Satisfaction and the potentially misleading power of counterfactual reasoning: a field study set before, during and after the COVID-19 lockdown

Thomas Martin & Daniel Sgroi

December 2022  No: 1443
Satisfaction and the potentially misleading power of counter-factual reasoning: a field study set before, during and after the COVID-19 lockdown

Thomas Martin$^1$ and Daniel Sgroi$^{1,2}$

$^1$University of Warwick
$^2$IZA Bonn

November 2022

Abstract

Does imagining what life could have been in the absence of a shock change current satisfaction? To answer this we collect field data through a survey that covers the period before, during and after the COVID-19 lockdown, exploiting the features of a natural experiment combined with induced variation stemming from a randomized control trial (RCT). Our data covers first year students studying before the COVID-19 pandemic, during the full COVID-19 lockdown period, and during the partial COVID-19 lockdown period. The RCT directs a subset of students to imagine how satisfied they could have been in the absence of COVID-19. The control group are instead asked about their current satisfaction. We find that imagining life in the absence of a shock (COVID-19) can impact current satisfaction: the higher individuals think their satisfaction would have been in the absence of the shock, the lower their current satisfaction. However, the natural experiment component of our study suggests that counterfactual reasoning may mislead. By comparing the satisfaction of COVID-19 students asked to imagine university life without COVID-19, with the reported satisfaction of equivalent students just before the arrival of COVID-19, we show students typically over-exaggerate how satisfied they would have been if a negative shock had not happened.

*Corresponding author’s email: t.i.martin@warwick.ac.uk
“The harvest is always richer in another man’s field” [Ovid, The Art of Love, 1 BC]¹

1 Introduction

The COVID-19 pandemic severely disrupted university life when, unexpectedly teaching in the 2020/21 academic year shifted online and opportunities for students to socialise were reduced substantially. The disruption impacted students’ mental health, physical activity (Giuntella et al., 2021; Barbieri et al., 2021), study and graduation plans (Aucejo et al., 2020) and satisfaction with e-learning (Keržić et al., 2021). Students who entered university in October 2020 experienced a year that they could not have imagined when they applied to university. We postulate that the COVID-19 shock may have led students to imagine how their university life could have been had COVID-19 not occurred which might in turn have an impact on their current life satisfaction.

The key research focus of this paper can be summarised in the form of two simple questions which we address using subtly different forms of variation: the first exploits a randomized control trial, the second an exogenous natural experiment. Our first core question asks: does imagining what life could have been in absence of a shock impact current satisfaction? Our central hypothesis is that if people imagine how satisfied they could have been, had a shock not occurred, then their level of imagined satisfaction serves as an internal reference point which may affect current judgments about satisfaction. We hypothesise that the more (less) satisfied someone imagines they could have been had the shock not occurred then the less (more) satisfied they are now. To test this hypothesis we use COVID-19 as a real-life unexpected shock - and compare satisfaction responses from a survey experiment that randomly asks respondents about their imagined satisfaction in absence of COVID-19 or their current satisfaction.²

Our second core question asks: are responses to questions about imagined satisfaction, in the absence of a shock, realistic? The timing of our survey covers both pre-COVID-19 and COVID-19 periods, which we use as an opportunity to test how realistic students’ imagined satisfaction in

¹Translated in Liveley (2005).
²The full survey is presented in Appendix B together with a precise description of which questions are seen by each group of students. The survey itself was established prior to the COVID-19 pandemic as an ongoing longitudinal source of data on student wellbeing with the potential to be used for randomized control trials and to span a long enough period to allow us to consider the impact of significant changes on student life.
absence of COVID-19 might be. To do this we compare imagined satisfaction in absence of COVID-19 (for students who experienced university during COVID-19) with the actual satisfaction of students prior to COVID-19. If the imagined satisfaction in absence of COVID-19 is greater (lesser) than the actual satisfaction of the pre-COVID-19 cohort then students’ imagined satisfaction in the absence of COVID-19 may be over (under) exaggerated compared to a realistic benchmark. Ex-ante it is unclear whether students will be realistic in their responses, or whether they will over or under-estimate imagined satisfaction. We argue that comparing satisfaction between COVID-19 and pre-COVID-19 cohorts is a valid comparison: if COVID-19 had not occurred there is no other reason to believe there will have been a systematic change in the satisfaction level of first year university students.

Our study covers three main cohorts of students. The pre-COVID-19 lockdown cohort: first year students who entered university in October 2019 and completed their first two terms of university before the onset of COVID-19. The full COVID-19 lockdown cohort: first year students who entered university in October 2020 and completed their first two terms of university during the full COVID-19 lockdown. The partial COVID-19 lockdown cohort: first year students who entered university in October 2021 and completed their first two terms of university during a partial-lockdown as COVID-19 restrictions were being lifted. In all three cohorts students are asked a range of questions about their current satisfaction. The current satisfaction respondents act as control respondents. The treated students, in the full and partial-lockdown cohorts, are randomly assigned to answer a set of questions that asks them to imagine how satisfied they would have been in absence of COVID-19.

We present three main findings. By comparing treated and control students in various ways, our results show that students: i) imagine that university life would have been substantively better in the absence of COVID-19, ii) report the lowest levels of current satisfaction in domains where they imagined life in the absence of COVID-19 life would be greatest, and iii) report imagined satisfaction levels in the absence of COVID-19 that seem to be unrealistically high.

This study follows the literature on potentially biased counterfactual reasoning initiated by

---

3Within the context of the UK university lockdown included the movement to e-learning alongside national measures which prohibited social interaction.
Kahneman and Tversky (1982), which develops the notion that individuals mentally simulate alternatives to reality when imagining what might have been, which can lead to bias. Our own work brings this notion of counterfactual reasoning to the literature that illustrates how big events, such as a job loss or health shock, impact satisfaction.\(^4\) Studies have tested how big events can impact current satisfaction (Clark and Oswald, 2002), perceptions of past satisfaction (O’Brien et al., 2012), as well as how effective people are at predicting future satisfaction if a big event occurs (Frey and Stutzer, 2014). Our work does so by asking how the imagined satisfaction of an alternative current state impacts current satisfaction. The importance of reference points that individuals judge themselves against when reporting satisfaction has been highlighted (see, for example, Hauret and Williams, 2019). This paper adds to this literature by testing how imagined satisfaction in an alternative state of the world potentially acts as an internal anchor and impacts current satisfaction. We also add to the methodological literature (see Graham and Coppock, 2021, for an overview of ‘imagine if this event didn’t occur’ style questions) by testing how accurately students can judge what satisfaction could currently be if a big event had not occurred. Finally, our results provide direct evidence which might support the literature on motivated reasoning (Bénabou and Tirole, 2006, 2016; Bénabou, 2015) which typically assumes that memory and/or beliefs are to some degree malleable by showing the extent to which counterfactual reasoning, based partly on memory and partly on imagination, can be subject to bias.

2 Satisfaction questions and forms of variation we exploit

2.1 Satisfaction questions

The satisfaction questions are split into two main types: i) current satisfaction and ii) imagined satisfaction in the absence of COVID-19. Both sets of satisfaction questions are asked across a range of domains: university life, physical health, mental health, social life, sleep, accommodation, and financial situation.\(^5\)

\(^4\)For a general overview on the literature on life satisfaction see Benjamin et al. (2014).

