Daily stressors and Food Choices: A Lab Experiment with Low-SES Mothers*

Nicolai Vitt ⓡ[1] (University of Bristol)

Jonathan James ⓡ (University of Bath)

Michèle Belot# ⓡ (Cornell University)

Martina Vecchi ⓡ (Pennsylvania State University)

Corresponding Author: Email: mb2693@cornell.edu

March 2021

Abstract

We investigate experimentally the effects of daily-like stressors on immediate and planned food choices, in a sample of low socioeconomic status (SES) mothers. We design a novel stress protocol that aims to mimic everyday stressors experienced by low socioeconomic status individuals. The protocol consists of budget and time allocation tasks to be performed under time and financial pressure. Immediate consumption is measured with in-laboratory consumption of low calorie and high calorie snacks; planned consumption is measured with an incentivized food shopping task. We find no evidence of a significant effect of the stressor on planned food consumption. We do find a notable increase in high-calorie snacking following the stress protocol but it is not precisely estimated. Overall, we find little support for the hypothesis that daily-life stressors induce unhealthy food choices.

JEL Classification: I12, D91

Keywords: Diet, Acute stress, Daily stressors, Lab experiment

---

*This project is funded by the European Union's Seventh Framework Programme, under grant agreement no. 607310. Nicolai Vitt further gratefully acknowledges funding from the Economic and Social Research Council, award 1651922. We thank EssexLab and the Università degli Studi di Firenze for allowing us to use their facilities, and Daacro Lab for the analysis of the saliva samples. We are grateful to Anne Banks, Rossella Corridori and Margherita Fabbri for their assistance with this project. We furthermore thank Ioan Balaban, Francesco Colombo, Nina Danilova, Rebecca Ravalli, Chiara Santantonio, Anamika Shaivya, Dragos Gruia, Iulia Popescu, Michael Zelenka, Roxana Dumitrescu, Shiobhan Low for their excellent research assistance. We thank Pablo Brañas-Garza, Francesca Cornaglia, Steven Dieterle, Daniela Glätzle-Rützler, Lorenz Götte, Silvia Maier, Mirko Moro, Arthur Schram, Daniel Schunk, Christiane Schwieren, Trent Smith, Andreas Steinhauer, as well as seminar participants at the Nudge-It meetings, the SGPE, SES, ESPE and EALE conferences, the Economics at Panmure House workshop, and at seminars in Edinburgh, Heidelberg, York and the European University Institute for their helpful comments.

[1] Author order was randomized using the AEA Author Randomization Tool - Date entered: 2020-10-09 07:37:35 Confirmation code: TTLWqjksCT8A

1 Introduction

In most developed countries, lower socioeconomic status (SES) groups tend to have a poorer diet and an unhealthier lifestyle (Pampel et al., 2010). A number of explanations have been proposed to explain why that is the case, such as income, relative prices of healthier and unhealthier foods, maternal employment, and technological change (see Cawley, 2011, for a review), or more recently behavioural anomalies, such as present bias (Loewenstein et al. 2012). Although these factors appear to play a modest role, there remains a large unexplained component for the socio-economic gradient in health behaviours.

In this paper, we investigate with a laboratory experiment the effects of temporary everyday-like stressors[2] on food choices of low SES individuals. Psychologists document that low SES individuals are more likely to experience both chronic and acute stressors, and perceive them as more severe (Grzywacz et al. 2005; Almeida et al. 2005, Adler et al. 1994). The famous epidemiological studies on UK public servants (Marmot et al., 1978, and Marmot et al., 1991) were among the first to show that individuals at the lower end of the hierarchy are more affected by high blood pressure and heart conditions, indicators usually associated with greater exposure to stress. Therefore, a plausible pathway for the socioeconomic gradient in health-related behaviours, and eating in particular, is stress. Several studies have provided evidence of associations between perceived stress and unhealthy eating behaviours (Richardson et al., 2015; Barrington et al., 2014 and Sims et al., 2008). These studies are correlational however, hence our study aims to shed light on the causality of these associations using a controlled experiment.

We focus here on how everyday-like stressors (daily hassles), encountered regularly, may affect food choices. In a low SES population, examples include how to spend and allocate a limited budget, how to meet the demands at work and at home, arguments, etc. We are interested in the impact of these daily stressors on low SES *mothers* in particular, since they play a crucial role in families' dietary choices, often being in charge of the family's food shopping and meal preparation (Harnack et al., 1998).

Why and how could daily stressors affect food choices? According to biologists, a stressor is a physical or psychological stimulus "that perturbs or threatens to perturb homeostasis" (Sapolsky, 2004). Depending on its duration, stress is classified as acute or

[2] An acute stressor is a short-term demand or pressure placed on the body. It can be of a physiological or psychological nature. In the following we use acute and short-term interchangeably when describing stressors and stress.

chronic. Daily stressors are a form of acute stress. Acute stressors require immediate attention and therefore shift resources (cognitive and/or physical) to dealing with the stressor. Consequently, individuals may pay less attention to other decisions (Allen and Armstrong, 2006). Stress has been found to causally affect a variety of economic behaviours and decisions (albeit induced using severe stressors), including temporarily altering time preferences and risk attitudes and thereby affecting the ability to make utility maximizing decisions (Delaney et al., 2014; Kandasamy et al., 2014; Buckert et al., 2017a; Buckert et al., 2017b; Buser et al., 2017; Zhong et al., 2018; Goette et al., 2015). Therefore, we conjecture that daily stressors could impact food choices by shifting the focus of cognitive resources to deal with the stressor and , thereby leading to more habitual behaviour, such as buying and consuming well known foods, and/or more impulsive choices. In a low-SES population both habitual and impulsive food choices are likely less healthy than choices made after thorough consideration.

A second reason why daily stressors could affect food choices is a more direct biological mechanism. The literature in biology suggests that short-term stress may affect the desire to eat as well as the choice of which foods to eat. Hormonal responses to a stressor have been frequently cited to cause cravings for energy-dense "comfort foods" and hence a temporary change in food preferences (e.g. Adam and Epel, 2007). Both short-term and chronic stress stimulate the release of cortisol (in humans) or of other glucocorticoids (in animals) which has been shown to affect food intake of rats (Zakrzewska et al., 1999; Dallman et al., 2004) and humans (Tataranni et al., 1996; George et al., 2010) when administered exogenously.

To study the impact of daily stressors on food choices, we propose a novel protocol aimed at mimicking this type of stressors in a laboratory setting. The protocol consists of time and budget allocation tasks, to be performed under time and financial pressure, and involving social incentives and distractions. The experiment involves 196 low-SES mothers, living in the area of Colchester in the UK. Half of the participants were assigned to a stress condition and the other to a control group. Participants in the control group were asked to also complete a cognitive task (reading short texts and answering simple questions), but without financial incentives, distractions or time pressure.

We evaluate the impact of the stress protocol both on immediate and planned eating choices. Immediate eating choices are measured by the in-laboratory consumption of high-calorie (muffins) and low-calorie (apple slices) snacks made available to the participants directly on their computer desk. Planned eating is measured based on incentivized choices in a "virtual supermarket", a computer-based tool with similar features to online supermarket platforms. Shopping choices require planning of future consumption and often involve larger choice sets, therefore impaired decision making (first channel) would be expected to affect these choices more than the less complex immediate consumption choices. In contrast, we would expect the biological mechanism to play a larger role in immediate food choices.

To further explore the role of impaired decision making under stress, we introduce a second experimental variation: the choice environment in the food shopping choice was either "simple" or "complex". Both the stress and the control groups were split into two further experimental conditions corresponding to the different choice environments. In the simple choice environment, items were displayed in 10 different categories (e.g. fruit, vegetable, eggs & dairy etc.). In the complex choice environment, items were displayed in a long list, grouped by category but without labelling of categories. If stress affects dietary choices by impairing individuals' decision-making, a more complex choice environment is expected to lead to less healthy food shopping choices under stress.

The first question we address is whether our protocol was successful, in the sense of replicating the type of "daily hassles" mothers could encounter. We would not expect "daily hassles" to trigger a large physiological response.[3] In our pre-analysis plan, we state that our primary outcome measure for assessing the stress response is the participants' subjective evaluation of how stressful the task was. We find that participants in the stress condition perceived the task as significantly more stressful than the control group. We also find evidence for a significance rise in their heart rate during the task, but not of their cortisol level, which indicates that the protocol was successful in inducing a mild stress response, consistent with what we would expect for daily hassles. Regarding the impact on eating choices, we find no statistically significant effects of this stressor on participants' food shopping choices or their immediate food intake. If we restrict the analysis to the subgroup of participants showing the largest cortisol response to the stressor, we still find no significant effects on behaviour. For the immediate food choice, the point estimates show a notable increase in high-calorie snacking following stress, but smaller than effects previously reported for severe stressors. For the planned food choice, the estimates provide evidence against large stress-induced increases in purchased calories, saturated fat and sugar. Our results suggest that daily-like stressors do not strongly affect planned dietary choices, while they may have a small (but quantitatively

[3] See Pre-Analysis Plan (page 19), we state that the effects on heart rate and cortisol are difficult to observe (as the timing of the measurements is of crucial importance and there are many factors affecting these physiological responses). Given that we did not aim at inducing a large stress response, we did not expect to be able to capture significant changes in heart rate or cortisol.

meaningful) impact on immediate snacking choices. When controlling for performance in the stress task, we find a higher intake of high-calorie snacks among those performing poorly, suggesting a combined role of stress and failure in eating behaviour.

A final take-away message from our study is related to statistical power. Stress responses (even to severe stressors) are known to be very noisy, so it is statistically challenging to detect them and, by the same token, to detect effects they may have on behaviour. This may be a reason why prior studies on the effects of stress on eating choices resort to rather severe stress-inducing protocols, which are not representative of stressors frequently encountered in everyday life. Since our research question relates to everyday-like mild stressors, by definition such responses are even harder to detect: an increase of 5% or 10% in daily calorie intake induced by daily stressors matters if added up over time. Based on our (ex post) power calculations, we calculate that the minimum sample size to significantly detect quantitatively meaningful effects of around 10% of the control group mean with 80% power would be of the order of thousands of participants. Thus, one additional take-away message from this paper is that the samples needed to study the effects of mild stress are very large, perhaps even unrealistically large.

This paper is structured as follows. Section 2 presents the experimental design. Section 3 describes the data collected during the experiment. Section 4 outlines the hypotheses tested. In section 5, we present our empirical analysis and our results. In Section 6, we discuss the external validity of the study as well as statistical power to detect quantitatively meaningful effects. Section 7 concludes.

2 Experimental Design

The study was conducted between 15 October and 19 October 2018 at the experimental laboratory of the University of Essex (EssexLab). Sessions lasted approximately two hours and started at 10:30 am, 2:00 pm and 5:00 pm.

The experimental design was pre-tested in June 2018 using a sample of 50 low-SES mothers in Florence, Italy.[4]

2.1 Sample and Randomization

We conducted a lab experiment with 196 participants,[5] recruited with the help of a marketing agency. They were recruited in the area of Colchester.[6] The specific eligibility criteria for participation in the study were: Aged between 18 and 45; Fluent in English; Being a mother whose youngest child is aged between 2 and 12 years old; Having a net annual household income below £35,000;[7] Not in a possession of a university degree and not currently enrolled at university; Not been pregnant in the past 6 months; Not having allergies or intolerances to foods used for the snack consumption choice; Not having any medical conditions which can affect diet.[8]

Those interested in participation were invited to complete an online screening questionnaire or to contact the experiment team by telephone. Eligible mothers were then invited to one of the experimental sessions; they received an information leaflet and a consent form by post. 31 of the 227 participants who signed up did not fulfil all eligibility criteria when answering the questions about eligibility on the day of the experiment and are hence excluded from the analysis, a sample of 196 low-SES mothers remains for our analysis, see Table 1 for details.

