METHODS

Supplement Table 1 Divisions of households by socio-demographic characteristics.

Group	Categories	Distribution within each group
Age of the main shopper	18-34	17.1%
	35-44	24.1%
	45-54	25.3%
	55-64	16.3%
	65+	17.3%
Total amount of grams of all alcohol purchased per day, averaged over the number of days between the first and the last purchase of alcohol for each household, adjusted for the number of adults in the household	≤1g	21.4%
	>1g to ≤2g	17.2%
	>2g to ≤5g	26.9%
	>5g to ≤10g	17.4%
	>10g	17.1%
Social class (from highest to lowest social class), based on National Readership Survey (2019)	AB	21.1%
	C1	39·1%
	C2	18·2%
	D	13·8%
	E	7·8%
Income per adult per household per year	£0-7·5k	22·1%
	>£7·5-12·5k	20.6%
	>£12·5-17·5k	22·1%
	>£17·5 to 25k	16.7%
	>£25k	18.6%

Generalized linear model to estimate volume of purchases

At the level of the household, we also used a generalised linear model to:

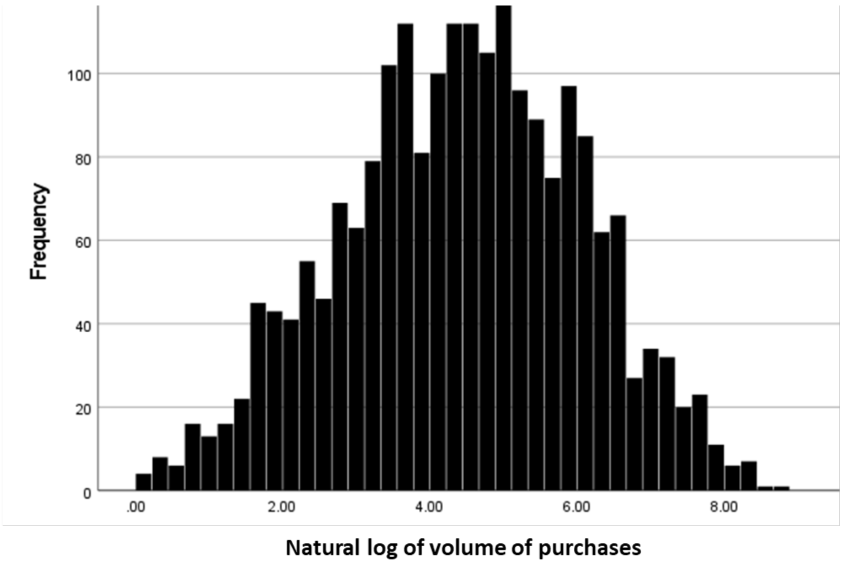
a. Estimate if there was any difference in the volume (ml) of subsequent purchases of the parent beers between parent-minus/nablab-plus and parent-minus/nablab-minus to answer question 1 further, to what extent does the introduction of nablab beers act as a gateway by increasing the volume of subsequent purchases of same branded higher strength beers?

b. Estimate if parent-plus purchased a higher volume (ml) of the new nablab product compared with parent-minus to answer question 2 further, to what extent does the previous purchase of higher strength beers affect the volume of purchases of newly introduced same branded nablab beers?

