Appendix 2 - Illustrative example

In this section, the discrimination-free prices are calculated to a simulated home insurance portfolio to illustrate the presented methodology. The variables in the simulated data are presented in Table 2.

Table 2: Summary of variables in the simulated data

Variable	Explanation	Type
Area	Size (m^2) of the apartment	Non-discriminatory
Housing type	Indicates whether the apartment is rented or owned	Non-discriminatory
Gender	Gender of the policyholder	Discriminatory

It is further assumed that the gender variable is correlated with the housing type variable[2].

In addition, the following assumptions – which are purely illustrative and might not reflect real-world phenomena – are made when simulating the claims data.

- The greater the size (m2) of the apartment, the greater the risk
- In rented apartments the risk is greater than in owned apartments
- Women possess greater risk than men.

Calculation of the discrimination-free prices

In figure 2 is presented the best-estimate prices for women and men, unawareness prices and discrimination-free prices as described in Table 1.

The best estimate prices and the unawareness prices are obtained using standard GLM methodology[3].

The discrimination-free price is derived from the best-estimate prices using the empirical marginal distribution of the gender with bias correction[4].

---

[2] In generating the data for the example, it is assumed that that $\mathbb{P}(woman \mid owned) = 0.4$ and $\mathbb{P}(woman \mid rented) = 0.85$

[3] Response variable (claims costs) is modelled with GLM using Tweedie distribution assumption and log-link function.

[4] Introducing discrimination-free prices might lead to situations where a bias occurs in the portfolio. This means that the total sum of the discrimination-free prices is lower or higher than the sum of the best estimate prices, in which cases the portfolio is underpriced, or overpriced respectively. In such situations, the marginal distribution used in the calculation of the discrimination-free prices can be adjusted to eliminate the portfolio bias.

In our example, the discrimination-free price for a policy *i* is given by:

$$\mathbb{E}(Y_i \mid X_i, man) * \frac{n_{man}}{n_{man} + n_{woman}} + \mathbb{E}(Y_i \mid X_i, woman) * \frac{n_{woman}}{n_{man} + n_{woman}}$$

ÖSTERMALMSTORG 1, 5TR
SE-114 42, STOCKHOLM, SVERIGE
FCG.GLOBAL

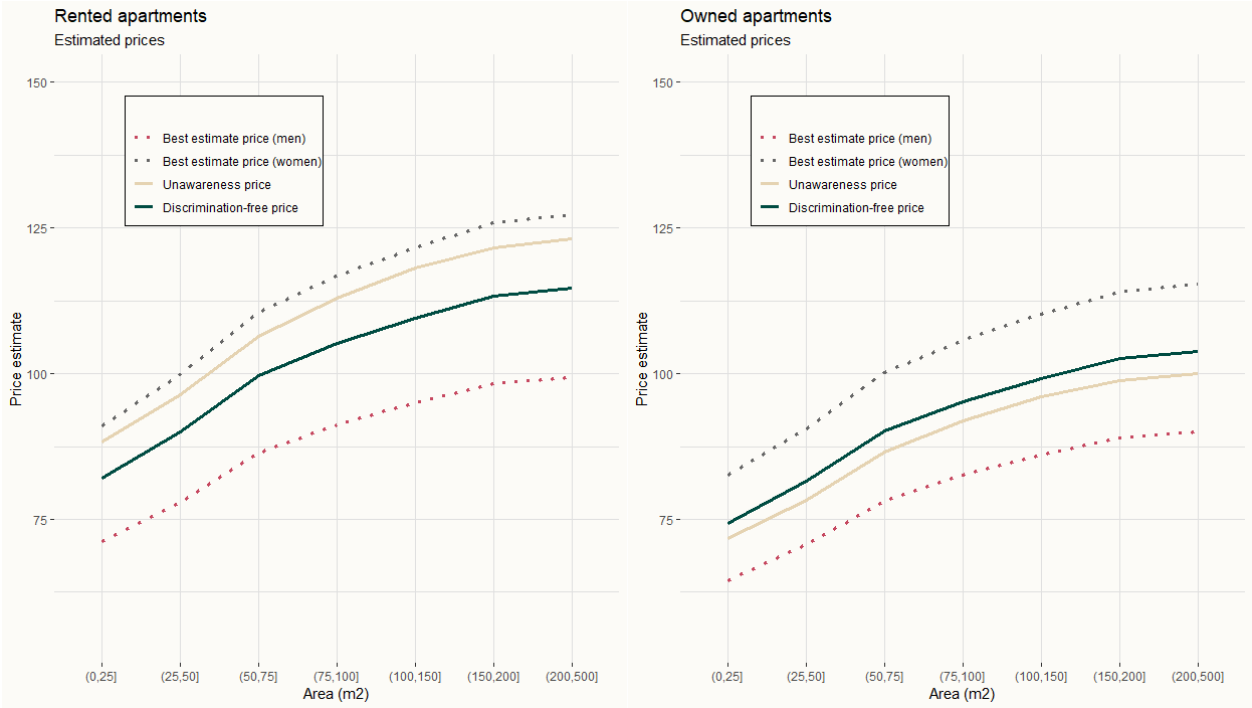

Figure 4: Prices between housing types

In Figure 4, the best estimate prices for men and women are presented with dotted lines. It is obvious that they discriminate between men and women, since gender is used as an explanatory variable, and it has explanatory power to the claims cost (see assumptions related to generating the data).

The unawareness price and the discrimination-free prices are presented with solid lines. The unawareness price lies above the discrimination-free price for the rented apartments (left) and below the discrimination-free price for the owned apartments (right). This means that the unawareness price implicitly allocates higher price to women, since policyholders with rented apartment are more likely to be women in this example.

Where n_{man} denotes the number of men in the portfolio and n_{woman} the number of women, respectively. The bias-corrected discrimination-free price for policy *i* is obtained by scaling the respective discriminatory-free price such that the sum of the bias-corrected discriminatory-free prices equal to the sum of the best estimate prices.

ÖSTERMALMSTORG 1, 5TR
SE-114 42, STOCKHOLM, SVERIGE
FCG.GLOBAL