We conducted 15 experimental sessions with 13 to 18 participants per session, spread over a period of five days. The experiment follows a 2x2 experimental design resulting in four experimental conditions, pre-assigned at the session level to ensure balance in terms of day of the week and time of day:

[4] The pre-test was conducted with ethical approval by the European University Institute and the University of Edinburgh and was pre-registered in the AEA RCT registry under the following trial ID: AEARCTR-0003089. Details can be found under https://www.socialscienceregistry.org/trials/3089/history/30976. Initially this pre-test was planned as the main experiment, but recruitment of participants proved too difficult to reach the necessary sample size.

[5] The lab experiment was conducted with ethical approval by the European University Institute and the University of Edinburgh. The experiment and the hypotheses tested in this study and in Belot et al. (2020b) were pre-registered in the AEA RCT registry under the following trial ID: AEARCTR-0003410. Details can be found under https://www.socialscienceregistry.org/trials/3410/history/35937.

[6] A marketing agency sent personalized letters to women in the Colchester area who matched our age restriction and lived in low SES neighbourhoods. The study was furthermore promoted to the participants of a previous experiment. Examples of the recruitment materials used to advertise the study can be found in Appendix A.

[7] This threshold is based on the median disposable household income of households with at least one child in the UK in 2015/16 of approx. £ 34,742. (Data Source: Office for National Statistics, 2017. Effects of Taxes and Benefits on Household Income, 2015-2016.)

[8] Specifically, we asked participants whether they have diabetes, an eating disorder or a metabolic disorder.

1) Stress Task & Simple Shopping Task
2) Stress Task & Complex Shopping Task
3) Control Task & Simple Shopping Task
4) Control Task & Complex Shopping Task

Participants were asked to indicate their preferred session slots, and were assigned to one of the slots solely based on scheduling concerns.

2.2 Procedure

A timeline of the experimental sessions is shown in Table 2.[9] Participants were asked not to consume any alcohol for 12 hours before the experiment, nor any product containing chocolate or cocoa for 6 hours before, nor any food or drink in the last hour before the experiment. Upon arrival at our lab facilities, participants' body weight and height (without shoes and heavy clothing) were measured by trained lab assistants. Throughout the experimental session, participants were asked to wear an armband monitoring their heart rate using an optical sensor. At the beginning of the experimental session, participants were asked to provide a first saliva sample (9 min before the start of the stress/control task).

Following this, participants were asked to complete a 10-minute task. The nature of the task depended on the session's randomly assigned experimental condition:

- In conditions 1 and 2 (stress conditions), participants were asked to complete an incentivised task aimed at inducing mild stress.
- In conditions 3 and 4 (control conditions), participants were asked to complete a task of a similar nature but with no stress inducing features.

Detailed descriptions of these tasks can be found below.

Following the first task, participants were asked to complete a food shopping task. They were given a fixed budget of £30 to purchase grocery items in a "virtual supermarket", a computer-based tool similar to online supermarket platforms. The complexity of the food shopping environment depended on the experimental condition assigned to the session:

[9] The full experimental instructions given to participants can be found in Appendix B.1.

- In conditions 1 and 3 (simple shopping task), products were listed separately in 10 different "logical" food categories such as fruit, vegetables, eggs & dairy, meat & fish etc.
- In conditions 2 and 4 (complex shopping task), products were shown in a long list without a clear listing logic.

Details of these food shopping tasks are outlined below.

After the food shopping task, participants were asked to provide a second saliva sample (29 mins after the start of the stress/control task) and then given a five-minute break. After the break, participants were asked to complete a questionnaire on demographics, family characteristics and behaviours, which might affect cortisol levels. During the break and the time given to complete the first questionnaire, participants were given permission to consume the snacks provided on their desks: high-calorie blueberry mini-muffins and low-calorie apple slices (not labelled with their calorie content or in any other way). After 20 minutes, the bowls of snacks were collected.[10]

Participants were then asked to complete a second questionnaire about food consumption and food preferences of the participant and their youngest child as well as the participant's food consumption during pregnancy. The questionnaire furthermore included questions about the stressfulness of the stress/control task, chronic stress, participants' coping behaviours when dealing with stress, and about potentially stressful events during the last 3 months as well as during the pregnancy. The data collected in this questionnaire is used in Belot et al. (2020b) to examine the link between chronic maternal stress during pregnancy and children's food preferences.

At the end of the experimental session, a final saliva sample was collected (85 mins after the start of the stress/control task). Before receiving their payment, participants were told that the snacks provided differed in calorie content (at the request of the ethics committee).

2.3 Stress Protocol

The goal of this study is to examine the effects of exogenously induced short-term mild stress on dietary choices mimicking the type of stressors low-SES mothers are likely to encounter in their daily lives.

[10] Six blueberry mini-muffins (mean weight: 136.3g) and approx. 160g of apple slices (mean weight: 158.6g) were provided for each participant. The total costs of the snacks provided were £0.75 for the muffins and £1.00 for the apples.

There exist a number of experimental studies examining the impact of stress on eating. These studies use stress protocols aimed at triggering marked physiological responses (for example, the Trier Social Stress Test by Kirschbaum et al., 1993, or arithmetic exercises), which make it difficult to extrapolate to real life situations low-SES individuals may face. Physiological responses to daily hassles are likely to be milder, which calls for a different protocol. Moreover, none of these studies are conducted in a population of low-SES individuals.

We propose a protocol aimed at mimicking the kind of stressors mothers are frequently exposed to in real life. We ask mothers to solve a series of time and money budgeting tasks, choosing the cheapest or the most time efficient option amongst all, as often required in real life. The protocol incorporates several features that have been shown to induce stress: financial incentives and losses, time pressure, social pressure and distractions.

Main task

The main task is computer-based and consists of a block of 15 short decision tasks.[11] Participants were given 10 minutes to complete as many tasks as they could, and faced time limits of 120 seconds for each of the 15 tasks visualized by a prominently placed countdown timer. The overall time constraint was expected to be binding for most of the participants and hence to induce time pressure.[12]

The decision tasks comprised of budget tasks and time management tasks, to reflect the often limited financial and time resources among low-SES mothers. No food products or other health-related items were included in the budget tasks, to avoid any priming effects. For the budget tasks, participants were asked to choose the cheapest way to purchase a given basket of household expenditure items from a list of options. For example, participants were asked to purchase five t-shirts choosing from a list of t-shirts, which included single items as well as value packs consisting of multiple items. For the time management task, participants were given a list of diary items and were asked to schedule these in a timetable provided. The items to be scheduled were of different lengths and a variety of constraints needed to be considered when scheduling them: some items needed to be scheduled at a specific time or within some given time window.

Distraction

[11] Sample screenshots of the stress tasks are shown in Figures B.2 – B.5 of Appendix B.2.
[12] 64.5% of participants in the stress group did not complete all 15 tasks.

Short incentivised pop-up knowledge questions unrelated to the main task appeared on screen throughout the course of the task block at pre-specified times unknown to participants. Buckert et al. (2017a) show that distractions induced via pop-up Stroop-tasks raise physiological stress levels of subjects.

Incentives

While the tasks were completed individually, incentives were based on the joint performance of social groups, to elicit stress through social pressure (Babcock et al., 2015; Kirschbaum et al., 1993; Dickerson and Kemeny, 2004). Each group consisted of two randomly matched participants in the same session. Group matching was anonymous, for ethical reasons. Incorrect answers and incomplete tasks were penalized, since uncertain financial pay-offs such as in economic competitions have been shown to induce stress (Buckert et al., 2017b; Buser et al., 2017; Zhong et al., 2018).

Each group was initially allocated £30 in the stress task block. The performance of each group member in the decision tasks and the pop-up knowledge questions determined how much of the initial £30 was "lost" by the group. Each participant could lose a maximum of £15 for the group, £13.50 from the decision tasks and £1.50 from the pop-up knowledge questions.[13]

Participants could incur time penalties, inducing stress through further time pressure (Kocher et al., 2013; Buckert et al., 2017a). Incentives were framed in terms of "losses" rather than "gains" to avoid inducing positive emotions.

The incentive structure ensured that participants' performance in every single task and pop-up question would affect the group's pay-off. This reduced the risk of participants giving up due to difficulties in solving some of the tasks. Participants were made aware of the joint incentive structure and that they were part of a group with another participant in the same session. However, the group assignments were not announced to the participants.

As mentioned above, this protocol was designed to mimic mild stressors often experienced by low-socioeconomic mothers: making decisions with consequences for others (e.g. for the family) subject to financial and time constraints as well as distractions (e.g. by children requiring attention). Child- and household related stressors, financial concerns as well

[13] In each of the 15 decision tasks, a participant could lose up to £0.90 to the group. There was no loss if the correct answer was submitted within 75 seconds of starting a decision task. If a correct answer was given more than 75 seconds after starting a task, £0.30 was lost. If a wrong answer was given or a task was not attempted or completed, £0.90 was lost. Each of the 10 pop-up knowledge questions was worth £0.15. If a participant gave a correct answer, there was no deduction. If a participant gave a wrong answer, £0.15 was lost to the group.

as time pressure are among the most frequently encountered and most severe daily hassles facing mothers with young children (Chamberlain and Zika, 1990; Olson and Banyard, 1993).

2.4 Control task

Participants in the control group were asked to complete a task which was comparable in length and of a similar nature, but was not aimed at inducing stress. Specifically, they were asked to answer 14 simple knowledge questions after reading seven short texts about a variety of topics.[14] The correct answers to each question could be found in the corresponding text. The questions were similar to those asked via pop-ups during the stress task. Participants were given 10 minutes for this task, there were no consequences from not completing all questions. The task was not incentivised and no "social groups" were formed.

2.5 Food Shopping Task

In both experimental conditions, participants were asked to complete a food shopping task. They were given 10 minutes[15] to allocate a fixed budget of £30 to food and drink items offered in a "virtual supermarket", a computer-based tool similar to online supermarkets adapted from a tool by Spiteri et al. (2019).[16] A variety of low-calorie and high-calorie food and drink items was available to choose from with prices matching market prices at a local supermarket. Participants were encouraged to make their shopping choices as they would during a weekly shop at their local supermarket.

The supermarket choice was incentivised to motivate participants to make choices representative of their normal shopping behaviour. 1 out of 15 participants was randomly selected to receive their chosen basket delivered to their home approximately two weeks after the experimental session, to ensure that current grocery stocks at home would not affect their choices. If participants had not spent the entire £30 budget, they were paid the difference in

[14] A sample screenshot of a control task is shown in Figure B.6 of Appendix B.2.

[15] In a related study using the same food shopping task with the same budget Belot et al. (2020a) found that 80% of those who had 10 minutes to complete the shopping task reported having just the right amount of time, with 20% reporting that they had too much time. None of the participants in that study reported having too little time. In this study we did not ask participants if they had enough time, but given the results in Belot et al. (2020a) and the observing of participants in this study this time constraint was not binding.

