People have been talking about this PNAS paper by Matthew Killingsworth: "Experienced well-being rises with income, even above $75,000 per year". Here's the abstract:

Past research has found that experienced well-being does not increase above incomes of $75,000/y. This finding has been the focus of substantial attention from researchers and the general public, yet is based on a dataset with a measure of experienced well-being that may or may not be indicative of actual emotional experience (retrospective, dichotomous reports). Here, over one million real-time reports of experienced well-being from a large US sample show evidence that experienced well-being rises linearly with log income, with an equally steep slope above $80,000 as below it. This suggests that higher incomes may still have potential to improve people's day-to-day well-being, rather than having already reached a plateau for many people in wealthy countries.

And here's Figure 1 from the paper, that makes the case:
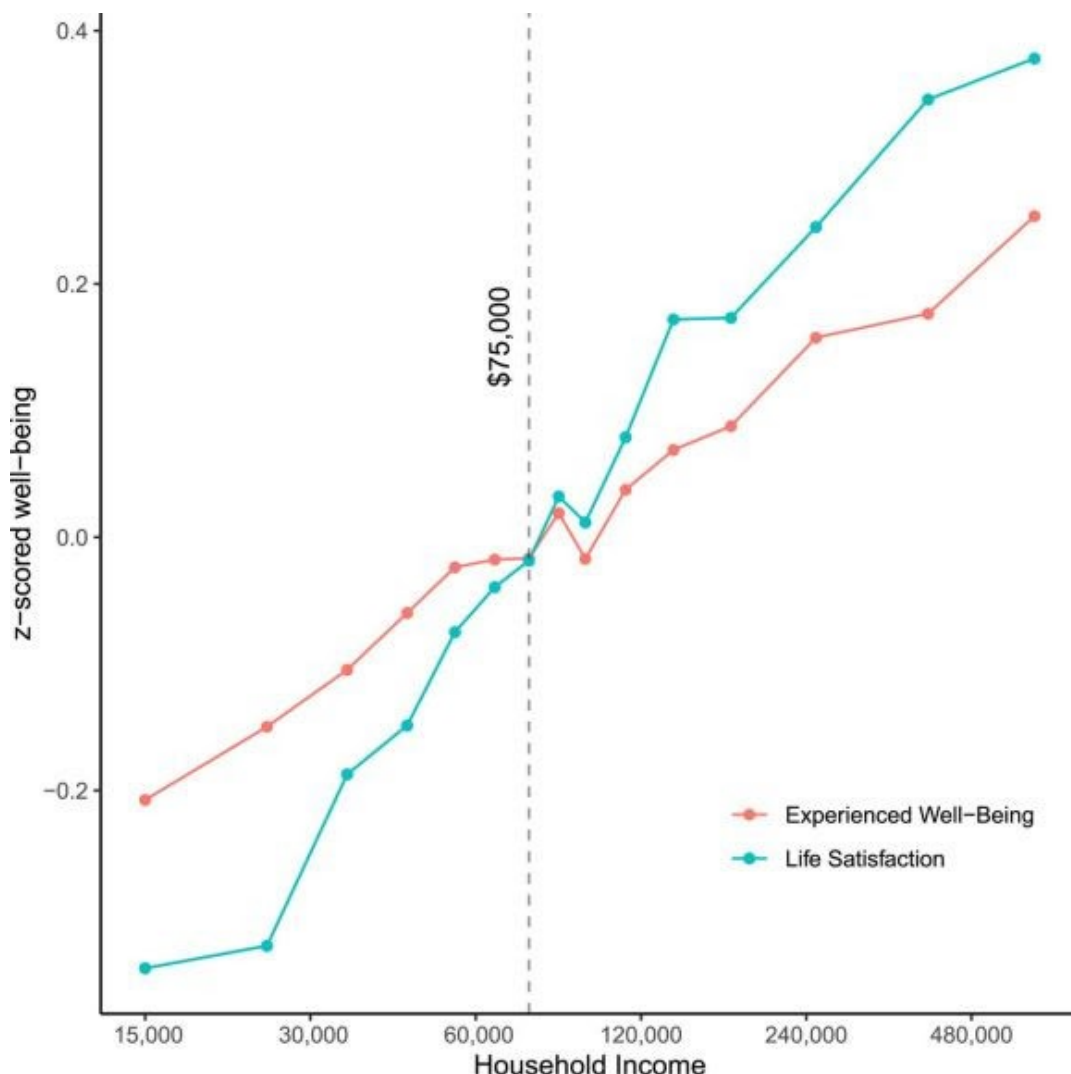
Killingsworth Figure 1

