<td>Baseline</td>
<td>7.62</td>
<td>8.43</td>
<td>7.36</td>
</tr>
<tr>
<td>Controls for holidays, birthdays</td>
<td>7.19</td>
<td>7.88</td>
<td>6.97</td>
</tr>
<tr>
<td>Controls for four price indices</td>
<td>7.22</td>
<td>7.96</td>
<td>7.01</td>
</tr>
<tr>
<td>Instrumenting food expenditure</td>
<td>7.77</td>
<td>8.55</td>
<td>7.52</td>
</tr>
<tr>
<td>Exogenous cutoff</td>
<td>7.90</td>
<td>8.68</td>
<td>7.60</td>
</tr>
<tr>
<td>Individuals: two-self model is better</td>
<td>7.60</td>
<td>8.40</td>
<td>7.29</td>
</tr>
</tbody>
</table>
### Robustness – by income quintile

<table>
<thead>
<tr>
<th>Specification</th>
<th>Mean $\hat{\sigma}_i$</th>
<th>1st</th>
<th>5th</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>7.62</td>
<td>9.06</td>
<td>6.53</td>
<td>1.39</td>
</tr>
<tr>
<td>Controls for holidays, birthdays</td>
<td>7.19</td>
<td>8.57</td>
<td>6.15</td>
<td>1.39</td>
</tr>
<tr>
<td>Controls for four price indices</td>
<td>7.22</td>
<td>8.60</td>
<td>6.15</td>
<td>1.40</td>
</tr>
<tr>
<td>Instrumenting food expenditure</td>
<td>7.77</td>
<td>9.23</td>
<td>6.66</td>
<td>1.39</td>
</tr>
<tr>
<td>Exogenous cutoff</td>
<td>7.90</td>
<td>9.45</td>
<td>6.69</td>
<td>1.41</td>
</tr>
<tr>
<td>Individuals: two-self model is better</td>
<td>7.60</td>
<td>9.11</td>
<td>6.51</td>
<td>1.40</td>
</tr>
</tbody>
</table>
Summary

- Within-person variation in diet quality is important, and similar in magnitude to between person variation.

- The two-selves model fits the data well, and provides a structural interpretation of the data, allows us to quantify variation in the sharing rule.

- The sharing rule varies between and within people:
  - Lower income and younger individuals suffer more from self-control problems than higher income and older people.
  - But failure to account for responses to changing prices and food budgets leads to an overestimate of the age and income gradient in self-control.
  - Low income and young individuals tend both to face more variation in their economic environment (prices and budgets) and to adjust their food purchasing behaviour more strongly in response to these changes.
Further work

- By placing more structure on the form of demand we can use to study the impacts of policy and other counterfactual situations on welfare

- In ongoing work we compare
  
  - purchasing for immediate vs. future consumption
  - identification based on comparing behaviour of same individual when making purchases with different consumption horizons

- Preliminary estimates suggest
  
  - on-the-go purchases associated with sugar temptation
  - effect is increasing in consumer BMI
“Sugar consumption and temptation in the soda market”

Rachel Griffith, Martin O’Connell and Kate Smith
Motivation

- Different theories of the two-selves variety lead to different welfare implications

- Impose more structure, our model is closest to Gul and Pesendorfer (2001, 2004)

- We exploit novel data that records the same consumers making decisions for immediate and future consumption:
  - temptation is likely to be stronger when consumption is immediate

- We consider choices made over soda:
  - there are sugary and diet varieties of most brands, and consumers might be tempted by the former
Data on purchases for immediate and future consumption

We use longitudinal data on the purchases of soda made by a panel of British individuals from 2010–16

Data record two types of purchase:

1. “At-home”: for future consumption; recorded by the main shopper and consumed by the household

2. “On-the-go”: for immediate consumption; recorded by members of the household and consumed by that member
Data on purchases for immediate and future consumption

- 523 single households who are observed buying soda (major brands) at-home and on-the-go at least once over the time period
- on average, 240 observations per individual, 41 per individual–year
- in general, the single households are older than the rest of the population
- conditioning on soda purchasers means that our sample of single individuals has an age distribution that is similar to the age distribution of all main shoppers
Table 1.1: *Soda brands and products*