[16] Sample screenshots of the food shopping task are shown in Figures B.7-B.9 of Appendix B.2.

cash up to a maximum of £2 to avoid pressure and discourage non-representative shopping choices aimed at spending exactly £30.[17]

To examine whether choice complexity leads to less healthy decisions under stress, sessions were pre-assigned to one of two supermarket choice environments (independently of the assignment to the stress or control group): a simple or a complex choice environment. In both choice environments, 156 grocery items from the following 10 different product categories were on offer: fruit, vegetables, eggs & dairy, meat & fish, bakery, pasta & rice, pantry, snacks, ready meals, drinks. In the simple choice environment, items were displayed on 10 different pages – one for each product category. In the complex choice environment, items were displayed on a single page, grouped by category but without labelling of categories. The order in which items were displayed within each category was randomized at the participant level to avoid order effects. Furthermore, the display order of categories and the first category shown when opening the supermarket tool was randomized.

2.6 Monetary Compensation

Participants were informed they could receive a compensation of £60 to £75. Participants in the control group (conditions 3 and 4) received a compensation of £60, participants in the stress group (conditions 1 and 2) received an average compensation of £67.86. Participants were only informed of their exact compensation at the end of the session to avoid any wealth effects between groups and during the immediate or planned food choices. Participants in the stress group did not receive any feedback on their (or their group's) performance in the stress task during the experiment. Of the 227 participants, 16 additionally received the food basket they selected during the food shopping task, worth up to £30.

2.7 Power calculations

We use the final sample size of 196 eligible participants and the mean, standard deviation and intraclass correlation coefficient (ICC) observed in the control group for our primary outcome variables to ex-post calculate the minimum detectable effects (MDE) of our experiment, taking the potentially clustered error structure at the session level into account.[18] Our calculations show that at a significance level of 5% comparisons between stress and control group have 80% power to detect effects of 0.43-0.54 standard deviations in the snack choice (25-49% of

[17] Under this incentive scheme, it was optimal for participants to aim to spend between £28 and £30. 97.5% of participants spent between £28 and £30 in the food shopping task.

[18] Details of our ex-post power calculations can be found in Appendix C.

the control group mean) and effects of 0.43-0.45 standard deviations in the food shopping choice (10-19% of the control group mean). These MDE ranges refer to the different outcome variables in our study, the exact MDE differ between the outcome variables of our study since the session-level clustered error structure as captured by the ICC differs between the outcome variables. The sample used in our experiment, which is larger than the samples of most previous studies of short-term stress and eating,[19] is therefore sufficient to detect sizable effects on snacking similar to those reported in Epel et al. (2001) and Habhab et al. (2009). However, the effects of mild everyday stressors are expected to be smaller than those found for severe stressors. We will discuss this in light of our analysis and results.

3 Measures

We have collected measures relating to the experience of the experimental treatments and the dietary decisions made by participants in the lab, as well as a range of control variables. In the following, we describe the measures used to answer our research questions.

3.1 Food Choice

a) Immediate food consumption

Immediate food consumption is measured by the snack choice faced by participants. For a duration of 20 minutes, participants were permitted to consume the snacks provided on their desks: high-calorie blueberry mini-muffins and low-calorie apple slices. Snacks were weighed before and after the experiment, the consumption quantities of each snack type (in grams) are the primary outcomes relating to the snack choice. Secondary outcomes are the total calories (in kcal), saturated fat (in grams) and sugar content (in grams) of the consumed snacks; these measures are deterministic functions of the two primary outcome variables.

b) Planned food consumption

Planned food consumption is measured by the food shopping choice made using a "virtual supermarket" tool (Spiteri et al., 2019). We construct measures of the nutritional content of the baskets based on the nutritional information of each of the products chosen. Primary outcomes are the energy (in kcal), the saturated fat (in grams) and the sugar content (in grams) of the

[19] Table D.1 in Appendix D provides an overview of the previous literature on short-term stress and eating.

chosen basket. The total weight of fruit and vegetables (in grams) chosen by the participant is considered as a secondary outcome.

3.2 Measures of stress

Given the focus of the study on daily hassles, we expect the stress response to be mild, and therefore to not necessarily trigger a detectable physiological response. For this reason, we pre-specified[20] the primary and second measures of stress as follows: The primary measure is a self-reported assessment of the stressfulness of the task; the secondary set of measures if based on measurements of heart rate and cortisol levels.

a) Self-reported measure:

Participants were asked in the final questionnaire to indicate their perceptions of the stress or control task. Specifically, they were asked whether they perceived the task as relaxing, easy, stressful, difficult, enjoyable and tiring. Each perception is rated on a 5-point Likert scale from 1 ("not at all") to 5 ("very much"). Of particular interest is the perceived stressfulness of the tasks, which is the primary measure of short-term stress in our analysis.

b) Physiological measures:

Key players in the physiological stress response aimed at re-establishing homeostasis are the autonomic nervous system (ANS) which triggers a rise in heart rate and blood pressure and the hypothalamic-pituitary-adrenal (HPA) axis which increases its release of the hormone cortisol. However, physiological responses tend to be noisy, and therefore hard to detect. This is likely to be even more relevant here since we focus on mild stressors. We collected two physiological measures of the response to stress: heart rate and salivary cortisol. The heart rate captures the response of the ANS to stress. Cortisol on the other hand captures the response of the HPA axis. We furthermore collected measures of salivary testosterone. Testosterone and cortisol levels are positively correlated (Mehta and Josephs, 2010), but have been suggested to capture different responses to stress. While a more pronounced increase in cortisol indicates a passive

[20] See page 19 of the detailed PAP.

coping style, a stronger increase in testosterone is indication of an active coping style (Salvador and Costa, 2009).

Heart Rate

Participants were asked to wear an armband with an optical heart rate sensor (Polar OH1) during the course of the lab experiment. Heart rate data was recorded in one-second intervals and stored on the internal memory of the sensor.[21] To capture the heart rate responses to the stress or control tasks, we defined a baseline period of 5 minutes, beginning with the start of the experiment, and a task period of 10 minutes, beginning with the start of the stress or control task. Comparison of the means during the baseline and the task period provides a measure of the heart rate response to the tasks.[22]

Salivary Cortisol and Testosterone

Participants were asked to provide three saliva samples during the course of the experimental session. The baseline sample was collected at the beginning of the experiment, 9 minutes prior to the start of the stress or control task. The second sample was collected 29 minutes after the start of the stress/control task and the final sample was collected 85 minutes after the start of the stress/control task. Cortisol reactivity to a stressor is found to peak between 10 and 40 minutes following the start of the stress protocol (Newman et al., 2007). Cortisol levels should revert to regular levels by the time the final sample is collected.

The samples were collected using synthetic swabs (Sarstedt Salivette® Cortisol), which were chewed by participants for 60 seconds and then placed in storage tubes. Samples were frozen immediately after collection and shipped under dry ice to Daacro Saliva Lab in Trier (Germany) for analysis. Samples were analysed in duplicate for salivary cortisol and testosterone concentrations.[23]

Comparing salivary cortisol concentrations of the baseline and the second saliva samples provides a measure of the cortisol response to the tasks. Cortisol responses can be problematic

[21] Due to technical problems with the sensors, heart rate data is not available for 29 participants.
[22] We report results for the absolute changes in heart rate between baseline and the task period. The reported results are robust to using the relative changes in heart rate instead.
[23] Absolute changes in salivary cortisol and testosterone levels are used in our analysis. The reported results are robust to using relative changes instead.

to induce and measure (Dickerson and Kemeny, 2004),[24] so we do not necessarily expect to find significant differences in cortisol responses.

3.3 Performance in the stress task

Participants in the stress group were asked to complete a 10-minute block of incentivised tasks. Each participant could lose between £0 and £15 for their randomly assigned group. Rescaling this measure to run from 0% (no correct answers given) to 100% (all tasks and pop-ups solved correctly) allows us to capture participants' performance in the stress task.

4 Hypotheses

The primary hypothesis we test is the following:

> *Hypothesis 1: Short-term stress leads to increased selection of foods high in calories, sugar and saturated fats, both in the context of immediate consumption ("snack choice") and planned consumption ("food shopping choice").*

We consider two channels for such stress-induced changes in dietary choices: (1) impaired decision making and (2) food cravings. The first channel is a cognitive mechanism. Coping with stress requires mental energy and time. Stress induces a reallocation of cognitive resources towards the stressor. As a consequence, less resources are available for processing other decisions (Allen and Armstrong, 2006), leading to more habitual and impulsive food choices.

The second channel is a biological one. The stress-induced release of cortisol in the hypothalamic-pituitary-adrenal (HPA) axis has been argued to cause cravings for energy-dense "comfort foods" (e.g. Adam and Epel, 2007), i.e. a temporary and endogenous change of individuals' food preferences. Exogenous administration of cortisol and other glucocorticoids have been shown to affect food intake in rats (Zakrzewska et al., 1999; Dallman et al., 2004) and humans (Tataranni et al., 1996; George et al., 2010).

Comparison between the effects on immediate and planned consumption can provide insights into the importance of these two potential mechanisms linking short-term stress and

[24] The paper presents a meta-analysis of 208 lab studies on stress and cortisol. They find large variation in measured effect sizes depending on time of day, measurement timing and type of task. They conclude: "Clearly, these findings refute the notion that all psychological stressors elicit cortisol responses. They also call into question the presence of a nonspecific physiological response to all stressors that include HPA activation"

dietary choices. Since stress-induced food cravings (the second channel) are expected to affect immediate consumption more than planned consumption, a larger effect of short-term stress on immediate consumption would point to a stronger relevance of temporary food cravings as a mechanism. If on the other hand a larger effect on planned consumption was observed, this would point to impaired decision making due to cognitive overload (the first channel) as an important mechanism.

In addition to the primary hypothesis described above, we furthermore test the following secondary hypothesis:

> *Hypothesis 2 - The impact of short-term stress on planned consumption ("food shopping choice") will be stronger among participants assigned to the complex choice environment.*

Hypothesis 2 derives from the potential mechanism of impaired decision making due to cognitive overload (first channel). Decisions in a complex choice environment require additional cognitive resources. We therefore expect cognitive depletion to harm decision making more in a complex than in a simple choice environment.

5 Empirical Analysis

5.1 Empirical Strategy

We estimate bivariate models of the following form:

$$Y_i = \beta_0 + \beta_1 T_i + \varepsilon_i$$

where Y_i denotes an outcome measure of the dietary choices and T_i is a dummy variable for the randomly assigned experimental treatment, taking value 1 for participants in a stress session and value 0 for participants in a control session. β_1 is the coefficient of interest and ε_i is an idiosyncratic error term.

To analyse the relevance of choice complexity, we further estimate the following models for the outcomes relating to the food shopping choice:

$$Y_i = \gamma_0 + \gamma_1 T_i + \gamma_2 C_i + \gamma_3 T_i C_i + \varepsilon_i$$

where C_i is a dummy variable for the randomly assigned choice environment, taking value 1 for participants in a complex choice session and value 0 for participants in a simple choice session. γ_1 captures the impact of short-term stress on the outcome variable, γ_2 the impact of choice complexity and γ_3 the additional impact of choice complexity when stressed.

To capture any potentially confounding factors we furthermore augment the above models by including a vector of control variables X_i:

$$Y_i = \beta_0 + \beta_1 T_i + X_i' \delta + \varepsilon_i$$

$$Y_i = \gamma_0 + \gamma_1 T_i + \gamma_2 C_i + \gamma_3 T_i C_i + X_i' \delta + \varepsilon_i$$

The set of control variables X_i includes dummy variables for the time of the experimental session, for the consumption of any food in the last hour, any drink in the last hour, any cocoa product in the last 6 hours and any big meal in the last 6 hours.[25] These control variables were chosen as they differed significantly between the stress and the control group.[26]

We estimate all models using ordinary least squares (OLS).[27] To account for potential error correlation among individuals in the same experimental session, we estimate standard errors robust to clustering at the session level. Due to the relatively small number of clusters, the wild cluster bootstrap approach proposed by Cameron et al. (2008) is also used. As a robustness check we furthermore estimate p-values using the randomization inference procedure set out in Young (2019) which accounts for the randomness in the assignment of the experimental conditions by simulating a large number of hypothetical treatment-control assignments. The results from this approach are similar to those presented below (see Appendix G).