On Twitter I noted—in the spirit of a seminar question—that the figure, consistent with the argument of the paper, puts the x-axis income groups on a log scale. The y-axis is in standardized z-score units for each of the two measures of happiness. There's nothing wrong with that per se, of course, and it often makes perfect sense to log-transform something like income, and to graph change on a log scale. (Don't @-me, inverse hyperbolic sine people. I know.) In the context of the paper, the main target is the idea that there's a ceiling above which more money doesn't make you happier—in this case, from previous research, that is taken to be $75,000 of income per year. There is of course no particular reason why this number should be a specific and absolute dollar figure—or at least, that's something that one would have to argue for. But

let's set that issue aside as it's not relevant here.

If you were an adherent of the ceiling view, you might reasonably say, look, even if the effect of income on happiness is linear in the log of income, that's basically the same as saying it's *not* linear in income, and that above some threshold or ceiling you'd need to increase your income by a *lot* in order to see any substantial increase in happiness. My colleague Steve Vaisey noted the same thing:

> Yes, money does buy happiness. However — to get that next sweet 1/10 standard deviation bump in well-being requires you to double your household income. https://t.co/8hfOr1QaTF
>
> — Stephen Vaisey (@vaiseys) January 25, 2021

We can see this if we re-draw the figure with income on a linear scale, which is what I did. The summary data for the paper is available on OSF at https://osf.io/cfnbv/. With the CSV of the data, we can do this:

```
library(tidyverse)

happy <- read_csv("~/Downloads/income_wellbeing.csv")

happy %>%
  select(household_income,
         experienced_wellbeing_zscore.Mean,
         life_satisfaction_zscore.Mean) %>%
  rename(income = household_income,
         `Experienced Wellbeing` = experienced_wellbeing_zscore.Mean,
         `Life Satisfaction` = life_satisfaction_zscore.Mean) %>%
  pivot_longer(`Experienced Wellbeing`:`Life Satisfaction`,
               names_to = "measure",
               values_to = "value") %>%
  ggplot(aes(x = income, y = value, color = measure)) +
  geom_vline(xintercept = 75000, color = "gray20") +
  geom_point() +
  geom_line() +
  annotate("text", x = 6e4, y = 0.2, label = "$75,000", angle = 90) +
  scale_x_continuous(labels = scales::label_dollar()) +
  labs(x = "Income bracket", y = "z-score", color = "Measure:") +
  theme_bw() +
  theme(legend.position = "top")
```

And get a figure that looks like this:
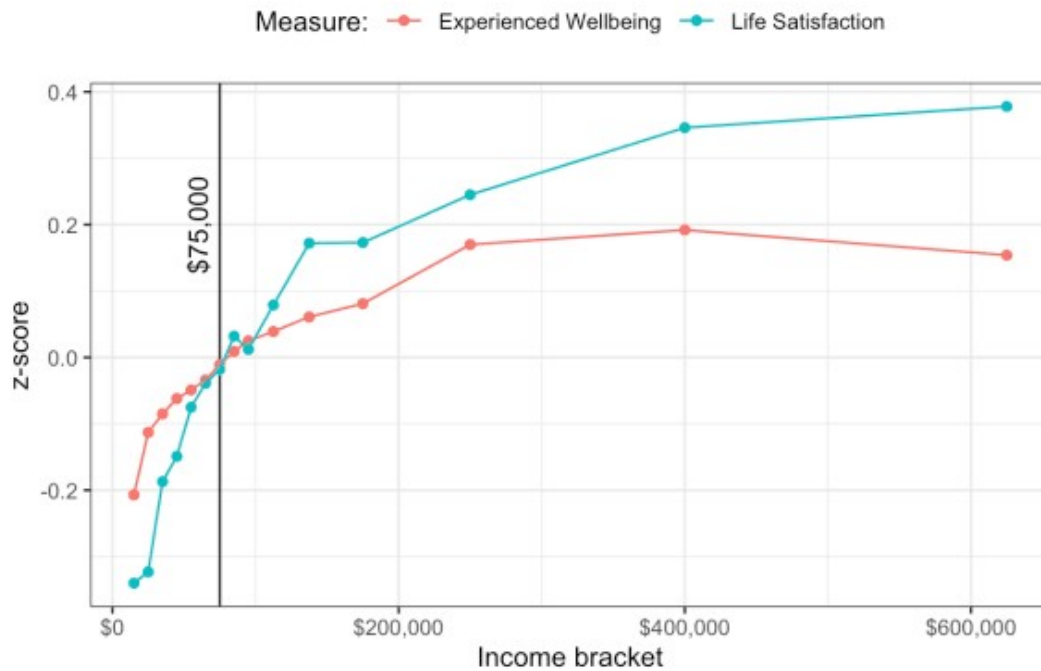
Killingsworth Figure 1, on a linear x-axis