<table>
<thead>
<tr>
<th>Brand:</th>
<th>At-home</th>
<th>On-the-go</th>
<th>Product:</th>
<th>Sugar content (g/100ml)</th>
<th>At-home</th>
<th>On-the-go</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coca Cola</td>
<td>15.8</td>
<td>21.8</td>
<td>Coca Cola Cherry: Diet</td>
<td>0</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Coca Cola Cherry: Regular</td>
<td>11.2</td>
<td>0.7</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Coca Cola: Diet</td>
<td>0</td>
<td>10.2</td>
<td>13.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Coca Cola: Regular</td>
<td>10.6</td>
<td>4.4</td>
<td>6.1</td>
</tr>
<tr>
<td>Dr. Pepper</td>
<td>0.9</td>
<td>1.3</td>
<td>Dr Pepper: Diet</td>
<td>0</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dr Pepper: Regular</td>
<td>9.5</td>
<td>0.7</td>
<td>1.2</td>
</tr>
<tr>
<td>Pepsi</td>
<td>11.0</td>
<td>6.6</td>
<td>Pepsi: Diet</td>
<td>0</td>
<td>9.2</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pepsi: Regular</td>
<td>10.9</td>
<td>1.8</td>
<td>1.7</td>
</tr>
<tr>
<td>Fanta</td>
<td>1.1</td>
<td>1.7</td>
<td>Fanta: Diet</td>
<td>0</td>
<td>0.4</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fanta: Regular</td>
<td>7.8</td>
<td>0.7</td>
<td>1.5</td>
</tr>
<tr>
<td>Sprite</td>
<td>0.8</td>
<td>1.1</td>
<td>Sprite: Diet</td>
<td>0</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sprite: Regular</td>
<td>9</td>
<td>0.4</td>
<td>0.8</td>
</tr>
<tr>
<td>Lucozade</td>
<td>3.7</td>
<td>4.3</td>
<td>Lucozade Energy</td>
<td>11.8</td>
<td>2.3</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lucozade Sport: Diet</td>
<td>0</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lucozade Sport: Regular</td>
<td>3.8</td>
<td>0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>Red Bull</td>
<td>0.4</td>
<td>1.0</td>
<td>Red Bull: Diet</td>
<td>0</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Red Bull: Regular</td>
<td>10.9</td>
<td>0.3</td>
<td>0.8</td>
</tr>
<tr>
<td>Outside Option</td>
<td>66.4</td>
<td>62.1</td>
<td>Outside Option</td>
<td>0</td>
<td>66.4</td>
<td>62.1</td>
</tr>
</tbody>
</table>

**Notes:** Numbers calculated using the sample of 523 individuals. Columns (2) and (3) shows the market share of each brand, listed in column (1), for at-home and on-the-go consumption. Column (5) shows the sugar content of the product listed in column (4). Columns (6) and (7) show the market share of each product, when purchased for at-home and on-the-go consumption.
Sugar content of at-home and on-the-go purchases

![Graph showing cumulative distribution of sugar content in purchases. The graph compares on-the-go purchases (blue line) and at-home purchases (red line). The x-axis represents sugar content of purchase (g/100ml), ranging from 0 to 10, and the y-axis represents cumulative distribution ranging from 0 to 1. The graph indicates that at-home purchases tend to have lower sugar content than on-the-go purchases.]
Consumer choice model with temptation

- Gul and Pesendorfer (2001, 2004) develop a framework to model consumers who may suffer from temptation when making decisions.

- Let $i$ index consumers; $t$ index purchase occasions; and $j$ index products.

- Define the following utility function:

$$U_{ijt} = u_{ijt} + v_{ijt},$$

$u_{ijt}, v_{ijt}$: von Neumann-Morgenstern utility functions;

- consumer does not suffer from temptation: $\max_j u_{ijt}$

- consumer does suffer from temptation: $\max_j U_{ijt}$
Random utility model

- We incorporate this idea in a random utility model to estimate empirically:
  - consumers derive utility from product characteristics
  - products contain a “tempting” characteristic: sugar

- On each purchase occasion consumers choose one product or decide not to buy a soda:
  - consumers suffer less from temptation when buying for future consumption: \( \max_j u_{ijt} \)
  - than when buying for immediate consumption: \( \max_j U_{ijt} = u_{ijt} + v_{ijt} \)
Empirical specification
Assume that indirect utility is quasilinear in the numeraire good and:

\[ u_{ijt} = \alpha_i (y_i - p_{jt}) + f(z_j, x_{jt}; \theta_i) + \epsilon_{ijt} \]

\[ v_{ijt} = \gamma_i Z_j \]

where \( \alpha_i \) is the marginal utility of an extra dollar spent on soda; \( \theta_i \) are preferences; and \( \gamma_i \) is temptation parameter, and:

\[ f(z_j, x_{jt}; \theta_i) = \beta_i z_j + \xi_{ibst} \]

Notation:

- \( y_i \): income of consumer \( i \)
- \( p_{jt} \): price of product \( j \) at time \( t \); \( x_{jt} \): other product characteristics
- \( z_j \): sugar content of \( j \)
- \( \epsilon_{ijt} \): type I extreme value iid demand shock
Identification

The unobserved brand effect, $\xi_{ibst}$, contains:

- brand effects
- segment–retailer and segment–quarter effects

which control for unobserved brand characteristics, and retailer and time varying demand shocks that may be correlated with price.