5.2 Pre-test of the Experimental Design

The experimental design was pre-tested in June 2018 using a sample of 50 low-SES mothers in Florence, Italy. Results from this pre-test using an eligible sample of 41 participants showed the novel stress protocol to be effective.[28] The stress task was perceived as significantly more stressful than the control task. The mean heart rate of participants in the stress group increased significantly by 7.0 bpm (8.5%) between baseline and the stress task, a difference-in-difference comparison relative to the control group showed a statistically significant increase by 10.0 bpm. A difference-in-difference comparison showed the stress protocol to induce a statistically significant and sizable increase in salivary cortisol levels. Comparison of salivary cortisol

[25] We control for these variables since they may correlate with salivary cortisol levels and with food consumption choices.

[26] For more details on these balance checks, see Table 3.

[27] The reported results are robust to estimation with a session random effects estimator.

[28] Results of the pre-test can be found in Appendix E.

levels before and after completion of the stress/control task showed an increase by 1.1 nmol/L (24.0%) in the stress group and a decrease by 1.1 nmol/L (22.2%) in the control group.

No significant impacts of short-term stress on food choices were observed in the pre-test. The consumed quantities in the snack choice and the nutritional content of baskets chosen in the food shopping choice were not found to differ significantly between stress and control group.[29]

5.3 Descriptive Statistics

Demographic characteristics of our sample can be found in Table 4. The only statistically significant difference between the stress and the control group is the age of the youngest child.[30] The average age of mothers is approximately 36 years and on average they have two children. 28% of mothers raise their youngest child by themselves. 61% of mothers are married or in a cohabiting relationship. 43% of participants completed GCSEs as their highest qualification, for 38% A levels are the highest qualification. 10% of mothers work full-time, 59% are in part-time employment and 24% are not employed.

Descriptive statistics of the dietary measures used in our analysis are in Table 5. On average, participants ate 41g (171 kcal) of the blueberry mini-muffins and 74g (41 kcal) of the apple slices offered during the snack choice. The average shopping basket selected during the food-shopping choice contained approximately 17,000 kcal, 460g of fat, 200g of saturated fat and 750g of sugar. We observe a small positive, but statistically insignificant, correlation between participants' immediate food intake and the nutrient content of their selected food-shopping basket. This underlines the importance of separately examining both types of food choices.

5.4 Effectiveness of Stress Protocol

The main goal of our stress protocol is to replicate stressors that low SES individuals may encounter in their daily lives. We document here the response of participants to the protocol, by examining their perceptions of the stress and control tasks as well as the response of heart rate and salivary cortisol to the tasks.

[29] In a bivariate regression, saturated fat content of the chosen food shopping baskets in the stress group was found to be lower at a 10% level of statistical significance. When controlling for factors unbalanced between treatment and control group, this difference was no longer found to be statistically significant.

[30] All results reported below are robust to the inclusion of child age as a control variable.

Perceived stressfulness of the tasks

Panel A of Table 6 shows participants' mean perceptions of the stress and control tasks. Perceived stressfulness of the task is significantly higher for the stress task. With a mean perceived stressfulness of 2.7 on the 5-point Likert scale, the stress task was perceived as mildly stressful. This is a considerable difference to the mean perceived stressfulness of 1.5 for the control task. The stress task was furthermore perceived as significantly less relaxing, less easy, more difficult, less enjoyable and more tiring (Table D.2 in Appendix D).

Heart rate response

Figure 1 shows the mean heart rate of participants in the stress and control group for minute intervals during the baseline (the first 5 minutes of the experiment), the pre-task phase, the task and the post-task phase. There are no significant differences in the heart rate levels during baseline between the stress and control group. The pre-task period shows slightly higher heart rate levels in the stress than in the control group, this is likely due to anticipation effects as participants were instructed about the tasks during this phase. Significant differences in heart rate appear immediately after the start of the stress and control task. During the first minute of the task, the mean heart rate of participants completing the stress task is 3.8 bpm (4.8%) above the heart rate of those completing the control task. The second minute of the task shows an even larger difference of 6.1 bpm (7.8%). In the remainder of the task stage the difference in mean heart rate reduces somewhat, likely due to participants adapting to the stress protocol.[31] The gap in heart rate between the two groups closes within minutes of completing the task, no significant differences are found during the post-task period.

Difference-in-difference comparison of the mean heart rate across the two groups and between the baseline and task stages are shown in Panel B of Table 6. For the control group, the mean heart rate is reduced by 3.7 bpm (4.6%) between baseline and the task stage likely due to an elevated heart-rate from physiological activity at arrival (e.g. the walk from the car park to the laboratory). For the stress group, we observe no significant change in mean heart rate from the baseline to the task stage – the downward trend observed in the control group is cancelled out in the stress group by the increase in heart rate caused by the short-term stressor. The difference-in-difference comparison shows a significantly increased heart rate during the stress task by 3.8 bpm relative to the control group. These differences in heart rate between

[31] This gradual closing of the heart rate gap between stress and control group indicates that the observed differences are due to stress, rather than differences in physiological demands of the stress task, which would be expected to persist throughout the task period.

stress and control task are sizeable considering the physiological requirements of both tasks were the same.

There are no straightforward benchmarks to compare this increase in heart rate to in low SES individuals. To put the effect in perspective, it is useful to compare it to normal heart rate variation observed in healthy individuals.[32] A recent large-scale study (Avram et al., 2019) monitoring heart rate in normal daily life finds a difference of 6.5 bpm between the daily peak in average hourly heart rate at 5pm and the lowest value at 5am. The difference between a weekday and a weekend is 0.7 bpm on average. A study tracking the heart rate and blood pressure of male – yet white-collar – workers over workdays and non-workdays (Vrijkotte et al, 2000) found a difference of 6.8 bpm in the workday heart rate between those reporting high and those reporting low work stress. Thus, the increase in heart rate induced in our experiment is 5 times as large as the one experienced when comparing a week day and a weekend day, but is half as large as the one experienced when comparing day and night or when comparing high and low work stress (white collar) individuals. Another benchmark we can use are the studies that use cognitive stressors such as mathematical tasks or tasks similar to the Trier Social Stress Test. For example, Delaney et al. (2014) use a cognitive stressor (IQ tests) and a physiological stressor (the cold pressor test where participants are asked to place their feet into ice-cold water). They find no impact of the cognitive stressor, but a 5% and 8% increase in diastolic and systolic blood pressure respectively for the physiological stressor. Thus, our stress protocol appears to induce a heart rate response that is stronger than the one induced by cognitive tasks in relatively educated samples, but at the lower end of the range induced by "normal stress" that would be experienced in daily life.

Cortisol response

Over the course of the experiment, we collected three saliva samples (at baseline, 29 minutes after the start of the task and at the end) from each participant to track the physiological response of the hypothalamic-pituitary-adrenal (HPA) axis to the stress and control tasks. Figure 2 shows the mean salivary cortisol concentrations of participants in the stress and control groups across these three measurements. For both groups, we observe a downward trend in salivary cortisol over the course of the experiment. With the exception of a marginally significant difference in baseline cortisol, we do not observe any differences in salivary cortisol

[32] Unfortunately, it was not possible to collect data on heart rate variability (HRV) with the devices used in our study since they only provide one data point per second and do not record the timing of individual heart beats.

between the stress and control group. Difference-in-difference comparison of mean cortisol concentrations across the two groups are shown in Panel C of Table 6. A comparison of the cortisol change from the baseline to the second measurement shows a marginally significant difference, with cortisol concentrations decreasing less in the stress group. As this result is entirely driven by differences in the baseline cortisol levels, it cannot be used as indication of a cortisol response to the short-term stressor.[33] Appendix Table D.3 and Figure D.1 report the analysis using relative cortisol changes. This analysis shows significant differences in relative salivary cortisol after the task between the two groups, which are likely driven by differences in the absolute cortisol levels at baseline.

It should be noted that this lack of finding of a cortisol difference is in contrast to the pre-test study conducted in Italy. While we cannot present any decisive evidence it is known that cortisol responses can be quite sensitive to many factors including weather (Kanikowska et al., 2019) and culture (Miller and Kirschbaum, 2019), both of which may explain the greater cortisol responses observed in Italy compared to the UK.[34]

Our findings show that the stress protocol was, as expected, perceived as mildly stressful. An increased heart rate during the stress task indicates a physiological response of the autonomic nervous system to this stressor. We do not observe a response of the hypothalamic-pituitary-adrenal (HPA) axis to our stress protocol. While everyday stressors have been linked with increased cortisol levels (see Saxbe, 2008, for a review), the magnitude of the cortisol response has been found to be lower for mild stressors and for women (Smyth et al., 1998).

Based on these results, we conclude that the stress protocol induced a stress response in individuals and is within the range of the stress response observed in daily life among healthy individuals.

5.5 Impact of Short-term Stress on Food Choices (Hypothesis 1)

In the following, we examine the impact of short-term stress by comparing the food choices made during the experiment by participants assigned to the stress and control group.

---

[33] The saliva samples were furthermore analysed for testosterone, an indicator of an active coping style (Salvador and Costa, 2009). The results of this analysis can be found in Appendix Table D.4 and Appendix Figure D.2. Similar to salivary cortisol, we observe a downward trend in testosterone over the course of the experiment for both groups, but no significant differences between the stress and the control group.

[34] Indeed, room temperatures were 23.6-26.9 °C (mean: 24.6 °C) in Florence and 21.8-24.0 °C (mean: 22.9 °C) in Colchester and the highest outdoor temperatures on the session days were 26-32 °C in Florence and 13-21 °C in Colchester. In line with the results shown by Kanikowska et al. (2019) this would imply higher cortisol levels in Italy relative to the UK.

Immediate food choices

Table 7 shows OLS results for the impact of short-term stress on immediate food consumption as captured by the snack choice during the experiment. Columns 1 and 3 correspond to bivariate models. Columns 2 and 4 correspond to augmented models, which control for the time of the experimental session and for the consumption of foods and drinks prior to the experiment. As shown in column 1, participants in the control group ate 38.2g of the high-calorie mini-muffins. Participants in the stress group consumed an additional 5.4g of the muffins. While this difference is not negligible in size, it is not statistically significant. When controlling for potentially confounding factors that differed between stress and control groups in column 2, we observe a similar difference in muffin intake of 5.1g, which again is not statistically significant.

As reported in column 3, participants in the control condition ate 72.8g of the low-calorie apple slices while those in the stress condition ate an additional 3.1g. This difference in apple intake increases to 5.7g when controlling for session time and for the prior consumption of foods and drinks in column 4. In both specifications, the difference in apple consumption between stress and control group is not precisely estimated. While we observe short-term stress to increase intake of both high- and low-calorie snacks, these increases are not statistically significant. When examining the total energy, saturated fat and sugar intake from both snack types (see Appendix Table D.5), we again find no statistically significant differences between stress and control group. Finally, we define our outcome variable as the share of snack intake from muffins (see Appendix Table D.20). This analysis still shows no statistically significant differences between stress and control group, but they are closer to significance relative to the absolute muffin intake estimations.