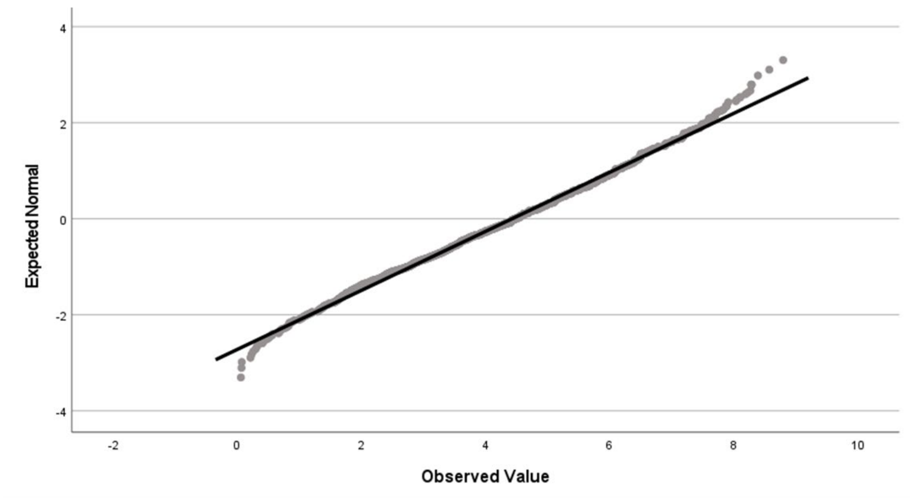
For these analyses, we calculated, at the level of the household, the volume of purchases of the nablab and parent products per adult per day of purchase, grouped by before or after the introduction of nablab, the event as described above. We examined the distributions of the volume of purchases of both nablab beer and parent beer after the introduction of nablab, and found them, as expected, to be dispersed. We took the natural log of the volumes, which resulted in a normal distribution, Supplement, Figures 1-4, pages 3-4.

The regression equation is:

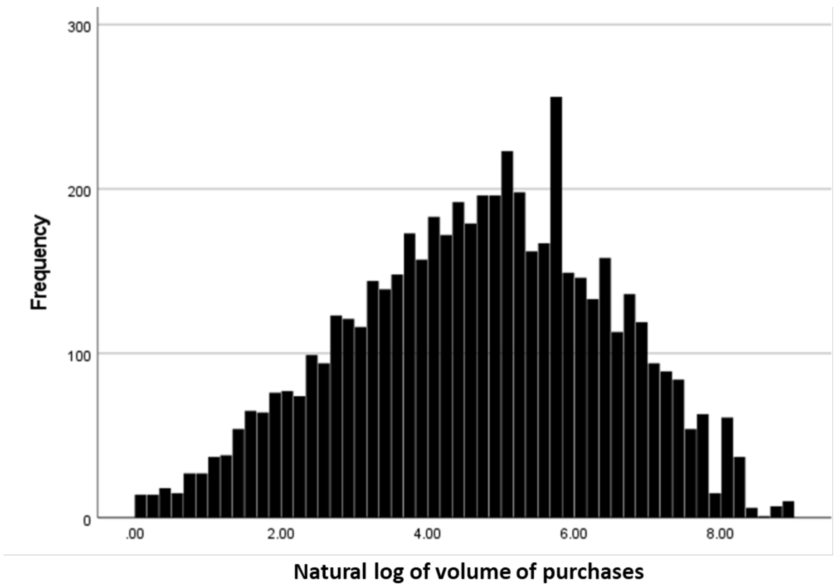
- Log natural (volume of purchase) = intercept + predictor (yes/no) + error.



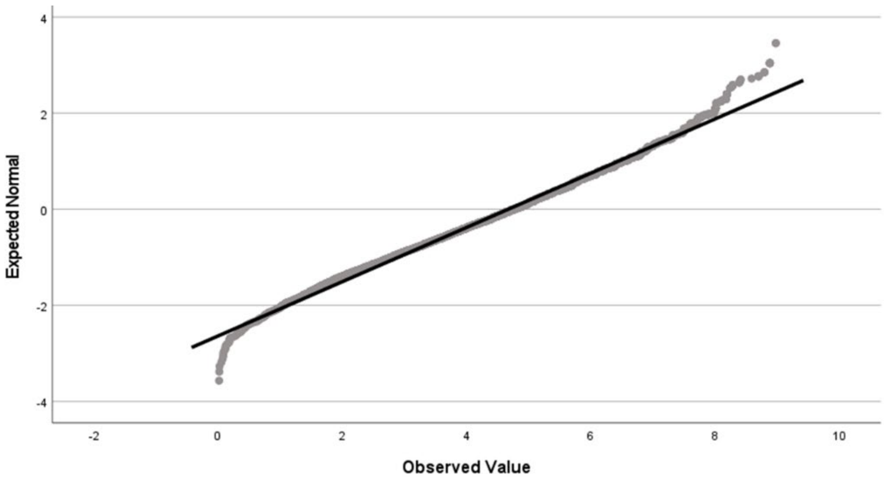
Supplement Figure 1 Distribution of the frequency of the natural log of average volume (ml) of purchases of the nablab products across all households, adjusted for the number of adults in each household, across all days of the study period since the introduction of nablab products