\(^5\)A full script can be found in Appendix B.
The current satisfaction with university life question asks subjects: *In general, between the start of term and now, I have been [enter a value on the slide between 0 - completely unsatisfied - to 100 - completely satisfied] with my ‘university life’.* The questions on the other domains are very similar and have the form: *In general, between the start of term and now, I have been [enter a value on the slide between 0 - completely unsatisfied - to 100 - completely satisfied] with my ‘[enter domain here]’*. The seven questions were asked in week four of each term so the reference period is the previous four weeks. The cohort was also asked equivalent questions about their satisfaction with their physical health, mental health, social life, sleep, accommodation, and financial situation where they were specifically directed to answer as if the COVID-19 situation did not exist.6

The imagined satisfaction question with university life is instead phrased as: *Imagine the whole COVID-19 situation DID NOT exist. In general, between the start of term and now, I think I would have been [enter a value on the slide between 0 - completely unsatisfied - to 100 - completely satisfied] with my ‘university life’.* The questions on the other domains are very similar and have the form: *Imagine the whole COVID-19 situation DID NOT exist. In general, between the start of term and now, I think I would have been [enter a value on the slide between 0 - completely unsatisfied - to 100 - completely satisfied] with my ‘[enter domain here]’*. The questions were asked in week four of each term so the reference period is the previous four weeks, and the domains include imagined non-lockdown satisfaction in physical health, mental health, social life, sleep, accommodation, and the respondent’s financial situation.

### 2.2 Survey timing

This paper uses a longitudinal survey which covers the onset and continuing influence COVID-19 on student life. There are six waves that cover the period from October 2019 to January 2022. The key survey questions tracks students’ current satisfaction as well as how students imagine their satisfaction would have been in the absence of COVID-19.

Figure 1 provides the timing of the key satisfaction questions. There are three main cohorts of first year students. The pre-lockdown cohort (the top two blocks) were asked current satisfaction

---

6The precise wording can be found in Appendix B, in Figures A8 and A9.
Figure 1: Timeline of key satisfaction questions

T1 current satisfactions Qs  T2 current satisfactions Qs
Wave 1: Oct 2019  Wave 2: Feb 2020

Control  T1 current satisfactions Qs  T2 current satisfactions Qs
Treatment  T1 imagined satisfactions Qs  T2 imagined satisfactions Qs
Wave 3: Oct 2020  Wave 4: Feb 2021

Control  T1 current satisfactions Qs  T1 current satisfactions Qs
Treatment  T1 imagined satisfactions Qs  T2 imagined satisfactions Qs
Wave 5: Oct 2021  Wave 6: Feb 2022

questions in term 1 (Wave 1: October 2019) and term 2 (Wave 2: February 2020). The full-lockdown cohort (the middle four blocks) were randomly asked either current satisfactions questions (control) or imagined satisfaction questions (treatment) in term 1 (Wave 3: October 2020) and term 2 (Wave 4: February 2021). The partial-lockdown cohort (the right four blocks) were randomly asked either current satisfactions questions (control) or imagined satisfaction questions (treatment) in term 1 (Wave 5: October 2021) and term 2 (Wave 6: February 2022).

2.3 First source of variation: A survey experiment

In this paper we are concerned with students’ perceived difference between satisfaction in the current COVID-19 situation and what would have been the case had COVID-19 not occurred. To test for the presence and size of this perceived difference we randomized the two types of satisfaction questions (current and imagined) in waves 3, 4 and 5. In the treatment group, students were asked questions on their imagined satisfaction in the absence of COVID-19. In the control group, students
were asked questions on their current satisfaction. Imagined satisfaction questions were not asked in waves 1 and 2 as COVID-19 was not anticipated at this time. These comparisons will be valid under the assumption that the randomization balanced observable and unobservable characteristics between students in the treatment and control. The treatment and control groups are shown in the top and bottom boxes in Figure 1.

2.4 Second source of variation: A natural experiment

We can also use the arrival of COVID-19 and the associated impact on student life as a natural experiment to try and test how ‘realistic’ responses to the imagined satisfaction questions might be. This is useful as the set of imagined satisfaction questions that are asked to students are purely hypothetical: there is no way to know how satisfied one would have been had COVID-19 not occurred. However, we can ask how similar are the imagined satisfaction levels reported by students in the COVID-19 cohort, as compared to the actual satisfaction reported by the pre-COVID-19 cohort. In relation to the timeline in Figure 1, this means that we can compare: i) the responses of the ‘imagined satisfaction’ questions in Wave 3 with the responses of the ‘current satisfaction’ questions in Wave 1 (term 1 comparisons), and ii) the responses of the ‘imagined satisfaction’ questions in Wave 4 with the responses of the ‘current satisfaction’ questions in Wave 2 (term 2 comparisons). These two comparisons will be valid under the assumption that Wave 1 (and 2) are similar to Wave 3 (and Wave 4) students.

3 Hypotheses

In general, a students’ current reported satisfaction with university life \( S \) can be represented as a weighted average \( S = \sum_{i \in \Theta} \phi_j r(C_j, I_j) \), where \( \Theta \) is the set of possible inputs to satisfaction containing a range of domains (such as social life, physical health, mental health), \( \phi_j \geq 0 \) for all \( j \), represents weights on each domain reflecting how much each contributes to reported satisfaction (the greater the weight, the more important the domain to student’s satisfaction with university life), and \( r(C_j, I_j) \) is a function that captures the students’ satisfaction with their current circum-
stance in a particular input \( (C_j) \) relative to some imagined circumstance in that input \( (I_j) \), such as if there was no COVID-19 lockdown.\(^7\)

For any given domain, students’ current satisfaction can partially reflect how the students’ current circumstances compare to a counterfactual situation they consider as an alternative. For instance, students’ satisfaction with their current social life during lockdown may be impacted by how they imagined their social life could be in a non-COVID-19 lockdown scenario.\(^8\) This may happen via ‘contrast effects’ which occur when a feeling is made more extreme when judged against some salient anchor (Plous, 1993). In other words, having a good social life with lockdown will feel good, but if someone imagines a better social life without lockdown, current social life will not feel quite so good. If the anchor provides an individual with an ‘upward counterfactual’ (I could have had a great social life in absence of lockdown) satisfaction is likely to be lesser than for an individual for whom the anchor provides a ‘downward counterfactual’ (I could have had a bad social life in absence of lockdown). Using the idea of contrast effects we make two hypotheses.

**Hypothesis 1** - relates to the **direction** of the relationship between imagined satisfaction in the absence of COVID-19 and current satisfaction \(: \) In life domains where the imagined non-lockdown comparison provides an upward (downward) counterfactual for the current lockdown circumstances, satisfaction will be lower (higher).

**Hypothesis 2** - relates to the **magnitude** of the relationship between imagined satisfaction in the absence of COVID-19 and current satisfaction \(: \) When individuals experience negative shocks such

---

\(^7\)Recent work on subjective wellbeing (Benjamin et al., 2014, for instance) suggests that when individuals report satisfaction about life in general the answer is a weighted average of satisfaction across salient inputs, such as income and health. Further, satisfaction with each input can be driven by an evaluation of that input against a salient anchor, such as a neighbour’s income, or the individual’s health one year ago.