Based on 95% confidence intervals from our bootstrapped estimation results of the augmented models, we can reject positive effect sizes above 51% (0.57 SDs) for the muffin intake and above 30% (0.50 SDs) for the apple intake. We are therefore able to significantly reject effects of a size similar to those reported in Epel et al. (2001) and Habhab et al. (2009).

Planned food choices

We now turn to examining the impact of short-term stress on planned future food consumption, by analysing the nutrient content of the baskets selected during the food shopping task. OLS results for the impact of the stress protocol on total energy, saturated fat and sugar content of the selected grocery items can be found in Table 8. Participants in the control group selected baskets with a mean energy content of 17,138 kcal, 202g of saturated fat and 775g of sugar. Results for the bivariate models in column 1 show participants in the stress group on average selected baskets containing 327 kcal less energy, 5g less saturated fat and 48g less sugar. When controlling for the timing of the experimental sessions and the intake of food and drink prior to arrival (including having a big meal and consuming cocoa), the differences in energy and saturated fat content are substantially reduced, to 103 kcal and 2g respectively. The difference in sugar content, on the other hand, slightly increases in absolute magnitude to 53g. The differences in nutrient content of baskets selected by the stress and control groups are not statistically significant in any of the specifications.[35] The selection of baskets with lower energy, saturated fat and sugar content in the stress group is contrary to our expectation under hypothesis 1. While these differences between stress and control group are not precisely estimated, the negative point estimates provide evidence against large positive effects. Indeed, based on 95% confidence intervals from our bootstrapped estimation results of the augmented models, we can reject positive effect sizes above 6% (0.26 SDs) for the calorie content, above 12% (0.28 SDs) for the saturated fat content, and above 7% (0.20 SDs) for the sugar content of the baskets chosen in the shopping choice. In further estimations (see Appendix Table D.6), we find no significant difference in the weight of fruit and vegetables purchased by the stress and control groups. These findings do not support the hypothesis that short-term mild stress leads to less healthy food choices in the context of planned consumption.

Summarizing, these results suggest no significant relationship between mild stress and planned food consumption. We find sizeable but not significant effects on immediate food consumption, possibly due to a larger sample size being required to detect effects of this size.

5.6. Role of choice complexity (Hypothesis 2)

We now examine whether the complexity of the food shopping choice affects the healthiness of the chosen grocery items. Columns 3 and 4 of Table 8 report estimation results for the stand-alone impact of choice complexity, columns 5 and 6 for the interaction between the complexity of the choice environment and the stress condition. Panel A shows a lower energy content of baskets selected in the complex choice environment, both among the stress and the control group. As shown in Panel B, we observe the saturated fat content of baskets selected in the

[35] A test of joint significance for all three primary outcomes also shows no statistically significant differences between the stress and control group.

complex choice environment to be lower in the control, but higher in the stress group. Panel C shows the sugar content of baskets chosen in the complex choice environment to be higher in the control, but lower in the stress group. The impact of choice complexity on the above outcomes is not statistically significant, both in stand-alone estimations and when interacted with the stress condition.

Thus, the evidence does not suggest stress to have differential effects on dietary choices depending on the complexity of the choices. We hence find no indication of a cognitive depletion channel between short-term stress and planned dietary choices.

5.7 Heterogenous effects

Daily stressors may affect different subpopulations and their dietary choices in different ways. We therefore examine whether certain coping styles or a strong physiological or psychological response to the experimental stressor alter the susceptibility of participants' dietary choices to the short-term stress protocol.[36] The self-assessed use of avoidance-based, emotion-oriented and task-oriented coping styles are not found to significantly alter the dietary choices made by the stress group (see Appendix Tables D.7 and D.8). Despite some significant coefficient estimates, we find no strong evidence that perceiving the task as stressful makes participants' food choices more susceptible to the stressor (see Appendix Tables D.9 and D.10). The physiological responses to the stressor as captured by the heart rate as well as the salivary cortisol and testosterone responses do not predict stronger susceptibility of participants' choices to the stress protocol (see Appendix Tables D.11 – D.14).[37]

In Appendix F, we furthermore find no evidence of emotional eating (Van Strien et al., 1986), time preferences and risk attitudes playing a role in the dietary response to stress.

[36] An alternative approach to estimating linear models with interaction terms of the experimental condition with the stress response is an instrumental variables (IV) estimation in an intention-to-treat framework. Instrumenting the stress response with the experimental assignment to stress or control group captures the effect on compliers. However, in this case of a single endogenous regressor and a single instrument, IV only rescales the reduced form coefficients and significance levels are asymptotically identical under the null hypothesis (Angrist, 2006). This is confirmed in IV estimations, which show similar significance levels as reported in Tables 7 and 8.

[37] In an additional analysis reported in Appendix Tables D.15 and D.16, we split the stress group in cortisol responders and cortisol non-responders using the 75th percentile of the cortisol response in the stress group as a threshold. No statistically significant differences between cortisol responders and non-responders as well as the control group are found, except for a lower sugar content of the baskets purchased in the food shopping choice.

5.8 The role of failure (exploratory evidence)

Finally, we present exploratory evidence on the impact of performance in the stress task on dietary choices.[38] We emphasise that the performance in the task is of course not exogenous and hence these results cannot be interpreted as causal. The performance in the task ranges from 0% to 100%, where 100% corresponds to solving all tasks and questions correctly, without any time penalty.[39] The performance in the task is positively related to participants' education and household income. Consistent with cortisol capturing a passive and testosterone capturing an active coping style (Salvador and Costa, 2009), performance is negatively related to the cortisol and positively related to the testosterone response to the stress task. We, however, observe no significant relation to perceived stressfulness of the task or the heart rate response. The results reported below are robust to controlling for these correlates or predictors of task performance.

Panel A of Table 9 shows OLS results for the impact of short-term stress on immediate food consumption, controlling for the performance in the stress task. We find that a very poor performance (0%) in the stress task[40] increases muffin intake by 30.8g (80.6% of the control intake) compared to the control group, which corresponds to an additional intake of 127.8 kcal, as shown in column 1. This difference is not only big in magnitude, but also statistically significant. An improved task performance by 1 percentage point significantly reduces the intake of muffins by 0.5 grams. Controlling for potentially confounding factors in column 2 leads to similar results. There is no significant effect on the intake of apple slices.

Panel B of Table 9 reports the impact of short-term stress on planned future food consumption, controlling for stress task performance. We estimate that a very poor performance (0%) in the stress task decreases the energy content of the selected basket by 2152 kcal (12.6%), the saturated fat content by 35g (17.3%) and the sugar content by 68g (8.8%), however these differences are not statistically significant. Performing well in the task increases energy, saturated fat and sugar content of the selected basket. When controlling for the timing of the experimental sessions and the intake of food and drink prior to arrival, the magnitude of these estimates is similar.

---

[38] Note that this was not part of the original pre-analysis plan and should therefore be considered as explorative evidence.

[39] Participants were not given feedback on their performance in the stress task or the resulting monetary compensation during the experiment, but likely formed expectations based on their perceptions of the task.

[40] A performance of 0% corresponds to no correct answers given in any of the 15 tasks or the 10 pop-ups. The lowest performance in our sample was 13%.

While performance in the task is of course not exogenous and hence these results cannot be interpreted as causal, we find a substantial and statistically significant correlation of a worse performance in the stress task and an increased consumption of high-calorie snacks. Participants in the lowest performance quartile of the stress group consume 47.2% more of the high-calorie snacks than participants in the control group, and 50.4% more than the top quartile of the stress group. This hints at the role of failure in mediating an impact of short-term stress on the intake of high-calorie foods. This effect seems in line with the literature, where solvable versus unsolvable tasks are often used to study stress-induced changes in dietary choices, possibly hinging on failure and stress combined more than stress per se to find an impact on food consumption (see Appendix Table D.1 for a review of the stressors used).

6. Discussion

6.1. External Validity

We have chosen stressors that aim to closely mimic stressors experienced by our target population in their everyday lives, and the design of the food shopping task was chosen for its familiarity. There are however potential limitations to the external validity of the study, which we discuss here briefly.

First, there is a possible concern surrounding the Hawthorne effect. The study was advertised to participants as a study on common household decisions, but participants may have realised that the study relates to food choices and made healthier choices than they normally would. While this is a concern, we study choices that are incentivised to reduce this possibility. Our experiment furthermore involves a control group which would be subject to the same Hawthorne effect.

Second, our sample is not a random sample of the population of interest. As is the case with any experiment, participants self-selected into our study sample and certain subsets of the population might be more likely to self-select into our sample. We compare the demographic characteristics of our sample to those of individuals matching the same demographic eligibility criteria in the nationally representative Understanding Society survey (Institute for Social and Economic Research, 2018).[41] We find little differences in the age of the mothers, the age of their youngest child, the number of children and the mothers' highest qualification. We observe

[41] Results of the comparison are shown in Table D.17 of Appendix D.

moderate differences in marital status, household income, received benefits and employment status. Specifically, participants in our sample are less likely to be single, have a somewhat lower household income, receive lower monthly benefit payments and are more likely to work part-time. Overall, we do not find large differences in the demographic characteristics of our experimental sample and a nationally representative sample from our population of interest.

Third, our planned shopping task only had 156 grocery items, much less than a real-life supermarket. Zizzo et al. (2016) found that participant's choices in a similar incentivised experimental food shopping task were negatively related to the current product stock at home, indicating that participants were making realistic choices aimed at refilling the grocery stock at home. To further assess the validity of our food shopping task, we compare the spending shares of participants on the 10 grocery categories in our experiment with grocery expenditure shares of a nationally representative sample in the food-diary based UK Living Costs and Food Survey 2016/17 (Department for Environment, Food and Rural Affairs, 2018), showing similar expenditure patterns in both.[42] While we observe somewhat more purchases of fruit, vegetables, pasta and rice and less purchases of snacks and ready meals in our experiment, these differences are not large and likely driven by our experimental task covering a larger product range in the former than in the latter categories (compared to a real-life supermarket). In total, the 10 grocery categories used in our task correspond to 79% of non-alcohol grocery spending for consumption at home. In addition, we find that the mean (median) number of categories that participants spent their budget in was 7.92 (8.00) and 94% of participants purchased items from at least 6 categories suggesting that participants were not just selecting items from one or two categories that would be unrepresentative of a standard shopping experience. To further investigate the validity of our food shopping task, we split the sample depending on the household size (up to 3 members versus more than three members). The effect of the treatment is not statistically significant for either of these groups, but we do find a stronger negative stress effect in small households and positive or zero stress effects in large households (Appendix D, Table D.19). If the budget of £30 in the shopping task was more realistic for small households, these split sample estimation results support our finding of no large positive effects of stress on shopping nutrient content.

Fourth, while the primary aim of these tasks was to induce mild stress, it is plausible they have a *priming* effect, i.e. that they remind mothers of their household responsibilities or their limited budget. Of course such a mechanism would plausibly also be at work in a real life

[42] Results of the comparison are shown in Table D.18 of Appendix D.

setting and is not unique to our experimental setting. Also, it is unclear in what direction such priming effects would work. The tasks could raise awareness about household responsibilities (which could lead to healthier choices), but also about existing financial hurdles (which could lead to unhealthier choices). The question of how priming alone could affect behaviour (without stress) is interesting and may warrant further research.

6.2. Power

Finally, we come back to the challenge of statistical power. As mentioned earlier, the sample size we have used is much larger than experimental studies examining the effects of stress, but prior studies did not aim at studying the impact of "realistic" every-day stressors. Rather the protocols were chosen to maximize the chances of inducing a stress response, and then study how that affected eating decisions. Here our goal was to study the impact of more realistic mild stressors.