Supplement Figure 2 Normal Q-Q plot of the frequency distribution of the natural log of average volume (ml) of purchases of the nablab products across all households, adjusted for the number of adults in each household, across all days of the study period since the introduction of nablab products



Supplement Figure 3 Distribution of the frequency of the natural log of average volume (ml) of purchases of the parent products across all households, adjusted for the number of adults in each household, across all days of the study period since the introduction of nablab products



Supplement Figure 4 Normal Q-Q plot of the frequency distribution of the natural log of average volume (ml) of purchases of the parent products across all households, adjusted for the number of adults in each household, across all days of the study period since the introduction of nablab products

Generalized linear model to estimate odds ratios – two other issues

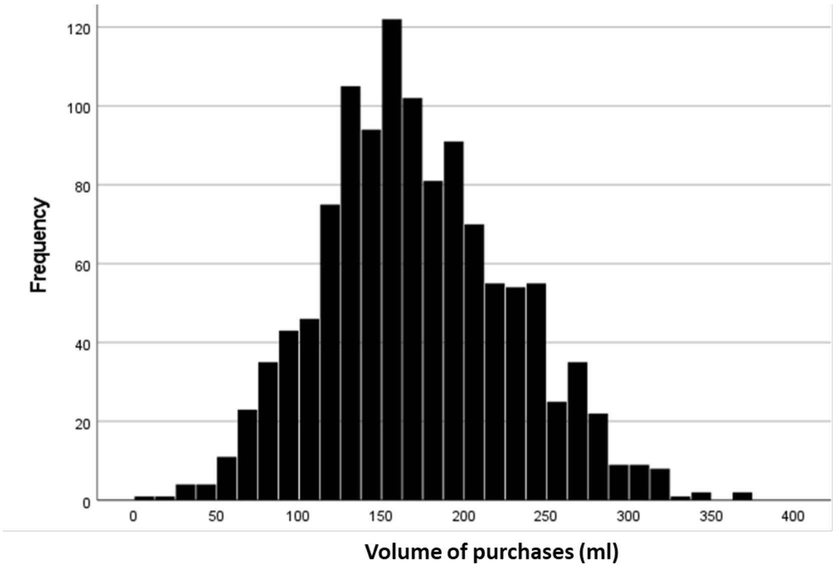
For question 2, we examined two other issues:

i. For the two nablab products (one a low alcohol product released before the other no-alcohol product) with the same parent product, we asked, did prior purchase of the low alcohol product impact on later purchases of the no-alcohol product? We tested this by repeating the above generalized linear model with the no-alcohol product as the dichotomized dependent variable (purchase yes/no) and both the parent product and the low alcohol product as the dichotomized independent predictor variables (purchase, yes/no); and

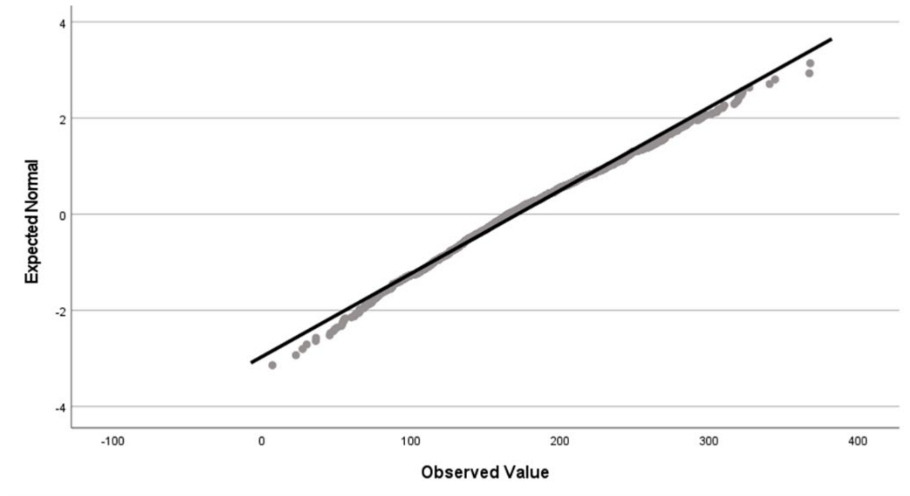
ii. Was there switching from a higher strength product to a different branded nablab product within the six nablab products that we investigated? We tested this by repeating the above generalized linear model to test the potential impact of previous purchases of each of all the parent products by adding them all to the model as independent predictor variables, dichotomized yes/no for purchasing, with models run separately for each of the dichotomized nablab products (purchase, yes/no) as dependent variables.