\(^8\)The basic idea can be explained through a utility framework. Suppose a student’s satisfaction in an input is a function of their current circumstances \( (C) \) and some imagined alternative scenario \( (I) \): \( r = r(C, I) \). For instance, \( C \) represents the student’s current social life during the lockdown, while \( I \) represents the social life they imagine in a non-COVID-19 lockdown scenario. We assume that as the student attains more, or better quality, of \( C \) (more/better social life) satisfaction increases at a diminishing rate \( (r_C > 0 \text{ and } r_{CC} < 0) \). In the setting of this paper, the imagined scenario cannot be realized since the student had no influence over the reality of lockdown. Therefore, we assume that as the student’s imagined amount or quality of \( I \) increases (more/better social life) satisfaction falls at a diminishing rate \( (r_I < 0 \text{ and } r_{II} > 0) \). If the student’s amount/quality of the imagined non-lockdown circumstance (the bad) increases then, to hold satisfaction constant, the current amount/quality of the current lockdown circumstance (the good) needs to increase, otherwise satisfaction will fall. To see this, set the total derivative of \( r \) equal to zero to give \( dr = r_C dC + r_I dI = 0 \), such that, \( dr \rangle = - \frac{r_C}{r_I} > 0 \) given \( r_C > 0 \) and \( r_I < 0 \). Setting up current versus imagined circumstances in this manner is consistent with past literature that shows imagining counterfactual ‘what ifs’ can have powerful consequences for current satisfaction.
a COVID-19 lockdown, the greater they think their satisfaction would have been in absence of the shock, the lower their current satisfaction.

4 Descriptive statistics

There are two key types of variation we exploit in this paper: i) the survey experiment used to compare current and imagined satisfaction and ii) the natural experiment used to compare imagined satisfaction in the absence of COVID-19 with past levels of actual satisfaction. Figure 2 presents satisfaction responses from wave 1 and 3 to illustrate how these comparisons can be made. The left histogram presents wave 1 (pre-COVID-19, Oct 2019) responses to the question on current satisfaction with university life. The middle histogram presents wave 3 (full lockdown, Oct 2020) responses to the question on current satisfaction with university life. The right histogram presents wave 3 (full lockdown, Oct 2020) responses to the question on imagined satisfaction in absence of COVID-19 with university life.

Figure 2: Histograms of satisfaction and imagined satisfaction

Notes: The left and middle panels are the histograms of satisfaction responses for university life in wave 1 and wave 3. The question on satisfaction with university life is: In general, between the start of term 1 and now, I have been [enter response 0-100] with my university life. The right panel is the imagined histogram of satisfaction responses for wave 3. The imagined satisfaction question is: Imagine the whole COVID situation DID NOT exist. In general, between the start of term and now, I think I would have been [enter a value on scale between 0 - completely unsatisfied - to 100 - completely satisfied] with my ‘university life’.

The first comparison, based on the survey experiment, compares the right (imagined satisfac-
tion) and the middle distribution (current satisfaction). By comparing these distributions we see students imagine that their satisfaction would have been substantially higher had COVID-19 not occurred: the right-hand distribution is shifted to the right compared to the middle distribution and has a mean (the vertical dashed lines) of 80/100 compared to 60/100. The second comparison, based on a natural experiment, compares the right and the left distribution. Students also imagined that their satisfaction would have been higher than the actual satisfaction reported by a comparable first year who studied in the year before COVID-19 occurred: the right-hand distribution is shifted to the right compared to the middle distribution and has a mean of 80/100 compared to 70/100.

The descriptive statistics for the estimation samples are provided in Table 2 in Appendix A. Across all 6 waves, roughly 60% of students are males, 40% from the UK, 40% from outside the EU, and the remainder within the EU but outside the UK. These characteristics are similar to the broader population of social science students.

4.1 Testing differences in actual versus imagined satisfaction: a survey experiment

We suggested that satisfaction (S) is a weighted sum of possible inputs $S = \sum_{j \in \Theta} \phi_j r(C_j, I_j)$. We first model the difference between actual and imagined satisfaction with general university life in each relevant wave using the regression specification:

$$S_{i,t} = \alpha_0 + \alpha_1 \text{Imagine}_{i,t} + X' \gamma + \epsilon_{i,t} \quad (1)$$

where $S_{i,t}$ is reported satisfaction with general university life of student $i$ in wave $t$ when asked about imagined non-lockdown satisfaction ($\text{Imagine}_{i,t} = 1$) or actual satisfaction ($\text{Imagine}_{i,t} = 0$). The error term is $\epsilon_{i,t}$. There is one model for each of wave 3, 4, and 5, which are the waves where imagined satisfaction questions were asked. Given that students were randomly assigned the imagined or actual satisfaction questions, we expect our treatment variable to be mean independent of the error: $E[\epsilon|\text{Imagine}] = E[\epsilon]$. That is, there is no reason to suspect that the imagined
satisfaction of students asked about imagined satisfaction should be systematically different from
the imagined satisfaction of students asked the current satisfaction question, had they been asked
the imagined satisfaction question. To help with precision, and as a robustness check, we estimate
a set of models with controls ($X$) included.

Model (1) will be estimated for three waves (3, 4 and 5), general satisfaction with university life
as well as six further domains (which may all contribute to general university life): physical health,
mental health, social life, sleep, accommodation, and financial situation. That is, there will be 21
models, one for each wave-domain. In total there are 21 parameters representing the difference
between imagined and actual satisfaction, which we reference as $\alpha_k^1$ for all $k = 1, \ldots, 21$. In terms
of the hypotheses set out in Section 3 we expect: i) $\alpha_k^1 > 0$ for all domains that provide an upward
counterfactual (imagined satisfaction is higher than current satisfaction) and ii) there will be an
inverse relationship between current satisfaction and the size of the gap between imagined and
current satisfaction ($\alpha_k^1$) when satisfaction responses are ordered from lowest to highest current
satisfaction.

### 4.2 Testing the realism of imagined satisfaction: a natural experiment

To test for possible over/no/under-exaggeration of responses to the imagined satisfaction question,
we compare the difference between imagined satisfaction in the absence of COVID-19 with actual
satisfaction of students who studied before COVID-19 for the nearest relevant wave. For example,
we can compare imagined satisfaction of those in wave 3 (October 2020) with actual satisfaction
of those in wave 1 (October 2019):

$$S_{i,t=1,3} = \gamma_0 + \gamma_1 \text{Im} \gamma_{i,t=3} + X'\gamma + \epsilon_{i,t=1,3}$$  (2)

where $S_{i,t=1,3}$ is the actual satisfaction of student $i$ in wave $t = 1$ ($\text{Im} \gamma_{i,t=3} = 0$) or imagined
satisfaction of student $i$ in wave $t = 3$ ($\text{Im} \gamma_{i,t=3} = 1$). The error term is $\epsilon_{i,t=1,3}$. In model
(2) $\gamma_1$ is an estimator of the difference between imagined satisfaction in the absence of COVID-

---

9 This assumption is supported by the balance check in Table 1: all variables are statistically similar between
treatment and control groups.
19 in term 1 during full-lockdown and actual satisfaction in term 1 pre-lockdown. We can also compare imagined satisfaction in the absence of COVID-19 in wave 4 (February 2021) versus actual satisfaction in term 2 pre-lockdown (February 2021). To check for possible imbalances in characteristics across waves a set of controls ($X$) are included in each model.