Ex post, we can use our data to calculate the minimum sample sizes required to detect small, but potentially quantitatively meaningful, effects with 80% power, assuming the same clustering structure with an average cluster size of 13.[43] Our calculations show that sample sizes of 3,978 / 1,118 are needed to detect effects on muffin intake / apple intake of 10% of the control group mean with 80% power. For the food shopping choice, sample sizes of 195 / 663 / 403 are required to have 80% power in detecting 10% effects on energy / saturated fat / sugar content. The required sample sizes to detect effect sizes of 5% are substantially larger. These power calculations show that very large (perhaps unrealistically large) sample sizes are required for any experimental studies of the impact of mild everyday stressors.

7 Conclusion

In this study, we examine the impact of mild short-term stress on food choices, both in the context of immediate and planned consumption, by conducting a lab experiment with 196 low-SES mothers. We introduce a novel incentivised stress protocol developed to mimic everyday stressors in low-SES families. At the start of the experiment, participants in the stress group were asked to complete this stress task, while participants in the control group were asked to complete a control task. After, participants were asked to purchase food items in a "virtual

[43] Details of our ex-post power calculations can be found in Appendix C

supermarket" as part of an incentivised food shopping choice and were offered high- and low-calorie snacks for immediate consumption. We use the nutritional content of the chosen food-shopping basket and the quantity of snacks eaten to determine the impact of short-term stress on planned and immediate food consumption choices. We asked participants about their perceptions of the stress or control task and measured their salivary cortisol as well as their heart rate over the course of the experimental sessions to assess the stressfulness of the stress task.

The novel stress protocol was perceived by participants as significantly more stressful than the control task. This is supported by a significant increase in the heart rate of participants in the stress group when compared to the control group. The task appeared to have been perceived as mildly stressful, in alignment with what we were expecting. We do not observe a significant difference in the cortisol levels of the stress and the control groups. This is not unexpected, as it is known that cortisol responses may not occur following mild stressors and, in general, can be problematic to measure. The task perceptions and heart rate data are reliable evidence that mild stress was induced among the stress group, but not among the control group. We do not find evidence of a significant impact of short-term stress on immediate or planned food choices. For the planned food choice, the negative point estimates provide evidence against large positive effects and do not support an effect of maternal day-to-day stress on household food consumption. For the immediate choice, we find a non-negligible increase in consumption of the high calorie snacks. Previous findings in the literature report a positive impact of short-term stress, induced using artificial stressors such as unsolvable mathematical tasks or the Trier Social Stress Test, on food consumption. Contrary to these studies, our stress protocol is more realistic as it mimics common everyday stressors and we focus on the population of low-SES mothers. On balance, our results do not significantly support the hypothesis that everyday mild stressors lead to unhealthier eating choices among low-SES mothers.

The complexity of the choice environment, participants' coping styles as well as the psychological and physiological response to the experimental stressor are not found to affect the susceptibility of dietary choices to short-term stress. If we control for performance in the task though, we find that poor performance leads to a higher intake of calorie dense foods (muffins), indicating that it may not be exposure to stress per se that matters but a combination of stress and failure. Further research is needed to investigate this possibility.

References

Adam, T.C. and Epel, E.S., 2007. Stress, eating and the reward system. *Physiology and Behavior*, 91(4), pp.449–458.

Adler, N.E., Boyce, T., Chesney, M.A., Cohen, S., Folkman, S., Kahn, R.L., and Syme, S.L., 1994. Socioeconomic Status and Health: The Challenge of the Gradient. *American Psychologist*, 49(1), pp.15–24.

Allen, T.D. and Armstrong, J., 2006. Further Examination of the Link Between Work-Family Conflict and Physical Health. *American Behavioral Scientist*, 49(9), pp.1204–1221.

Almeida, D.M., Neupert, S.D., Banks, S.R., and Serido, J., 2005. Do daily stress processes account for socioeconomic health disparities? *The Journals of Gerontology: Series B*, 60(S2), pp.S34-S39.

Angrist, J.D., 2006. Instrumental variables methods in experimental criminological research: What, why and how. *Journal of Experimental Criminology*, 2(1), pp.23–44.

Avram, R., Tison, G.H., Aschbacher, K., Kuhar, P., Vittinghoff, E., Butzner, M., Runge, R., Wu, N., Pletcher, M.J., Marcus, G.M. and Olgin, J., 2019. Real-world heart rate norms in the Health eHeart study. *npj Digital Medicine*, 2(58).

Babcock, P., Bedard, K., Charness, G., Hartman, J., and Royer, H., 2015. Letting down the team? Social effects of team incentives. *Journal of the European Economic Association*, 13(5), pp.841–870.

Barrington, W.E., Beresford, S.A.A., McGregor, B.A. and White, E., 2014. Perceived Stress and Eating Behaviors by Sex, Obesity Status, and Stress Vulnerability: Findings from the Vitamins and Lifestyle (VITAL) Study. *Journal of the Academy of Nutrition and Dietetics*, 114(11), pp.1791–1799.

Belot, M., James, J. and Spiteri, J., 2020a. Facilitating healthy dietary habits: An experiment with a low income population. *European Economic Review*, 129, p.103550.

Belot, M., James, J., Vecchi, M., and Vitt, N., 2020b. Maternal Stress During Pregnancy and Children's Food Preferences. *Unpublished manuscript*.

Buckert, M., Oechssler, J., and Schwieren, C., 2017a. Imitation under stress. *Journal of Economic Behavior & Organization*, 139, pp.252–266.

Buckert, M., Schwieren, C., Kudielka, B.M., and Fiebach, C.J., 2017b. How stressful are economic competitions in the lab? An investigation with physiological measures. *Journal of Economic Psychology*, 62, pp.231–245.

Buser, T., Dreber, A., and Mollerstrom, J., 2017. The impact of stress on tournament entry. *Experimental Economics*, 20(2), pp.506–530.

Cameron, A.C., Gelbach, J.B., and Miller, D.L., 2008. Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics*, 90(3), pp.414–427.

Cawley, J., 2011. The Economics of Obesity. In: J. Cawley, ed. *The Oxford Handbook of the Social Science of Obesity*. Oxford University Press, pp.120–137.

Chamberlain, K. and Zika, S., 1990. The minor events approach to stress: Support for the use of daily hassles. *British Journal of Psychology*, 81(4), pp.469–481.

Dallman, M.F., La Fleur, S.E., Pecoraro, N.C., Gomez, F., Houshyar, H., and Akana, S.F., 2004. Minireview: Glucocorticoids - Food intake, abdominal obesity, and wealthy nations in 2004. *Endocrinology*, 145(6), pp.2633–2638.

Delaney, L., Fink, G., and Harmon, C., 2014. Effects of stress on economic decision-making: Evidence from laboratory experiments. *Journal of Behavioral Economics for Policy, Vol. 3, No. 1, pp.30-36, 2019*.

Department for Environment, Food and Rural Affairs, 2018. *Living Costs and Food Survey, 2016-2017*. London: Familiy food datasets.

Dickerson, S.S. and Kemeny, M.E., 2004. Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research. *Psychological Bulletin*, 130(3), pp.355–391.

Epel, E., Lapidus, R., McEwen, B., and Brownell, K., 2001. Stress may add bite to appetite in women: A laboratory study of stress-induced cortisol and eating behavior. *Psychoneuroendocrinology*, 26(1), pp.37–49.

George, S.A., Khan, S., Briggs, H., and Abelson, J.L., 2010. CRH-stimulated cortisol release and food intake in healthy, non-obese adults. *Psychoneuroendocrinology*, 35(4), pp.607–612.

Goette, L., Bendahan, S., Thoresen, J., Hollis, F., and Sandi, C., 2015. Stress pulls us apart: Anxiety leads to differences in competitive confidence under stress. *Psychoneuroendocrinology*, 54, pp.115–123.

Grzywacz, J.G., Almeida D.M., Neupert S.D., and Ettner S.L., 2004. Socioeconomic Status and Health: A Micro-level Analysis of Exposure and Vulnerability to Daily Stressors. *Journal of Health and Social Behavior*, 45(1), pp.1-16.

Habhab, S., Sheldon, J.P., and Loeb, R.C., 2009. The relationship between stress, dietary restraint, and food preferences in women. *Appetite*, 52, pp.437–444.

Harnack, L., Story, M., Martinson, B., Neumark-Sztainer, D., and Stang, J., 1998. Guess who's cooking? The role of men in meal planning, shopping, and preparation in US families. *Journal of the American Dietetic Association*, 98(9), pp.995–1000.

Institute for Social and Economic Research, 2018. *Understanding Society: Waves 1-8, 2009-2017 and Harmonised BHPS: Waves 1-18, 1991-2009*. University of Essex. Available via UK Data Service, SN: 6614.

Kandasamy, N., Hardy, B., Page, L., Schaffner, M., Graggaber, J., Powlson, A.S., Fletcher, P.C., Gurnell, M., and Coates, J., 2014. Cortisol shifts financial risk preferences. *Proceedings of the National Academy of Sciences*, 111(9), pp.3608–3613.

Kanikowska, D., Roszak, M., Rutkowski, R., Sato, M., Sikorska, D., Orzechowska, Z., Breborowicz, A. and Witowski, J., 2019. Seasonal differences in rhythmicity of salivary cortisol in healthy adults. *Journal of Applied Physiology*, 126(3), pp.764–770.Kirschbaum, C., Pirke, K.-M., and Hellhammer, D.H., 1993. The `Trier Social Stress Test' - A Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting. *Neuropsychobiology*, 28, pp.76–81.

Kocher, M.G., Pahlke, J., and Trautmann, S.T., 2013. Tempus Fugit: Time Pressure in Risky Decisions. *Management Science*, 59(10), pp.2380–2391.

Loewenstein, G., Asch, D.A., Friedman, J.Y., Melichar, L.A., and Volpp, K.G., 2012. Can Behavioural Economics Make us Healthier? *British Medical Journal*, 344, pp.23–25.

Marmot, M.G., Rose, G., Shipley, M., and Hamilton, P.J.S., 1978. Employment grade and coronary heart disease in British civil servants. *Journal of Epidemiology and Community Health*, 32(4), pp.244–249.

Marmot, M.G., Smith, G.D., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E., and Feeney, A., 1991. Health inequalities among British civil servants: the Whitehall II study. *Lancet*, 337, pp.1387–1393.

Mehta, P.H. and Josephs, R.A., 2010. Testosterone and cortisol jointly regulate dominance: Evidence for a dual-hormone hypothesis. *Hormones and Behavior*, 58(5), pp.898–906.

Miller, R. and Kirschbaum, C., 2019. Cultures under stress: a cross-national meta-analysis of cortisol responses to the Trier Social Stress Test and their association with anxiety-related value orientations and internalizing mental disorders. Psychoneuroendocrinology, 105, pp.147-154.

Newman, E., O'Connor, D.B., and Conner, M., 2007. Daily hassles and eating behaviour: The role of cortisol reactivity status. *Psychoneuroendocrinology*, 32(2), pp.125–132.

Office for National Statistics, 2017. *Effects of Taxes and Benefits on Household Income, 2015-2016*. 2nd Edition. Available via UK Data Service, SN: 8201.

Olson, S.L. and Banyard, V., 1993. 'Stop the World So I Can Get off for a While': Sources of Daily Stress in the Lives of Low-Income Single Mothers of Young Children. *Family Relations*, 42(1), pp.50–56.