Interrupted time series analysis to estimate impact on purchase volume analysed at the level of the study day

The assignment of the day of purchase to the one of the study days, implied that the recalculated study day across all six nablab products included different days of the week and different seasons of the year. This largely reduced any effect of seasonality of the dependent variable, the volume of subsequent purchases of all parent products. Any residual seasonality was removed with the ratio to moving average method (Makridakis et al. 1983; McLaughlin 1984). We examined the distribution of the dependent variable and, not unexpectedly, since it was calculated at the level of the study day, found it to be normally distributed, Supplement Figures 5-6.



Supplement Figure 5 Distribution of the frequency of the average volume (ml) of purchases of the parent products per study day of all households, adjusted for the number of adults in each household.



Supplement Figure 6 Normal Q-Q plot of the distribution of the frequency of the average volume (ml) of purchases of the parent products per study day of all households, adjusted for the number of adults in each household.

We found the series of the dependent variable over time to be stationary (Augmented Dickey Fuller test, $p<0.01$) and without autoregression (Box-Llung Q statistic, $p=0.107$). We used a generalized linear model, with the regression equation:

Volume of purchase of parent product = intercept + time + event + error,

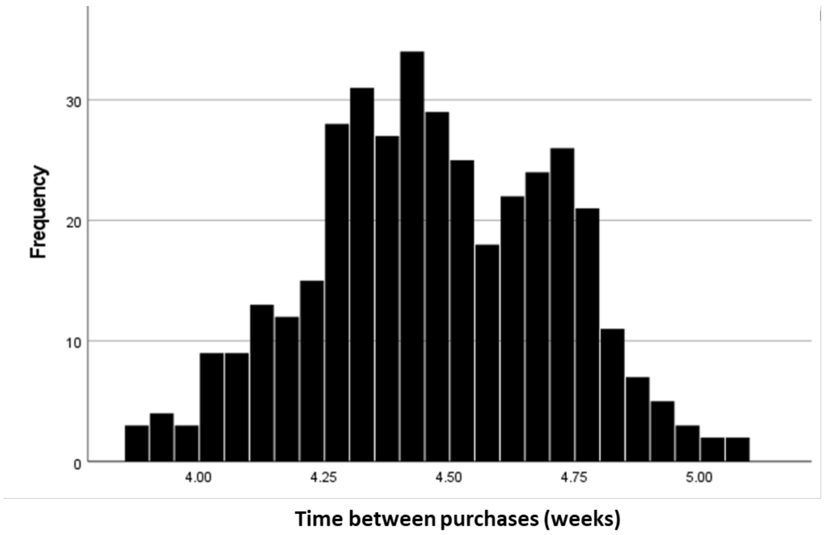
where time is the days before and after the event, with the event the dummy-coded variable for the introduction of the new nablab product.

Makridakis, S., S. C. Wheelwright, and V. E. McGee. 1983. *Forecasting: Methods and applications*. New York: John Wiley and Sons.