5 Results

5.1 Balancing of covariates within the: i) survey experiment and ii) natural experiment

In this paper we make two sets of comparisons based on a survey experiment (section 3.1) and a natural experiment (section 3.2). To check for evidence that students in the treatment and control in each set up are similar we test for statistical balance over a range of characteristics. Table 1 presents balancing tests between gender and overseas status as well as academic expectations that were collected in the same survey.

<table>
<thead>
<tr>
<th></th>
<th>Randomisation: actual v. imagined</th>
<th>Natural experiment: base v. imagined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>obs control obs treatment diff. s.e. p-value</td>
<td>obs control obs treatment diff. s.e. p-value</td>
</tr>
<tr>
<td>Individual characteristics:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>male($=1$)</td>
<td>120 0.575 124 0.605 -0.030 0.063 0.637</td>
<td>239 0.598 124 0.605 -0.007 0.054 0.905</td>
</tr>
<tr>
<td>UK($=1$)</td>
<td>120 0.492 124 0.573 -0.081 0.064 0.207</td>
<td>239 0.431 124 0.573 -0.142 0.055 0.010</td>
</tr>
<tr>
<td>Academic expectations:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mark expectations</td>
<td>111 72.018 115 69.487 2.531 1.481 0.170</td>
<td>202 71.847 115 69.487 2.360 1.508 0.119</td>
</tr>
<tr>
<td>hours of needed work</td>
<td>110 6.256 115 6.365 -0.109 0.375 0.772</td>
<td>228 6.426 115 6.365 0.061 0.871 0.944</td>
</tr>
<tr>
<td>needed concentration</td>
<td>111 81.820 115 82.191 -0.371 1.470 0.801</td>
<td>233 82.927 115 82.191 0.736 1.294 0.570</td>
</tr>
</tbody>
</table>

Notes: The left hand panel presents the balance of characteristics in the randomization of wave 3 between students who answered the current satisfaction questions (termed the ‘actual’ group) and those who answered the imagined satisfaction questions (termed the ‘imagined group’). The right hand panel presents the balance of characteristics in the natural experiment between students who answered the current satisfaction questions in wave 1 (termed the ‘base’ group) and those who answered the imagined satisfaction questions in wave 3 (termed the ‘imagined group’).

The left panel presents the balancing tests for the survey experiment for wave 3 of the survey: all
characteristics balance. For example, the top row shows that in the control 58% of the sample were male and in the treatment 60% were male, and the difference of $-0.03$ is statistically insignificant (with a standard error of 0.06 and a p-value of 0.637). The rest of the characteristics also balance. This indicates that the treatment and control groups in the survey experiment are comparable.

The right panel presents the balancing tests for the natural experiment comparing wave 1 against wave 3. There is no statistical difference in genders, however, respondents in wave 3 are more likely to be home students (when we control for overseas status in our regressions it has no impact on the results). There are no statistical differences in academic expectations. As a robustness check, we will control for overseas status, along with the other characteristics, in the main regressions results.

5.2 Results from the survey experiment: actual versus imagined satisfaction

Our first key results is that student satisfaction dropped substantively during COVID-19. Figure 3 plots average current satisfaction (the black dots with 95% confidence bars). The black dots show that in pre-COVID-19, wave one (October 2019) and two (February 2020), satisfaction is approximately 70. Average current satisfaction then dipped to 60 and then 40 during term 1 and term 2 of the COVID-19 lockdown, which are waves 3 (October 2020) and 4 (February 2021). Average current satisfaction then rose when COVID-19 restrictions at university began to be removed to about 65 in waves 5 (October 2021) and 6 (February 2022).

Our second key result is that students imagined that they would have been more satisfied had COVID-19 not occurred. Figure 3 also plots average imagined satisfaction in the absence of COVID-19 (the red dots with 95% confidence bars). In all waves where imagined satisfaction question are asked (wave 3, 4, and 5), average actual satisfaction lies below average imagined satisfaction. For example, in wave 3 (October 2020) actual satisfaction is 60 while imagined satisfaction in the absence of COVID-19 is 80. The mean difference of $-20$ for wave 3 (along with its 95% confidence bar) is presented in green. The dashed horizontal line in Figure 3 represents the line of no difference between actual and imagined satisfaction. In wave 4 (February 2021)
actual satisfaction (40) was 40 units below imagined satisfaction in the absence of COVID-19 (80). In wave 5 (October 2021) actual satisfaction (65) was 10 units below imagined satisfaction in the absence of COVID-19 (75). For all three waves, the means are below zero as are the 95% confidence bars: actual satisfaction lies below imagined satisfaction.

Figure 3: Actual, imagined, and benchmark satisfaction with general university life

The green dots and bars are equivalent to the treatment effect point estimates and 95% confidence intervals from the OLS estimates of model (1). Controlling for observable characteristics has no substantive impact on the estimated differences. See Table 4 in Appendix A for the full regression results.

The broad patterns seen for average satisfaction with general university life are also reflected in other domains. Figure 4 is a similar plot to Figure 3 but plots satisfaction measures over all domains: general university life, physical health, social life, accommodation, mental health, financial situation, and sleep. The far left plot repeats Figure 3 with black vertical lines splitting the responses in the pre-COVID-19 lockdown (waves 1 and 2), full COVID-19 lockdown (wave 3 and 4), and partial COVID-19 lockdown (waves 5 and 6). A similar pattern to satisfaction with general university life can be seen for satisfaction with physical health, social life, accommodation, and mental health. The two exceptions to the general pattern are satisfaction with financial situation.
and sleep. Controlling for observable characteristics has no substantive impact on the estimated differences: see Table 4 in Appendix A for the full regression results.

![Figure 4: Actual, imagined, and benchmark satisfaction across all domains](image)

**5.3 Discussion of the results on actual versus imagined satisfaction**

In this section we discuss what may be driving the differences we observed between actual and imagined satisfaction in the absence of COVID-19. To consider whether the evidence in Figures 3 and 4 supports hypothesis 1, it is useful to discuss what students consider when answering satisfaction questions. For each input, all else constant, the larger the upward (downward) counterfactual, the salient anchor provides for the current circumstance, the more (less) satisfaction that input will contribute to overall satisfaction. Further, the more (less) specific the satisfaction question is, the less (more) scope there is for the number of salient inputs and anchors. If a question is very broad in nature, for example about general university life, many academic and non-academic inputs may contribute to a possible answer. If the question is very specific, for example about the enjoyment of a particular lecture, the inputs are likely to be fewer.
The evidence in Figures 3 and 4 support hypothesis 1: if imagined non-lockdown comparisons provide an upward (downward) counterfactual for current lockdown circumstances, current satisfaction will be lower. Non-lockdown life is likely to provide an upward counterfactual for most inputs into satisfaction with university life. Without lockdown, students were able to choose to attend in-person lectures/classes, visit the library, and become involved in extra-curricula academic and non-academic activities. Lockdown restricted the majority of academic and non-academic choices, likely making non-lockdown life look better by comparison (an upward counterfactual). This may be part of the reason why satisfaction in lockdown is lower than satisfaction in the comparable time pre-lockdown. In all domains where the imagined non-COVID inputs provide mainly upward counterfactuals, we see the same pattern: drops in lockdown satisfaction compared to pre-lockdown satisfaction in social life, physical health, mental health, and accommodation. This provides supportive evidence for hypothesis 1.

