

Data Description

Methodology

- We use OLS estimation to analyze the effect of being a superhost compared to non-accredited hosts on the listing's price. In our regression equation, the DV the "price" and IV is the "superhost dummy". We also include covariates such as the number of reviews, house size and etc.
- To overcome endogeneity problem, we use sample matching method. To be more specific, we match the treated (superhosts) and untreated (non-accredited hosts) groups based on the covariates included in the model like a real lab experiment, both conditions are approximately the same except the difference in the type of host (superhost vs non-accredited).
- We match both groups with optimal matching with a ratio of 1. The main goal of this approach is to find the matched samples with the smallest average absolute propensity score distance across all the matched pairs.

See the distribution of propensity scores before and after the matching:

Results

- When we use simple OLS, the coefficient of superhost is -7.1. However, we find no statistical evidence to support our hypothesis.
- OLS after propensity score matching yields to a coefficient of -24.99 ($t = 4.24$, $p < 0,001$). Unlike our expectations, we find that on average a superhost charges 25 Euros less than its non-accredited counterpart for the same house. Our hypothesis at the beginning was superhosts are charging higher prices than ordinary hosts, however analysis revealed that the direction of the relationship is the opposite. This might be because, superhosts feel themselves under pressure and set their prices lower so that they have higher reviews and continue as a superhost.
- Interestingly, houses in the center have a lower price than houses located outside the center.