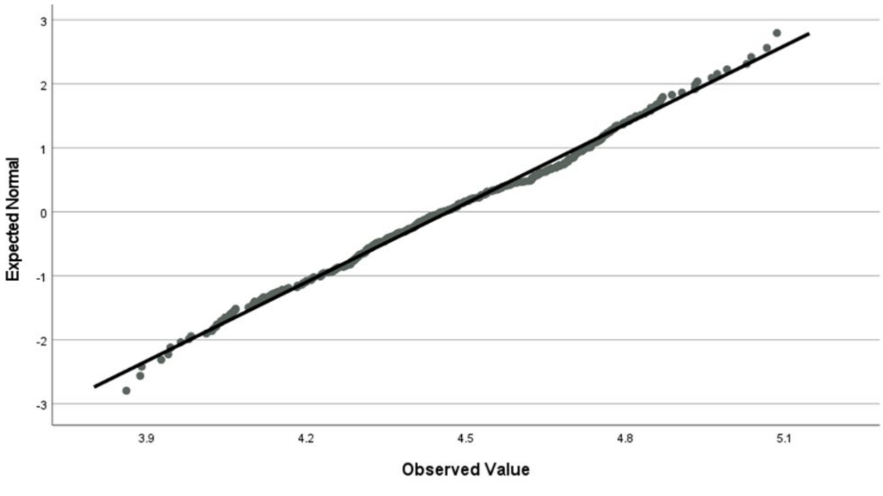
McLaughlin, R. L. 1984. *Forecasting techniques for decision making*. Rockville, Md.: Control Data Management Institute.

Frequency of purchases

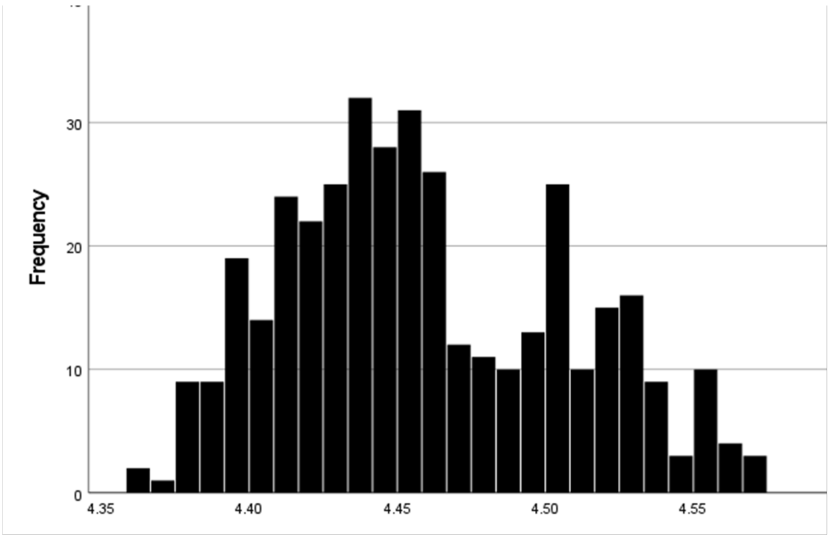
We checked for the stability of frequency of purchases over time and for any differences in frequency between parent products and new nablab products by: first, at household level, calculating the number of days between a purchase of a parent product and the next purchase of a parent product (and the same for the new nablab products); and, second, then averaging the number of days of difference across all households by study day, converting the difference, which we call frequency of purchase, to weeks, by dividing by 7. As with the volume of purchase, the recalculated study day largely reduced any effect of seasonality. Any residual seasonality was removed with the ratio to moving average method (Makridakis et al. 1983; McLaughlin 1984). Examination of the distribution of the dependent variables, frequency of purchase found them to be normally distributed, Supplement Figures 7-10. We used linear regression to estimate the coefficient of regression (with 95% confidence intervals) of the frequency of purchase with study day. We used non-paired t-tests to estimate the differences in frequency of purchase between parent products before and after the event (introduction of first nablab) and paired t-tests to estimate the differences in frequency of purchase between parent products and new nablab products after the event, reporting means and the differences in means with respective 95% confidence intervals.



Supplement Figure 7 Distribution of the frequency of the average time between purchases (weeks) of the parent products per study day of all households.