There is no drop in satisfaction with sleep and finance between pre-lockdown and lockdown satisfaction, which seems to run contrary to hypothesis 1. However, if students are considering various inputs and salient anchors, the upward and downward counterfactuals relevant to the sleep and finance domains may cancel each other out. Assume the more hours of sleep students
get the more satisfied they are with sleep. In lockdown, on the one hand, students may get more sleep due to fewer opportunities to socialise (imagined non-lockdown satisfaction provides a downward counterfactual), while, on the other hand, increased mental and physical health issues may reduce sleep (imagined non-lockdown satisfaction provides an upward counterfactual). A similar combination may occur when students respond to questions on satisfaction with finance. For instance, students may require less money to socialise (imagined non-lockdown provides an upward counterfactual), but also have fewer opportunities to work to earn money (imagined non-lockdown provides a downward counterfactual). It may also simply be that students’ satisfaction with finances did not change between pre-lockdown and lockdown as they were not impacted financially by the lockdown.

Hypothesis 2 predicts that in life domains where the imagined non-lockdown comparison provides a bigger upward counterfactual for the current lockdown circumstances, the lower current satisfaction will be lower. Following model (1) in Section 4.1, for each wave-domain, let $s^k_a$ represent the actual satisfaction reported for all $k = 1, ..., 21$ and let $\alpha^k_1$ represent the gap between actual and imagined satisfaction in the absence of COVID-19. Hypothesis 2 suggests that if we order $s^k_a$ from lowest to highest satisfaction, then $\alpha^k_1$ will be ordered from smallest gap to highest gap. To test this Figure 5 re-arranges the data to look for evidence for hypothesis 2.

Figure 5 is constructed by: i) selecting all waves and domains where both imagined and actual satisfaction are asked, - waves 3, 4, and 5 across all domains -, ii) ordering them along the x-axis from lowest actual satisfaction to highest actual satisfaction - i.e. the black dots -, and iii) adding lines of best fit through the ordered data. We argue that Figure 5 provides supportive evidence of hypothesis 2: the upward sloping green line of best fits shows, on average, that in wave-domains with the lowest satisfaction are those with the biggest difference between actual and imagined satisfaction. That is, in life domains where the imagined non-COVID-19 lockdown comparison provides a larger upward (downward) counterfactual in comparison to current COVID-19 lockdown circumstances, the lower/higher current satisfaction will be.
5.4 Results from the natural experiment: bench-marking imagined satisfaction

The COVID-19 lockdown also provides us with a unique opportunity to test how realistic students’ imagined non-COVID-19 lockdown satisfaction might be. We have already shown that: i) imagined satisfaction lies above current satisfaction and ii) the larger the gap between imagined and current the lower current satisfaction is likely to be. Next we ask whether their imagined levels of satisfaction are realistic. That is, does the fact that students are currently going through COVID-19 lockdown remove their ability to realistically imagine satisfaction with university life without the COVID-19 lockdown. To do this we compare the imagined non-COVID-19 lockdown satisfaction of the 2020/21 COVID-19 lockdown cohort with the actual satisfaction of the 2019/20 pre-COVID-19 lockdown cohort. If the imagined non-COVID-19 lockdown satisfaction of the 2020/21 cohort is greater/lesser than the actual non-COVID-19 lockdown satisfaction of the 2019/20 cohort then students over/under estimate their non-COVID-19 lockdown satisfaction. We argue this is a valid comparison as if COVID-19 had not occurred there is no reason to believe there will have been a structural shock that impacted satisfaction levels. Ex-ante, it is not clear whether students will over or underestimate imagined satisfaction.

We can use Figure 3 to show that students typically exaggerate how satisfied they would have been had they not been in COVID-19 lockdown. The blue dots (and 95% confidence bars) compare the imagined non-lockdown satisfaction of the 2020/21 lockdown cohort with the actual satisfaction of the 2019/20 pre-lockdown cohort. For example, in wave three (October 2020) students imagined their non-lockdown satisfaction would be 80 which is 10 units higher than the satisfaction in wave one (October 2019, pre-lockdown). We interpret this as students over-exaggerating how satisfied they would have been by 10 units. A similar result is shown in wave 4 (a 10-unit exaggeration) and wave 5 (a 5-unit exaggeration). Further, Figure 5 plots a gently upward sloping blue line providing suggestive evidence that in wave-domains with the lowest satisfaction/biggest gap there exists the greatest over-exaggeration. It is possible that this is not over-exaggeration and that the current lockdown cohort would have been happier than those who went before them. We suggest this is not likely as there is no reason to believe one first year cohort will be systematically happier
than a previous cohort and in regressions that control for any differences in characteristics the estimates of over-exaggerated happiness remain unchanged. We argue it is more likely that the over-exaggeration we are estimating is a flip side to current levels of low satisfaction: if you find yourself in a low satisfaction state you imagine any state other than the one you are in would be superior.

6 Conclusion

Does imagining what life could have been in absence of a shock impact current satisfaction? To answer this question we use COVID-19 as a natural experiment together with a randomized control trial administered through a survey designed to generate reported measures of life satisfaction from students suffering the effects of the COVID-19 period. The survey covers pre-COVID-19 lockdown, full COVID-19 lockdown, and partial COVID-19 lockdown. We find that imagining life in absence of a shock (COVID-19) can impact current satisfaction: when individuals experience negative shocks, the higher they think their satisfaction would have been in absence of the shock, the lower their current satisfaction. In other words this seems to be a validation of the well-known proverb “the grass is always greener on the other side of the fence” or the very similar Ovid quotation presented at the start of the introduction. As suggested by the proverb, counterfactual reasoning of this sort may be misleading. In the case of our study this implies that individuals may over-exaggerate how satisfied they would have been if a negative shock had not happened.

Our results suggest that any policy which can correct this over-exaggeration would also boost current satisfaction. Whether this is best served by providing general education on the potential biases that may feature in counterfactual reasoning, or more specific attempts to target misinformation during any large shock, we leave for further study.
References