So you can see why an advocate of the threshold or ceiling view of income satisfaction might be unconvinced that the log-linearity of happiness in income is much to be concerned about. Sure, it's still growing, but after the initial steep increase in happiness that comes with getting some money, across most of the range of achievable incomes the increase is negligible. I don't have a dog in this fight either way.

This figure raises a further issue, though. If you look at the two lines you can see that while the green (life satisfaction) measure is the same as in the paper's figure, the red (Experienced Wellbeing) one is different. In particular it does actually dip downwards for the highest observed income class, which might be relevant for the thesis of the paper. There are very few observations up that high in the income distribution, so everything's very noisy. Still, if we redraw the original figure based on the data supplied in the replication package we can see they're not the same series:

```
1
2 happy %>%
3   select(household_income,
4          experienced_wellbeing_zscore.Mean,
5          life_satisfaction_zscore.Mean) %>%
6   rename(income = household_income,
7          `Experienced Wellbeing` = experienced_wellbeing_zscore.Mean,
8          `Life Satisfaction` = life_satisfaction_zscore.Mean) %>%
9   pivot_longer(`Experienced Wellbeing`:`Life Satisfaction`,
10              names_to = "measure",
11              values_to = "value") %>%
12   ggplot(aes(x = income, y = value, color = measure)) +
13   geom_vline(xintercept = 75000, color = "gray20") +
14   geom_point() +
15   geom_line() +
16   annotate("text", x = 6.9e4, y = 0.2, label = "$75,000", angle = 90) +
17   scale_x_log10(labels = scales::label_dollar()) +
18   labs(x = "Income bracket", y = "z-score", color = "Measure:",
19        title = "Logged x-axis") +
20   theme_bw() +
21   theme(legend.position = "top")
```