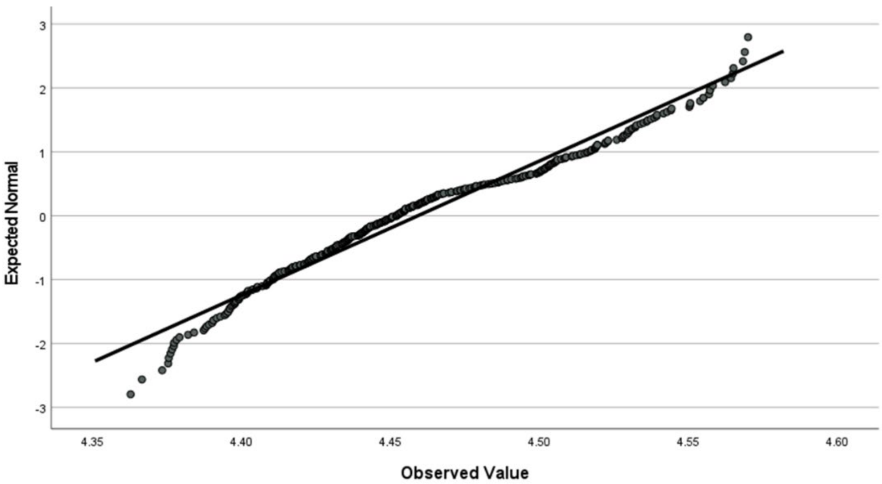


Supplement Figure 8 Normal Q-Q plot of Distribution of the frequency of the average time between purchases (weeks) of the parent products per study day of all households.



Time between purchases (weeks)

Supplement Figure 9 Distribution of the frequency of the average time between purchases (weeks) of the nablab products per study day of all households.



Supplement Figure 10 Normal Q-Q plot of Distribution of the frequency of the average time between purchases (weeks) of the parent products per study day of all households.

Makridakis, S., S. C. Wheelwright, and V. E. McGee. 1983. *Forecasting: Methods and applications*. New York: John Wiley and Sons.

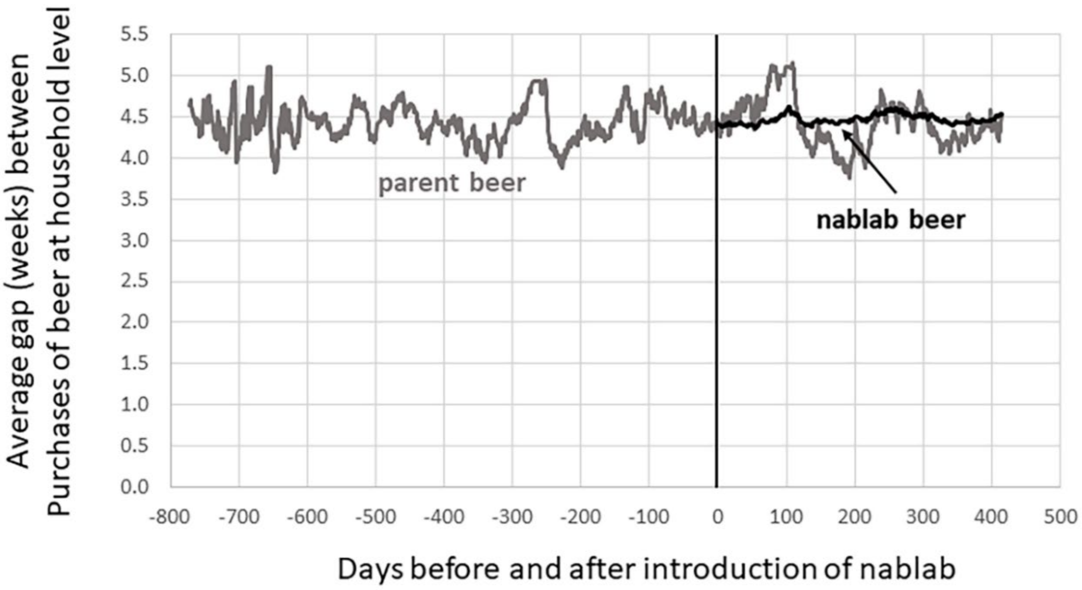
McLaughlin, R. L. 1984. *Forecasting techniques for decision making*. Rockville, Md.: Control Data Management Institute.