Pampel, F.C., Krueger, P.M., and Denney, J.T., 2010. Socioeconomic Disparities in Health Behaviors. *Annual Review of Sociology*, 36, pp.349–370.

Richardson, A.S., Arsenault, J.E., Cates, S.C. and Muth, M.K., 2015. Perceived stress, unhealthy eating behaviors, and severe obesity in low-income women. *Nutrition Journal*, 14(122).

Salvador, A. and Costa, R., 2009. Coping with competition: Neuroendocrine responses and cognitive variables. *Neuroscience and Biobehavioral Reviews*, 33(2), pp.160–170.

Sapolsky, R.M., 2004. Social Status and Health in Humans and Other Animals. *Annual Review of Anthropology*, 33, pp.393–418.

Saxbe, D.E., 2008. A field (researcher's) guide to cortisol: tracking HPA axis functioning in everyday life. *Health Psychology Review*, 2(2), pp.163–190.

Sims, R., Gordon, S., Garcia, W., Clark, E., Monye, D., Callender, C. and Campbell, A., 2008. Perceived stress and eating behaviors in a community-based sample of African Americans. *Eating Behaviors*, 9(2), pp.137–142.

Smyth, J., Ockenfels, M.C., Porter, L., Kirschbaum, C., Hellhammer, D.H. and Stone, A.A., 1998. Stressors and mood measured on a momentary basis are associated with salivary cortisol secretion. *Psychoneuroendocrinology*, 23(4), pp.353–370.

Spiteri, J., James, J., and Belot, M., 2019. A computer-based incentivized food basket choice tool: presentation and evaluation. *PLoS ONE*, 14(1), p.e0210061.

Tataranni, P.A., Larson, D.E., Snitker, S., Young, J.B., Flatt, J.P., and Ravussin, E., 1996. Effects of glucocorticoids on energy metabolism and food intake in humans. *American Journal of Physiology - Endocrinology and Metabolism*, 271(2), pp.E317–E325.

Van Strien, T., Frijters, J.E.R., Bergers, G.P.A., and Defares, P.B., 1986. The Dutch Eating Behavior Questionnaire (DEBQ) for Assessment of Restrained, Emotional, and External Eating Behavior. *International Journal of Eating Disorders*, 5(2), pp.295–315.

Vrijkotte, T.G.M., Van Doornen, L.J.P. and De Geus, E.J.C., 2000. Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability. *Hypertension*, 35(4), pp.880–886.

Young, A., 2019. Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results. The Quarterly Journal of Economics, 134(2), pp.557–598.

Zakrzewska, K.E., Cusin, I., Stricker-Krongrad, A., Boss, O., Ricquier, D., Jeanrenaud, B., and Rohner-Jeanrenaud, F., 1999. Induction of obesity and hyperleptinemia by central glucocorticoid infusion in the rat. *Diabetes*, 48(2), pp.365–370.

Zhong, S., Shalev, I., Koh, D., Ebstein, R.P., and Chew, S.H., 2018. Competitiveness and Stress. *International Economic Review*, 59(3), pp.1263–1281.

Zizzo, D.J., Parravano, M., Nakamura, R., Forwood, S., and Suhrcke, M., 2016. The impact of taxation and framing on diet: An online field study with breakfast cereals and soft drinks. *Appetite 107:694*

Table 1: Sample size

	Stress		Control		Total
	Simple	Complex	Simple	Complex	
Total:	63	58	60	46	227
Not eligible due to:					
- Child's age	0	1	1	0	2
- HH income	0	0	1	1	2
- University degree	0	0	1	0	1
- Food allergy / intolerance	1	2	1	3	7
- Medical condition (diet-related)	2	0	0	0	2
- Depression	2	6	7	2	17
Eligible:	58	49	49	40	196

Table 2: Timeline of the experiment

1) Arrival at the lab:
   - Heart rate monitors fitted
   - Body measurements
   - Mouth rinsed to prepare for saliva sample

2) Instructions (6-10 min)
   1st Saliva Sample

3) Instructions (9 min)
   Stress / Control Task (10 min)

4) Instructions (6 min)
   Food Shopping Choice (10 min)

5) Instructions (3 min)
   2nd Saliva Sample (29 min after start of stress / control task)

6) Break (5 min)
   - Low and high-calorie snacks available at desks

7) 1st Questionnaire (15 min)
   - Snacks still available at desks

8) Collection of snack bowls (4 min)
   - Mouth rinsed to prepare for 3rd saliva sample

9) 2nd Questionnaire (30 min)

10) Instructions (2 min)
    3rd Saliva Sample (85 min after start of stress / control task)

Table 3: Balance of control variables across groups

	(1) Stress	(2) Control	(1) vs (2) P-Value
Session time:			
10:30	0.43	0.33	0.14
14:00	0.22	0.42	0.00
17:00	0.35	0.26	0.19
Room temperature	22.90	22.89	0.92
	(0.69)	(0.47)	
Diet - mother:			
vegetarian	0.07	0.06	0.79
	(0.25)	(0.23)	
vegan	0.00	0.00	.
	(0.00)	(0.00)	
allergies	0.02	0.02	0.85
	(0.14)	(0.15)	
intolerances	0.03	0.00	0.11
	(0.17)	(0.00)	
other	0.05	0.02	0.36
	(0.21)	(0.15)	
Snack position: apples - right	0.44	0.45	0.88
	(0.50)	(0.50)	
Previous experiment	0.21	0.16	0.39
	(0.41)	(0.37)	
Food - last 1hr	0.18	0.06	0.01
	(0.38)	(0.23)	
Big meal - last 6hrs	0.19	0.09	0.05
	(0.39)	(0.29)	
Cocoa - last 6hrs	0.08	0.03	0.14
	(0.28)	(0.18)	
Drink - last 1hr	0.26	0.13	0.03
	(0.44)	(0.34)	
Alcohol - last 24hrs	0.18	0.15	0.55
	(0.38)	(0.36)	
Caffeine - last 6hrs	0.50	0.51	0.89
	(0.50)	(0.50)	
Medication - last 24hrs	0.36	0.34	0.69
	(0.48)	(0.48)	
Exercise - last 6hrs	0.27	0.21	0.35
	(0.45)	(0.41)	
Smoker	0.30	0.28	0.78
	(0.46)	(0.45)	
Cigarettes per day	8.56	8.12	0.73
	(4.85)	(4.75)	
Any allergies	0.38	0.36	0.73
	(0.49)	(0.48)	
Regular medication	0.27	0.33	0.41
	(0.45)	(0.47)	
Oral contraceptive	0.27	0.22	0.46
	(0.45)	(0.42)	
Menopause	0.01	0.00	0.36
	(0.10)	(0.00)	
Endocrine disorders	0.00	0.00	1.00
	(0.14)	(0.15)	
N	107	89	196

Table 4: Demographic characteristics

	(1) Stress	(2) Control	(1) vs (2) P-Value
Age - mother	35.70	36.24	0.53
	(5.96)	(5.83)	
Age - youngest child	5.94	7.10	0.01
	(2.79)	(3.09)	
No. of children	1.97	2.18	0.15
	(1.01)	(1.01)	
Single parent	0.28	0.28	0.99
	(0.45)	(0.45)	
Marital status:			
single	0.21	0.26	0.48
married	0.44	0.39	0.52
cohabiting	0.17	0.21	0.42
other	0.18	0.13	0.42
Monthly HH net income:			
<1000 GBP	0.13	0.09	0.37
1000-2000 GBP	0.46	0.49	0.61
>2000 GBP	0.41	0.42	0.95
Monthly benefits:			
none	0.17	0.11	0.26
1-650 GBP	0.56	0.65	0.23
>650 GBP	0.27	0.24	0.66
Highest qualification:			
none	0.08	0.10	0.68
GCSE: <5 A*-C passes	0.17	0.10	0.18
GCSE: ≥5 A*-C passes	0.26	0.33	0.33
A levels	0.38	0.37	0.86
professional	0.10	0.10	0.97
Employment status:			
full-time	0.11	0.08	0.43
part-time	0.54	0.64	0.17
self-employed	0.08	0.06	0.45
not employed	0.26	0.22	0.55
N	107	89	196

Table 5: Descriptive statistics

	Mean	SD	Min	Max
Snack choice:				
Muffins (g)	41.15	34.80	0.00	142.00
Apples (g)	74.46	48.82	0.00	180.00
Energy (kcal)	211.64	151.95	0.00	667.19
Fat (g)	9.13	7.67	0.00	31.38
Saturated fat (g)	1.15	0.97	0.00	3.98
Carbohydrates (g)	29.73	19.49	0.00	89.31
Sugar (g)	20.68	12.41	0.00	58.51
Protein (g)	1.99	1.53	0.00	6.51
Salt (g)	0.20	0.15	0.00	0.65
Shopping choice:				
Energy (kcal)	16960.10	4046.76	7094.90	27061.19
Fat (g)	463.14	197.50	113.29	1167.79
Saturated fat (g)	199.43	95.32	31.32	555.60
Carbohydrates (g)	2304.84	827.72	106.24	4619.40
Sugar (g)	748.31	245.26	87.00	1675.11
Protein (g)	771.13	169.59	401.07	1199.40
Salt (g)	34.04	9.63	5.23	61.60
Fruit & veg (g)	2838.55	1820.31	0.00	9164.00
N	196			

Table 6: Response to stress and control task

Panel A: Perceived stress	Stress	Control	Difference	P	P(Wild)
	2.745	1.517	1.228	0.000	0.000
	(0.064)	(0.094)			
N	106	89	195		

Panel B: Heart rate	Stress	Control	Difference	P	P(Wild)
(1) Baseline	80.854	80.807	0.047	0.981	0.978
	(0.638)	(1.840)			
(2) Task	80.917	77.104	3.813	0.033	0.028
	(0.927)	(1.307)			
Difference: (2)-(1)	0.062	-3.703	3.766		
P	0.916	0.000	0.001		
P(Wild)	0.884	0.000	0.001		
N	94	73	167		

Panel C: Salivary cortisol	Stress	Control	Difference	P	P(Wild)
(1) Baseline	4.262	4.892	-0.630	0.230	0.252
	(0.322)	(0.384)			
(2) Post-Task	3.500	3.556	-0.056	0.871	0.869
	(0.240)	(0.241)			
(3) End	3.095	3.025	0.070	0.876	0.889
	(0.282)	(0.336)			
Difference: (2)-(1)	-0.761	-1.335	0.574		
P	0.001	0.000	0.089		
P(Wild)	0.007	0.000	0.086		
N	107	89	196		

Note: Standard errors of the stress and control means were clustered at the session level and are shown in parentheses. For the differences in mean, p-values based on standard errors clustered at the session level and p-values based on a wild bootstrap clustered at the session level are reported. Perceived stress was scored from 1 for 'not at all' to 5 for 'very much'. Heart rate means were calculated based on heart rate data collected every second.