**Online Appendix**

**Appendix A: Additional Tables**

Table 2: Descriptive statistics of first years by wave

<table>
<thead>
<tr>
<th></th>
<th>Wave 1</th>
<th>Wave 2</th>
<th>Wave 3</th>
<th>Wave 4</th>
<th>Wave 5</th>
<th>Wave 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs. mean s.d.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual demographics:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male(=1)</td>
<td>239</td>
<td>0.60</td>
<td>0.49</td>
<td>244</td>
<td>0.59</td>
<td>0.49</td>
</tr>
<tr>
<td>home/UK(=1)</td>
<td>239</td>
<td>0.43</td>
<td>0.50</td>
<td>244</td>
<td>0.53</td>
<td>0.50</td>
</tr>
<tr>
<td>home/EU(=1)</td>
<td>239</td>
<td>0.10</td>
<td>0.31</td>
<td>244</td>
<td>0.13</td>
<td>0.33</td>
</tr>
<tr>
<td>overseas(=1)</td>
<td>239</td>
<td>0.46</td>
<td>0.50</td>
<td>244</td>
<td>0.34</td>
<td>0.47</td>
</tr>
<tr>
<td>Satisfaction:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>general university</td>
<td>239</td>
<td>71.39</td>
<td>17.38</td>
<td>120</td>
<td>60.43</td>
<td>21.35</td>
</tr>
<tr>
<td>physical health</td>
<td>239</td>
<td>66.21</td>
<td>21.80</td>
<td>120</td>
<td>59.54</td>
<td>23.35</td>
</tr>
<tr>
<td>social life</td>
<td>239</td>
<td>70.20</td>
<td>20.35</td>
<td>120</td>
<td>53.68</td>
<td>24.05</td>
</tr>
<tr>
<td>accommodation</td>
<td>239</td>
<td>75.33</td>
<td>19.49</td>
<td>120</td>
<td>66.18</td>
<td>26.78</td>
</tr>
<tr>
<td>mental health</td>
<td>239</td>
<td>69.20</td>
<td>23.19</td>
<td>120</td>
<td>60.36</td>
<td>23.32</td>
</tr>
<tr>
<td>finance</td>
<td>239</td>
<td>67.60</td>
<td>24.37</td>
<td>120</td>
<td>67.29</td>
<td>23.69</td>
</tr>
<tr>
<td>sleep</td>
<td>239</td>
<td>56.05</td>
<td>22.99</td>
<td>120</td>
<td>55.24</td>
<td>25.65</td>
</tr>
<tr>
<td>Imagined satisfaction:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>general university</td>
<td>124</td>
<td>79.69</td>
<td>14.94</td>
<td>62</td>
<td>82.02</td>
<td>14.67</td>
</tr>
<tr>
<td>physical health</td>
<td>124</td>
<td>73.42</td>
<td>19.27</td>
<td>62</td>
<td>78.98</td>
<td>16.93</td>
</tr>
<tr>
<td>social life</td>
<td>124</td>
<td>80.39</td>
<td>16.28</td>
<td>62</td>
<td>85.18</td>
<td>13.79</td>
</tr>
<tr>
<td>accommodation</td>
<td>124</td>
<td>76.44</td>
<td>17.35</td>
<td>62</td>
<td>79.00</td>
<td>19.42</td>
</tr>
<tr>
<td>mental health</td>
<td>124</td>
<td>72.74</td>
<td>20.84</td>
<td>62</td>
<td>76.00</td>
<td>16.90</td>
</tr>
<tr>
<td>finance</td>
<td>124</td>
<td>65.12</td>
<td>24.12</td>
<td>62</td>
<td>72.85</td>
<td>23.22</td>
</tr>
<tr>
<td>sleep</td>
<td>124</td>
<td>58.72</td>
<td>22.01</td>
<td>62</td>
<td>62.21</td>
<td>19.30</td>
</tr>
</tbody>
</table>

Notes: descriptive statistics by wave. Wave 1 and 2 are the the pre-lockdown cohort: the 2019/20 cohort of first year students who completed their first two terms of university pre-lockdown. Waves 3 and 4 are the full-lockdown cohort: the 2020/21 cohort of first year students who completed their first two terms of university during lockdown. Waves 5 and 6 are the partial-lockdown cohort: the 2021/22 cohort of first year students who completed their first two terms of university during lockdown. For the pre-lockdown cohort we survey the students in two terms, which are 10 weeks in length. Term 1 runs from early October 2019 to mid December 2019. Term 2 runs from mid January 2020 to late March 2020. For the full-lockdown cohort we survey the students in two terms which are also 10 weeks in length. Term 1 runs from early October 2020 to mid December 2020. Term 2 runs from mid January 2021 to late March 2021. For the partial-lockdown cohort we survey the students in two terms which are also 10 weeks in length. Term 1 runs from early October 2021 to mid December 2021. Term 2 runs from mid January 2022 to late March 2022.
## Table 3: Satisfaction over time: with and without controls

<table>
<thead>
<tr>
<th>Survey wave</th>
<th>gen uni</th>
<th>soci</th>
<th>phy</th>
<th>ment</th>
<th>sleep</th>
<th>acc</th>
<th>fina</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1.568)</td>
<td>(0.353)</td>
<td>(0.060)</td>
<td>(0.336)</td>
<td>(0.380)</td>
<td>(0.072)</td>
<td>(0.077)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>(2.233)</td>
<td>(2.272)</td>
<td>(2.544)</td>
<td>(2.544)</td>
<td>(2.566)</td>
<td>(2.531)</td>
<td>(2.738)</td>
<td>(2.782)</td>
</tr>
<tr>
<td>(2.629)</td>
<td>(2.615)</td>
<td>(2.769)</td>
<td>(3.052)</td>
<td>(2.927)</td>
<td>(2.874)</td>
<td>(2.772)</td>
<td>(2.714)</td>
</tr>
<tr>
<td>(2.068)</td>
<td>(2.078)</td>
<td>(2.688)</td>
<td>(2.738)</td>
<td>(2.794)</td>
<td>(2.595)</td>
<td>(2.630)</td>
<td>(2.836)</td>
</tr>
<tr>
<td>Oct. 2019 aver. satis.</td>
<td>71.4</td>
<td>71.4</td>
<td>70.06</td>
<td>70.02</td>
<td>66.04</td>
<td>66.19</td>
<td>69.43</td>
</tr>
<tr>
<td>Observations</td>
<td>782</td>
<td>782</td>
<td>784</td>
<td>781</td>
<td>782</td>
<td>779</td>
<td>784</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.15</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.03</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td>Individual Characteristics</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
</tbody>
</table>

Notes: The regression estimates in the table are for the specification $s_{i,t} = \delta + Z'\pi + X'\gamma + \epsilon_{i,t}$ where $Z$ is a vector of time dummies $\{OCT \ 2020, FEB \ 2021, OCT \ 2021, FEB \ 2022\}$ and $s_{i,t}$ is the response to the relevant satisfaction questions on general satisfaction, social life, physical health, mental health, sleep, accommodation, finance. Wave 1 and 2 are the the pre-lockdown cohort: the 2019/20 cohort of first year students who completed their first two terms of university pre-lockdown. Waves 3 and 4 are the full-lockdown cohort: the 2020/21 cohort of first year students who completed their first two terms of university during lockdown. Waves 5 and 6 are the partial-lockdown cohort: the 2021/22 cohort of first year students who completed their first two terms of university during lockdown. For the pre-lockdown cohort we survey the students in two terms, which are 10 weeks in length. Term 1 runs from early October 2019 to mid December 2019. Term 2 runs from mid January 2020 to late March 2020. For the full-lockdown cohort we survey the students in two terms which are also 10 weeks in length. Term 1 runs from early October 2020 to mid December 2020. Term 2 runs from mid January 2021 to late March 2021. For the partial-lockdown cohort we survey the students in two terms which are also 10 weeks in length. Term 1 runs from early October 2021 to mid December 2021. Term 2 runs from mid January 2022 to late March 2022.
## Table 4: Actual versus imagined satisfaction by wave: with and without controls