Stability of changes over time

To consider stability of changes at the level of the individual household, we analyzed parent-plus/nablab-plus households. For each of the six nablab beers, and at the level of each household, we first set day 1 as the day of first purchase of a new nablab beer; we then created a new dependent variable, a ratio, calculated as (volume of purchases of new nablab beer) divided by (volume of purchases of new nablab beer plus volume of purchases of parent beer) for each day from day 1 onwards. For each new nablab beer, we then calculated the mean of these ratios across all households from day 1 onwards. We then averaged these means across the six nablab beers, plotted the results over time and undertook a regression analysis with time (day) as the independent variable.

RESULTS

Supplement Figure 11 plots the frequency of purchases over time. Overall, across both parent and nablab products, the mean frequency, i.e., the time between purchases per household, was 4.43 weeks (95% CI=4.42 to 4.44). We found that the frequency of purchases of the sum of the parent products did not differ over time (regression coefficient for frequency with study day=-6.3^{-5} (95% CI=-13.5^{-5} to 1.1^{-5}), with no differences in frequency before the event (frequency = 4.41 weeks, 95% confidence intervals, CI=4.39 to 4.43) and after the event (frequency = 4.44 weeks, 95% confidence intervals, CI=4.42 to 4.46), difference in frequencies = 0.02, (95% CI= -0.006 to 0.054). We found that the frequency of purchases of the nablab products did not differ by time since the event, (regression coefficient for frequency with study day =1.7^{-4} (95% CI=-1.5^{-4} to 4.9^{-4}) with, after the event, no differences in frequency of purchases between parent products (frequency = 4.44 weeks, 95% confidence intervals, CI=4.42 to 0.46) and nablab products (frequency = 4.45 weeks, 95% confidence intervals, CI=4.44 to 4.46), difference in frequencies = 0.01, (95% CI= -0.01 to 0.03).



Supplement Figure 11 Average gap (weeks) between purchases of parent and nablab beers at household level by study days before and after introduction of nablab (vertical line).

Supplement Table 2 Results of models, one for each new nablab beer as dependent variable, testing the potential impact of previous purchases of each of the parent products (independent predictor variables) for purchasing, separately, each of the nablab products. Dependent and independent variables entered as dichotomous variables, present (at least one purchase) or absent (no purchase). Odds ratios estimate likelihood of at least one purchase of new no and now alcohol beer comparing presence of at least one purchase of a parent product to absence.

Dependent variable (nablab product number)	Independent variables (parent product number, equivalent to nablab beer product number – see Table footnote). Bold: same-branded parent product.	Odds ratios (95% confidence intervals). Bold: confidence intervals do not cross zero.
1	**1**	**2.526 (2.172 to 2.937)**
	3	1.236 (.646 to 2.366)
	4	**2.310 (1.778 to 3.001)**
	5	.788 (.368 to 1.687)
	6	**2.065 (1.271 to 3.354)**
2	**2**	**1.679 (1.288 to 2.190)**
	3	.953 (.234 to 3.882)
	4	.833 (.410 to 1.691)
	5	1.188 (.378 to 3.731)
	6	.750 (.184 to 3.058)
3	2	1.073 (.660 to 1.746)
	3	**8.203 (3.756 to 17.919)**
	4	1.434 (.618 to 3.326)
	5	2.030 (.618 to 6.671)
	6	**2.977 (1.122 to 7.899)**
4	2	1.132 (.869 to 1.476)
	3	1.001 (.316 to 3.170)
	4	**2.645 (1.785 to 3.918)**
	5	.865 (.275 to 2.719)
	6	1.046 (.383 to 2.856)
5	2	.861 (.503 to 1.474)
	3	.852 (.114 to 6.385)
	4	1.115 (.406 to 3.067)
	5	**4.973 (1.995 to 12.398)**
	6	2.788 (.848 to 9.162)
6	2	.886 (.456 to 1.721)
	3	1.444 (.191 to 10.909)
	4	2.332 (.923 to 5.892)
	5	.000 (.000 to .)
	6	**4.728 (1.424 to 15.694)**

Products 1 and 2 had the same parent product. Thus: for product 1 as low-alcohol dependent variable, parent product 2 was not included as an independent variable; for product 2 as no-alcohol dependent variable, parent product 1 was not included as an independent variable; for no-alcohol dependent variables 3 to 6, no-alcohol parent product 2 was included as an independent variable.