Table 7: Impact of acute stress on snack intake

	Muffins (g)		Apples (g)	
	(1)	(2)	(3)	(4)
Stress	5.396	5.081	3.053	5.736
	(6.691)	(5.782)	(6.152)	(6.691)
	[0.499]	[0.471]	[0.670]	[0.457]
Controls for:				
Session time	No	Yes	No	Yes
Time since food/drink	No	Yes	No	Yes
Control mean	38.202		72.798	
Control SD	34.323		42.863	
N	196	196	196	196

Note: Standard errors clustered at the session level are shown in parentheses. P-values based on a wild bootstrap clustered at the session level are shown in brackets. Significance levels correspond to the largest p-value obtained from both methods and are indicated as follows: * p<0.1, ** p<0.05, *** p<0.01

Table 8: Impact of acute stress on food shopping

Panel A	Energy (kcal)					
	(1)	(2)	(3)	(4)	(5)	(6)
Stress	-326.627	-103.019			-222.230	-20.000
	(451.987)	(445.395)			(698.664)	(606.634)
	[0.534]	[0.817]			[0.828]	[0.970]
Complex			-464.333	-317.335	-340.412	-214.320
			(410.513)	(359.590)	(667.002)	(657.182)
			[0.311]	[0.443]	[0.672]	[0.779]
Stress * Complex					-221.648	-201.619
					(823.634)	(789.510)
					[0.824]	[0.838]
Controls for:						
Session time	No	Yes	No	Yes	No	Yes
Time since food/drink	No	Yes	No	Yes	No	Yes
Control mean			17138.414			
Control SD			3917.050			
N	196	196	196	196	196	196

Panel B	Saturated fat (g)					
	(1)	(2)	(3)	(4)	(5)	(6)
Stress	-5.403	-2.101			-7.775	-4.740
	(9.896)	(11.681)			(11.900)	(13.779)
	[0.621]	[0.860]			[0.584]	[0.787]
Complex	[0.584]	[0.868]			-10.009	-11.634
			-7.121	-8.880	(14.614)	(14.216)
			(10.168)	(10.128)	[0.603]	[0.582]
Stress * Complex			[0.531]	[0.467]	5.366	5.072
					(20.138)	(21.111)
					[0.814]	[0.854]
Controls for:						
Session time	No	Yes	No	Yes	No	Yes
Time since food/drink	No	Yes	No	Yes	No	Yes
Control mean			202.380			
Control SD			90.529			
N	196	196	196	196	196	196

Panel C	Sugar (g)					
	(1)	(2)	(3)	(4)	(5)	(6)
Stress	-48.356	-52.720			-18.233	-22.833
	(42.017)	(44.609)			(57.111)	(58.356)
	[0.287]	[0.280]			[0.793]	[0.765]
Complex			6.000	10.612	42.809	42.745
			(45.085)	(48.717)	(49.249)	(50.884)
			[0.903]	[0.861]	[0.424]	[0.485]
Stress * Complex					-66.574	-63.893
					(82.171)	(95.763)
					[0.468]	[0.636]
Controls for:						
Session time	No	Yes	No	Yes	No	Yes
Time since food/drink	No	Yes	No	Yes	No	Yes
Control mean			774.705			
Control SD			255.610			
N	196	196	196	196	196	196

Note: Standard errors clustered at the session level are shown in parentheses. P-values based on a wild bootstrap clustered at the session level are shown in brackets. Significance levels correspond to the largest p-value obtained from both methods and are indicated as follows: * p<0.1, ** p<0.05, *** p<0.01

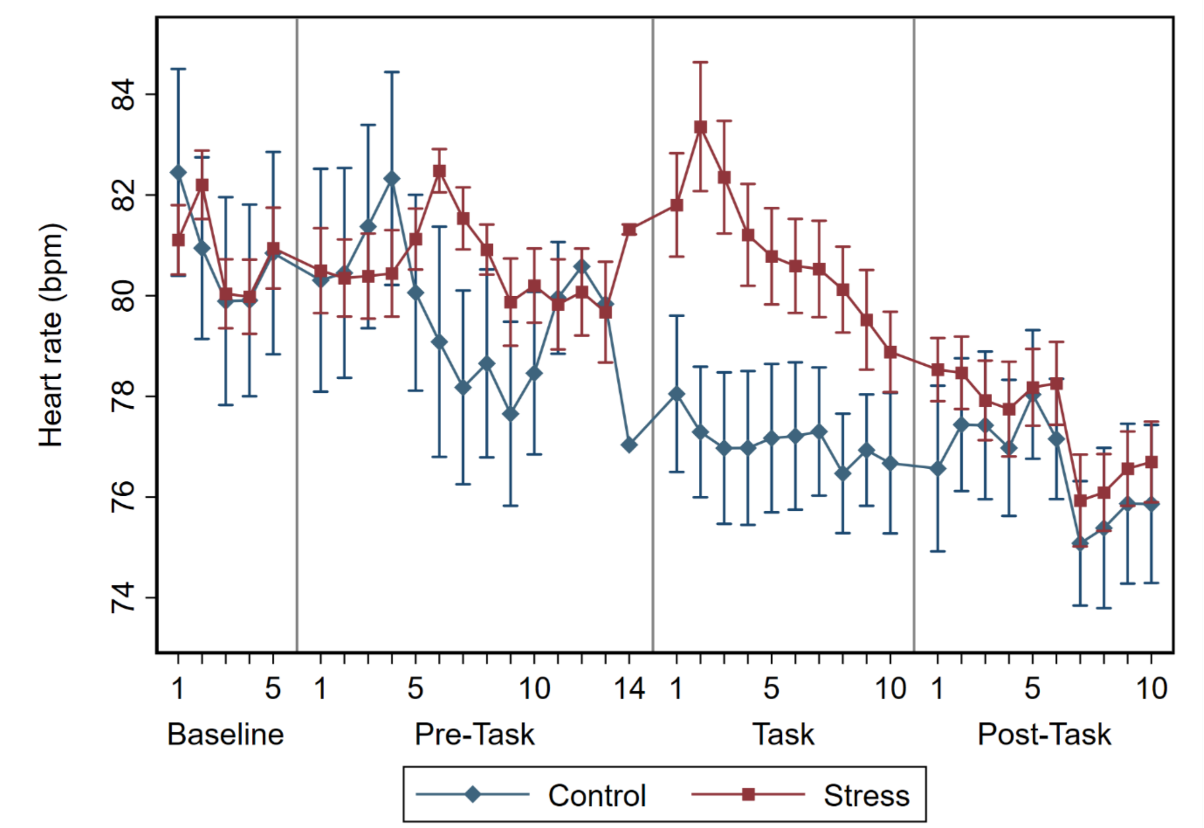
Table 9: The role of participants' stress task performance

Panel A: Snack intake	Muffins (g)		Apples (g)	
	(1)	(2)	(3)	(4)
Stress	30.803**	31.214***	4.903	5.510
	(8.550)	(7.081)	(10.494)	(12.288)
	[0.010]	[0.004]	[0.664]	[0.648]
Stress task performance	-0.502***	-0.516***	-0.037	0.004
	(0.087)	(0.090)	(0.226)	(0.234)
	[0.003]	[0.003]	[0.849]	[0.989]
Controls for:				
Session time	No	Yes	No	Yes
Time since food/drink	No	Yes	No	Yes
Control mean	38.202		72.798	
Control SD	34.323		42.863	
N	196	196	196	196

Panel B: Food shopping	Energy (kcal)		Saturated fat (g)		Sugar (g)	
	(1)	(2)	(3)	(4)	(5)	(6)
Stress	-2152.804	-1813.934	-34.934	-32.259	-68.013	-70.393
	(1381.700)	(1400.023)	(28.304)	(30.955)	(67.729)	(71.876)
	[0.171]	[0.239]	[0.320]	[0.423]	[0.334]	[0.352]
Stress task performance	36.098	33.798	0.584	0.596	0.389	0.349
	(25.494)	(26.609)	(0.478)	(0.520)	(1.173)	(1.201)
	[0.209]	[0.247]	[0.388]	[0.381]	[0.765]	[0.741]
Controls for:						
Session time	No	Yes	No	Yes	No	Yes
Time since food/drink	No	Yes	No	Yes	No	Yes
Control mean	17138.414		202.380		774.705	
Control SD	3917.050		90.529		255.610	
N	196	196	196	196	196	196

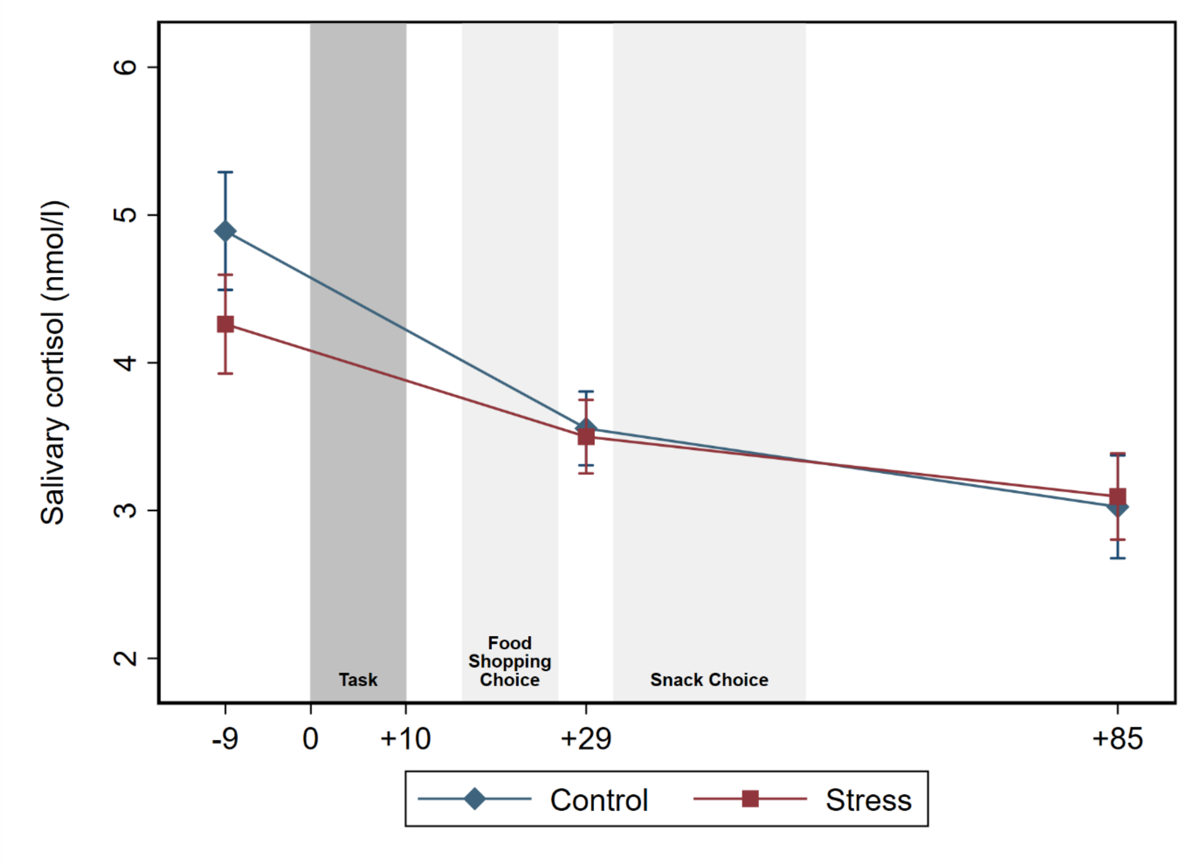
Note: Standard errors clustered at the session level are shown in parentheses. P-values based on a wild bootstrap clustered at the session level are shown in brackets. Significance levels correspond to the largest p-value obtained from both methods and are indicated as follows: * p<0.1, ** p<0.05, *** p<0.01

Figure 1: Heart rate response to stress / control task



Note: Means were calculated for minute intervals based on heart rate data collected every second. Bands indicate +/- 1 standard error. The length of the pre-task period differed across sessions (between 9 min 45 s and 13 min 39 s), but it did not differ significantly between stress and control session.

Figure 2: Salivary cortisol response to stress / control task



Note: Bands indicate +/- 1 standard error.