<table>
<thead>
<tr>
<th>Treatment in wave 3:</th>
<th>gen uni.</th>
<th>soc.</th>
<th>phy</th>
<th>ment</th>
<th>sleep</th>
<th>acc.</th>
<th>fina.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imaged satisfaction (yyy)</td>
<td>10.730***</td>
<td>19.920***</td>
<td>27.616***</td>
<td>28.120***</td>
<td>13.756***</td>
<td>13.312***</td>
<td>3.592</td>
</tr>
<tr>
<td>(2.346)</td>
<td>(4.035)</td>
<td>(3.221)</td>
<td>(3.520)</td>
<td>(3.464)</td>
<td>(3.482)</td>
<td>(3.474)</td>
<td>(2.830)</td>
</tr>
<tr>
<td>Average current sati.</td>
<td>69.85</td>
<td>69.85</td>
<td>66.65</td>
<td>66.65</td>
<td>66.65</td>
<td>66.61</td>
<td>56.7</td>
</tr>
<tr>
<td>observations</td>
<td>251</td>
<td>251</td>
<td>251</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>250</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.22</td>
<td>0.22</td>
<td>0.22</td>
<td>0.22</td>
<td>0.22</td>
<td>0.22</td>
<td>0.22</td>
</tr>
<tr>
<td>Individual Characteristics</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Treatment in wave 4:</td>
<td>gen uni.</td>
<td>soc.</td>
<td>phy</td>
<td>ment</td>
<td>sleep</td>
<td>acc.</td>
<td>fina.</td>
</tr>
<tr>
<td>Imaged satisfaction (yyy)</td>
<td>10.703***</td>
<td>10.045***</td>
<td>12.606***</td>
<td>11.744***</td>
<td>11.431***</td>
<td>10.139***</td>
<td>9.113***</td>
</tr>
<tr>
<td>(3.099)</td>
<td>(3.070)</td>
<td>(3.221)</td>
<td>(3.173)</td>
<td>(3.520)</td>
<td>(3.464)</td>
<td>(3.482)</td>
<td>(2.830)</td>
</tr>
<tr>
<td>Average current sati.</td>
<td>68.5</td>
<td>68.5</td>
<td>67.39</td>
<td>67.39</td>
<td>65.5</td>
<td>65.5</td>
<td>60.8</td>
</tr>
<tr>
<td>observations</td>
<td>173</td>
<td>173</td>
<td>173</td>
<td>173</td>
<td>173</td>
<td>173</td>
<td>173</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.06</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>Individual Characteristics</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
</tbody>
</table>

Notes: The regression estimates in the table are for the specification $S_{i,t} = \alpha_0 + \alpha_1 \text{Imagine}_{i,t} + X'\gamma + \epsilon_{i,t}$, where $S_{i,t}$ is reported satisfaction with general university life of student $i$ in wave $t$ when asked about imagined non-lockdown satisfaction ($\text{Imagine}_{i,t} = 1$) or actual satisfaction ($\text{Imagine}_{i,t} = 0$). There is one model for each of wave 3, 4, and 5, which are the waves where imagined satisfaction questions were asked. Wave 1 and 2 are the the pre-lockdown cohort: the 2019/20 cohort of first year students who completed their first two terms of university pre-lockdown. Waves 3 and 4 are the full-lockdown cohort: the 2020/21 cohort of first year students who completed their first two terms of university during lockdown. Waves 5 and 6 are the partial-lockdown cohort: the 2021/22 cohort of first year students who completed their first two terms of university during lockdown. For the pre-lockdown cohort we survey the students in two terms, which are 10 weeks in length. Term 1 runs from early October 2019 to mid December 2019. Term 2 runs from mid January 2020 to late March 2020. For the full-lockdown cohort we survey the students in two terms which are also 10 weeks in length. Term 1 runs from early October 2020 to mid December 2020. Term 2 runs from mid January 2021 to late March 2021. For the partial-lockdown cohort we survey the students in two terms which are also 10 weeks in length. Term 1 runs from early October 2021 to mid December 2021. Term 2 runs from mid January 2022 to late March 2022.
<table>
<thead>
<tr>
<th>gen uni.</th>
<th>soci.</th>
<th>phy</th>
<th>ment</th>
<th>sleep</th>
<th>acc.</th>
<th>fina.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment in wave 3 v. wave 1:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2.230)</td>
<td>(2.289)</td>
<td>(2.543)</td>
<td>(2.524)</td>
<td>(2.515)</td>
<td>(2.566)</td>
<td>(2.536)</td>
</tr>
<tr>
<td>[0.000]</td>
<td>[0.000]</td>
<td>[0.000]</td>
<td>[0.013]</td>
<td>[0.008]</td>
<td>[0.000]</td>
<td>[0.000]</td>
</tr>
<tr>
<td>Average benchmark sati. wave 1</td>
<td>71.4</td>
<td>71.4</td>
<td>70.02</td>
<td>70.02</td>
<td>66.19</td>
<td>66.19</td>
</tr>
<tr>
<td>observations</td>
<td>369</td>
<td>369</td>
<td>369</td>
<td>369</td>
<td>366</td>
<td>366</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.08</td>
<td>0.08</td>
<td>0.12</td>
<td>0.14</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>Individual Characteristics</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Treatment in wave 4 v. wave 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[0.000]</td>
<td>[0.000]</td>
<td>[0.000]</td>
<td>[0.000]</td>
<td>[0.000]</td>
<td>[0.002]</td>
<td>[0.001]</td>
</tr>
<tr>
<td>Average benchmark sati. wave 2</td>
<td>69.8</td>
<td>69.8</td>
<td>66.5</td>
<td>66.5</td>
<td>64.3</td>
<td>64.3</td>
</tr>
<tr>
<td>observations</td>
<td>195</td>
<td>195</td>
<td>195</td>
<td>195</td>
<td>195</td>
<td>195</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.31</td>
<td>0.32</td>
<td>0.31</td>
<td>0.31</td>
<td>0.68</td>
<td>0.13</td>
</tr>
<tr>
<td>Individual Characteristics</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Treatment in wave 5 v. wave 1:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2.626)</td>
<td>(2.592)</td>
<td>(2.743)</td>
<td>(2.685)</td>
<td>(3.648)</td>
<td>(3.053)</td>
<td>(2.927)</td>
</tr>
<tr>
<td>[0.002]</td>
<td>[0.003]</td>
<td>[0.002]</td>
<td>[0.002]</td>
<td>[0.002]</td>
<td>[0.003]</td>
<td>[0.008]</td>
</tr>
<tr>
<td>Average benchmark sati. wave 1</td>
<td>71.4</td>
<td>71.4</td>
<td>70.02</td>
<td>70.02</td>
<td>66.19</td>
<td>66.19</td>
</tr>
<tr>
<td>observations</td>
<td>332</td>
<td>332</td>
<td>331</td>
<td>331</td>
<td>331</td>
<td>331</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.04</td>
<td>0.05</td>
<td>0.03</td>
<td>0.06</td>
<td>0.81</td>
<td>0.02</td>
</tr>
<tr>
<td>Individual Characteristics</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
</tr>
</tbody>
</table>

Notes: The regression estimates in the table are for the specification $S_{i,t} = 13 \gamma + \gamma I_{Imagine_{i,t=3}} + X' \gamma + \epsilon_{i,t=1,3}$, where $S_{i,t=1,3}$ is the actual satisfaction of student $i$ in wave $t = 1$ (Imagine$_{i,t=3} = 0$) or imagined satisfaction of student $i$ in wave $t = 3$ (Imagine$_{i,t=3} = 1$). Wave 1 and 2 are the pre-lockdown cohort: the 2019/20 cohort of first year students who completed their first two terms of university pre-lockdown. Waves 3 and 4 are the full-lockdown cohort: the 2020/21 cohort of first year students who completed their first two terms of university during lockdown. Waves 5 and 6 are the partial-lockdown cohort: the 2021/22 cohort of first year students who completed their first two terms of university during lockdown. For the pre-lockdown cohort we survey the students in two terms, which are 10 weeks in length. Term 1 runs from early October 2019 to mid December 2019. Term 2 runs from mid January 2020 to late March 2020. For the full-lockdown cohort we survey the students in two terms which are also 10 weeks in length. Term 1 runs from early October 2020 to mid December 2020. Term 2 runs from mid January 2021 to late March 2021. For the partial-lockdown cohort we survey the students in two terms which are also 10 weeks in length. Term 1 runs from early October 2021 to mid December 2021. Term 2 runs from mid January 2022 to late March 2022.
Appendix B: Survey