Identification of temptation parameter, $\gamma_i$:

- within segment variation in choices for sugary and diet varieties across purchases made for immediate and future consumption

We can control for individuals’ exposure to advertising of products.
Preference heterogeneity

We allow for unobserved preference heterogeneity by specifying a distribution of random coefficients:

- random coefficients on: price, sugar, and brand effects
- goal of estimation is to recover parameters governing distribution

We are interested in how the temptation parameter varies with individuals’ characteristics:

- allow $\gamma_i$ to vary by gender and SES
### Price elasticities

<table>
<thead>
<tr>
<th>Brand</th>
<th>At-home</th>
<th>On-the-go</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coca Cola</td>
<td>-2.3</td>
<td>-3.4</td>
</tr>
<tr>
<td>Dr. Pepper</td>
<td>-2.2</td>
<td>-3.7</td>
</tr>
<tr>
<td>Pepsi</td>
<td>-1.5</td>
<td>-3.2</td>
</tr>
<tr>
<td>7 Up</td>
<td>-1.9</td>
<td>-3.2</td>
</tr>
<tr>
<td>Fanta</td>
<td>-2.0</td>
<td>-3.5</td>
</tr>
<tr>
<td>Schweppes</td>
<td>-1.6</td>
<td>-3.6</td>
</tr>
<tr>
<td>Shloer</td>
<td>-3.7</td>
<td></td>
</tr>
<tr>
<td>Sprite</td>
<td>-2.2</td>
<td>-3.6</td>
</tr>
<tr>
<td>Tango</td>
<td>-1.4</td>
<td>-3.1</td>
</tr>
<tr>
<td>Irn Bru</td>
<td>-2.0</td>
<td>-3.4</td>
</tr>
<tr>
<td>Lucozade</td>
<td>-2.8</td>
<td>-3.6</td>
</tr>
<tr>
<td>Red Bull</td>
<td>-5.5</td>
<td>-4.8</td>
</tr>
</tbody>
</table>
Temptation parameter

<table>
<thead>
<tr>
<th></th>
<th>WTP for sugary varieties</th>
<th>Additional WTP when tempted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>7p</td>
<td>34p</td>
</tr>
<tr>
<td>Women</td>
<td>-4p</td>
<td>25p</td>
</tr>
<tr>
<td>Low SES</td>
<td>12p</td>
<td>31p</td>
</tr>
<tr>
<td>High SES</td>
<td>-36p</td>
<td>22p</td>
</tr>
</tbody>
</table>

- willingness to pay (WTP) for sugary varieties $= \frac{\beta_i}{-\alpha_i} \times 10$
- additional willingness to pay (WTP) when tempted $= \frac{\gamma_i}{-\alpha_i} \times 10$

expressed in pence relative to average price/L of $1.$
Summary and next steps

We use novel data to investigate the extent to which consumers are “tempted” by more sugary varieties when buying soft drinks.

Show variation in the degree of temptation by individual characteristics.

Next steps include enriching the model

▶ e.g. how might advertising affect individuals temptation?
Advertising exposure

We have detailed data on all adverts of soft drinks shown on TV:

▶ know the time, channel, duration, show pre/post

We have information on TV viewing habits of the main shopper:

▶ can combine this with the advertising data to construct an individual specific measure of advertising exposure

We can use this to study:

▶ how advertising affects demand for different products
▶ whether it alters people’s preferences for sugar
Inattention and advertising

- Firms might seek to exploit consumers’ biases
- How does advertising affect consumer decision making?
  - we estimate demand for potato chips, and show that advertising shifts consumers’ willingness to pay for healthier potato chips

<table>
<thead>
<tr>
<th>Advertising level</th>
<th>None</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>willingness to pay for healthier product, % of mean price</td>
<td>1.6 [1.2, 2.0]</td>
<td>-0.2 [-0.4, 0.2]</td>
<td>-1.5 [-1.8, -1.1]</td>
</tr>
</tbody>
</table>

*numbers in [] are confidence intervals*

* Dubois, Griffith and O’Connell (2017)*
Overall Summary

- Considerable evidence of self-control problems
- Relatively little known about how large, what variation across individuals, what drives variation, ...
- More structure needed to make welfare interpretations
- Promising and important area of future research