Logged x-axis

Measure: ● Experienced Wellbeing   ● Life Satisfaction
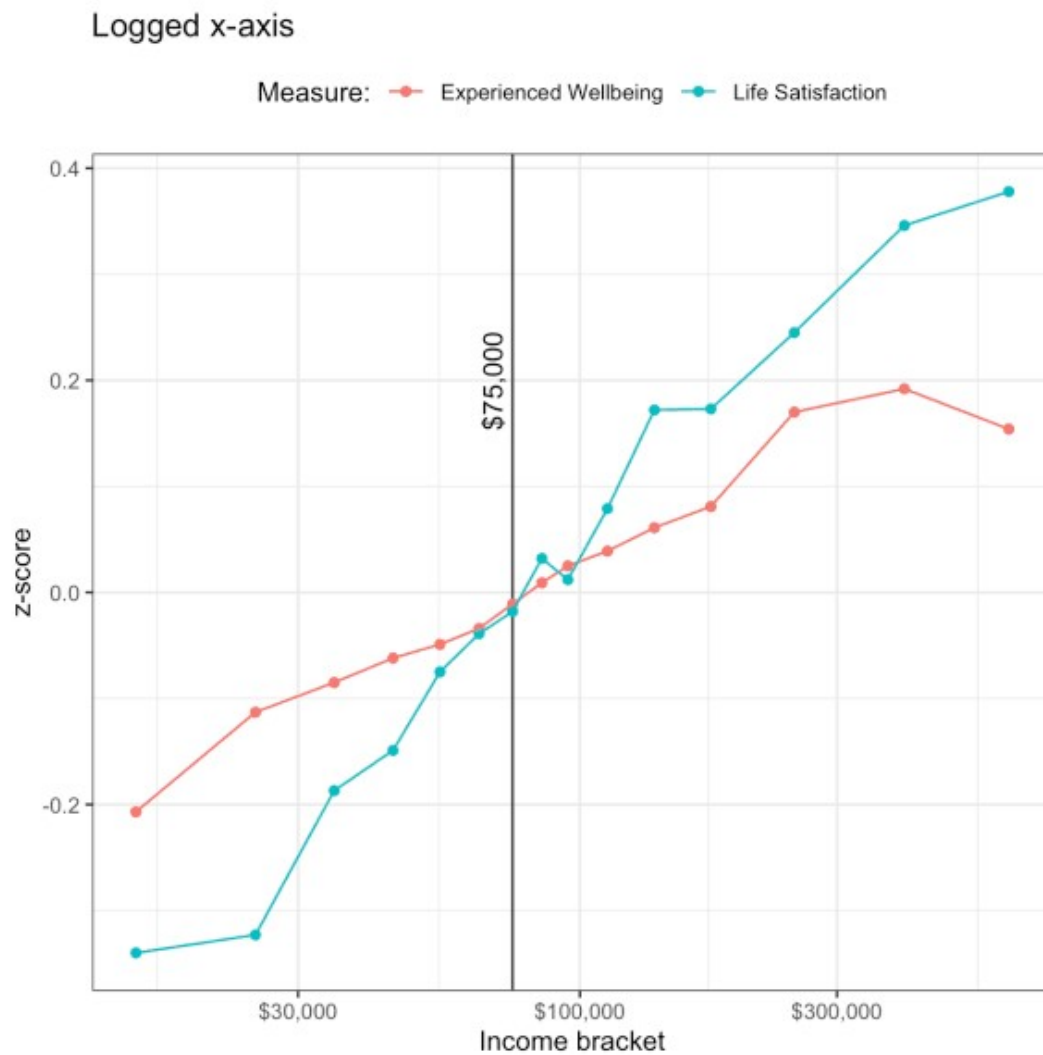
Killingsworth Figure 1 redrawn. Notice amongst other things the dip in the last value of Experienced Wellbeing.

At first I thought I'd chosen the wrong variable to draw, but looking again at the data I can't see a plausible alternative.

```
 1 happy %>%
 2   select(contains("experienced")) %>%
 3   colnames()
 4
 5 ## 1] "experienced_wellbeing.Mean"              "experienced_wellbeing_zscore.
 6 Mean"
 7 ## [3] "experienced_wellbeing.std.error"        "experienced_wellbeing_zscore.
 8 std.error"
 9 ## [5] "experienced_wellbeing.personcount"      "experienced_wellbeing_zscore.
10 personcount"
```

I could still be missing something, of course, and would welcome any corrections. Another possibility is

that the data uploaded to the OSF replication package isn't quite the one used in the paper. This is the sort of thing that happens quite often.

A third possibility is suggested by a remark in the caption to Figure 1:

> Mean levels of experienced well-being (real-time feeling reports on a good–bad continuum) and evaluative well-being (overall life satisfaction) for each income band. Income axis is log transformed. Figure includes only data from people who completed both measures.

"Figure includes only data from people who completed both measures" could explain the discrepancy. The z-score means in the replication package are, presumably, calculated from all the observations for each measure. But if the figure is showing a subset of the two (i.e. only observations from people who answered both questions) then the z-score means across income levels will be slightly different, depending on who is excluded. I don't quite see the rationale for subsetting the data in this way for the purposes of the drawing Figure 1, though. It seems that you'd want to just use all the available observations for each measure, given that there aren't any constraints on trying to balance the two groups or anything like that. Including it would introduce a non-linearity in the log-income relationship, though. That might well be just measurement error, given the vagaries of income reporting and small-n noisiness at very high incomes, but it would directly cut against the main claim of the paper.