The figures below are screen captures taken from the survey which give an indication of the precise wording and ordering seen by subjects. The survey includes a number of questions relating to specific elements of the undergraduate programme which are not referenced in the main text but are included to give a full idea of the scale and scope of the survey (recalling that the survey was initiated prior to the COVID-19 pandemic to serve as an ongoing source of longitudinal data about student wellbeing). The key randomization in the survey occurs from the section called ‘First year Welfare’ in Figure A6. Roughly 50% of undergraduates are given the questions from this section (which concludes in Figure A7), while the remainder instead are given the alternative ‘First year Welfare: no-covid’ questions listed in Figures A7 and A8. The randomisation gave two key benefits. First, it gave us a valid comparison between students’ current satisfaction and imagined non-COVID-19 satisfaction. Second, it kept survey completion time down as students were not required to answer both sets of satisfaction questions.
SURVEY INSTRUCTIONS

Introduction: Please answer the following questions on your phone, tablet or computer. It should take approximately 8-10 minutes to complete. The survey is designed to understand and ultimately improve students’ experience at university.

Confidentiality: All data will be anonymised and will be used to develop departmental policies and future research. Any statistic created from the information you share will be reported at the aggregate level so no individual can be identified. This was explained in the participation information leaflet.

Risk and Benefits: There are no risks associated with your participation in this study.

Participant’s rights: Your participation is voluntary and you have the right to change your mind and end the survey at any point. You can also choose at the outset not to participate in the survey. The information you provide in here will be managed and analysed uniquely by the research team and will not be shared with any third party.

Contact: If you have any questions please email economics.wellbeing@warwick.ac.uk.

Do you agree to participate in this study?

☐ Yes
☐ No

Gender, L100, home/overseas

I am:
☐ male
☐ female
☐ prefer not to say

I am a core economics (L100) student:
☐ yes
☐ no

Figure A1: (a) page 1 (b) page 2

Perceived education production function: part 1

Think about how you spend time in a typical week during term. Suppose that after essential activities (e.g. sleeping, shopping and eating) you have 60 hours left. In those 60 hours you need to fit non-academic activities (e.g. leisure, sport and job searching activities) and academic activities split into: i) lectures/seminars (online or face-to-face), ii) study with asynchronous videos, and iii) study without asynchronous videos. In a typical week during term, how many hours do you plan to spend on each activity? [Total needs to sum to 60.] (a) (b)

Figure A2: (a) page 3 (b) page 4
how you would perform. [End of year mark is your average across all your year 1 modules.]

Per week non-academic activities
- lectures/seminars
- study with videos
- study without videos

With your planned time allocation you expect to get $\{q[Q0274]/ChoiceNumericEntryValue/s\}$ how tell us your expected end of year mark in three different scenarios. In each scenario you increase a particular academic activity by 1 hour, assuming time in non-academic activities drops by 1 hour and time in other academic activities remains fixed (so total hours remains fixed at 60). [End of year mark is your average across all your year 1 modules.]

Perceived education production function: part 2 (3 hours)

Based on the weekly hours you just stated (summarised below) what is your expected end of year mark? There is no right or wrong answer, just try your best to imagine how you would perform. [End of year mark is your average across all your year 1 modules.]
With your planned time allocation you expect to get $s_{q:0274}/ChoiceNumericEntryValue()$. Now let us your expected end of year mark in three different scenarios. In each scenario you increase a particular academic activity by 5 hours, assuming time in non-academic activities drops by 5 hours and time in other academic activities remains fixed (so total hours remain at 60). (End of year mark is your average across all your year 1 modules.)

Perceived education production function: part 3

Suppose that you take a "current knowledge test" which covers all content from your core modules up to now at Warwick. 100 randomly chosen first-year students (who are taking the same core modules as you) take the same test. What do you think your relative performance would be? (0 lowest performer, 100 highest performer.)

Welfare of friends: part 1

Think about the 10 students you know best in your cohort at Warwick. In your opinion how many of those 10 students are experiencing conditions ranging from mild anxiety to severe depression. Importantly, this is your opinion and can include medically or non-medically diagnosed conditions.

Welfare of friends: part 2

of the $s_{q:0302}/ChoiceNumericEntryValue()$ students you thought may be experiencing conditions ranging from mild anxiety to severe depression, what percentage do you think are likely to report these conditions to the university (wellbeing services). (If you put 0 in the previous question, please leave the percentage blank.)

First year Welfare

In general between the start of the academic year and now, I have been ______ with my university life.

In general between the start of the academic year and now, I have been ______ with my physical health.

In general between the start of the academic year and now, I have been ______ with my social life.
In general, between the start of the academic year and now, I have been ........ with my mental health.

In general, between the start of the academic year and now, I have been ........ with my financial situation.

In general, between the start of the academic year and now, I have been ........ with the amount of sleep I had.

First year Welfare: no-covid

Imagine the whole covid situation DID NOT exist. In general, between the start of the academic year and now, I think I would have been ........ with my accommodation situation.

Imagine the whole covid situation DID NOT exist. In general, between the start of the academic year and now, I think I would have been ........ with my mental health.

Imagine the whole covid situation DID NOT exist. In general, between the start of the academic year and now, I think I would have been ........ with my social life.

Imagine the whole covid situation DID NOT exist. In general, between the start of the academic year and now, I think I would have been ........ with my university life.
Imagine the whole covid situation DID NOT exist. In general, between the start of the academic year and now, I think I would have been ....... with the amount of sleep I had.

Imagine the whole covid situation DID NOT exist. In general, between the start of the academic year and now, I think I would have been ....... with the academic support I received from the department.

Imagine the whole covid situation DID NOT exist. In general, between the start of the academic year and now, I think I would have been ....... with the pastoral support I received from the department.

(a) EC109 academic performance and effort
Now consider your performance and effort in EC109. What is the mark you expect to get on EC109 test 1?

How many hours of EC109 work per week, on average during term 1, do you think you NEED TO put into EC109 to achieve your expectation? DO NOT include the 2 hours of lectures/seminars in your answer.

How many hours of EC109 work per week, on average during term 1, do you think you WILL put into EC109 to

(b) Subjective wages and employment
Consider the following hypothetical situation. You graduate with a .... class degree and decide NOT to go into further education, but to apply for jobs. Think about the kind of jobs that are available to you. What do you think is the probability you will be employed in a job one year from the date of graduation. The alternative to being employed is being unemployed.

On a scale of 1-100, for an average hour of EC109 work, how focused do you think you NEED TO be to achieve your expectation? 0 is your lowest focus possible. 100 is your highest.

On a scale of 1-100, for an average hour of EC109 work, how focused do you think you WILL be to achieve your expectation? 0 is your lowest focus possible. 100 is your highest.

(a) (b)

Figure A9: (a) page 17 (b) page 18

Figure A10: (a) page 19 (b) page 20
Figure A11: (a) page 21