Regression Decomposition

Expert Testimony

Predicting out of sample

Recall you can predict out of sample using the following formula:

$$\hat{Y}_{\text{test}} = X_{\text{test}} \cdot \hat{\beta}(X_{\text{train}}, Y_{\text{train}})$$

More generally, given a model estimated on data A, it can be evaluated on data B:

$$\hat{Y}_{AB} = X_B \cdot \hat{\beta}_A$$

Viennese hotel prices in Budapest?

Hotels in Budapest are cheaper than in Vienna.

Is it because of their characteristics (location, quality), or because Budapest is a cheaper city?

Hotel price model for Vienna

$$E(\text{price}|\text{distance}) = 38.14 - 9.03 \cdot \text{distance} + 26.42 \cdot \text{rating}$$

Model

AKA "coefficients," "parameters," or "weights": (38.14, -9.03, 26.42)

Data

AKA "features," "predictors," or "independent variables": (distance, rating)

Naive solution

Run a regression on pooled data and include a city dummy.

This assumes that the model is the same in both cities.

Hotel prices in two cities

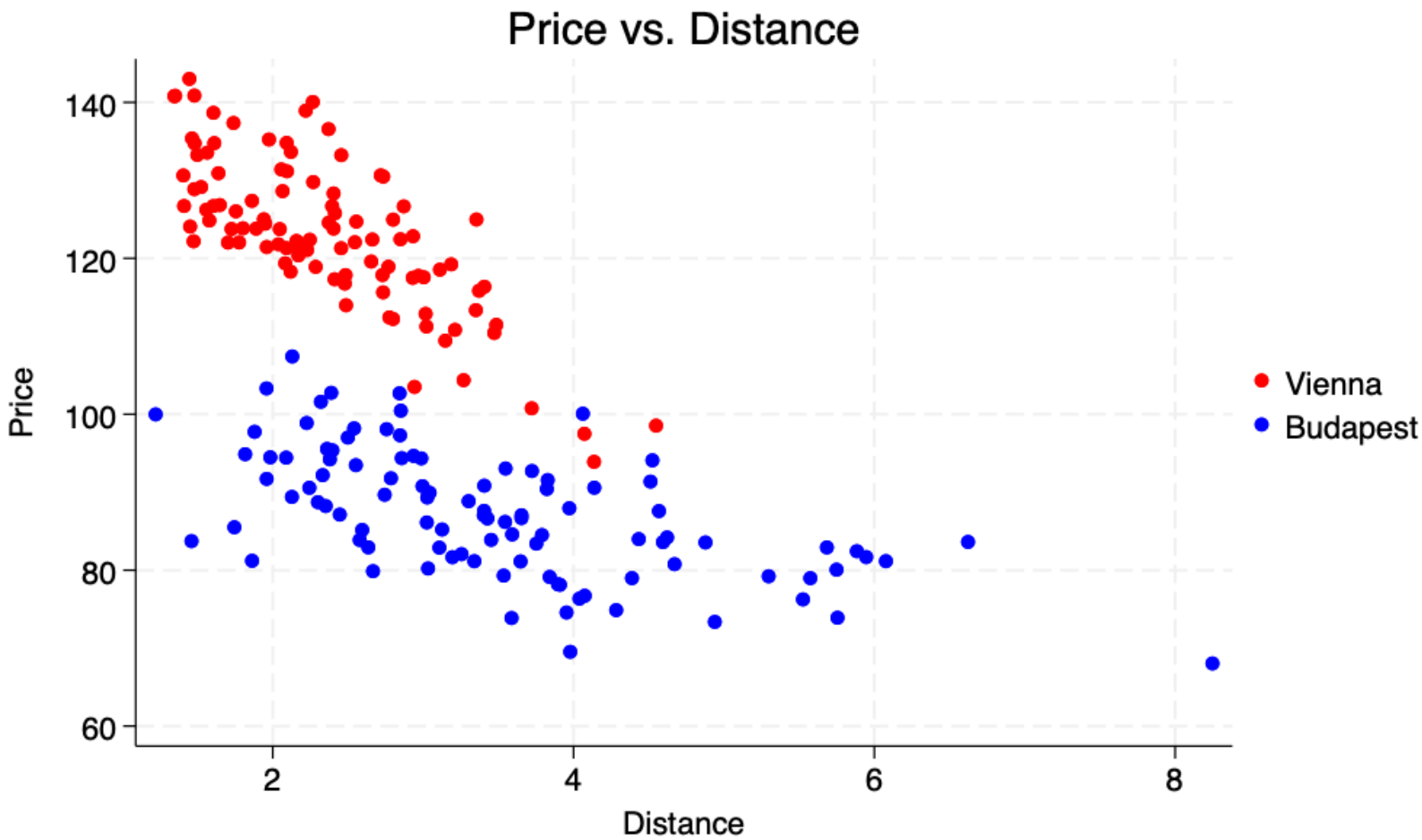

Figure 1: Hotel prices in two cities

The models are clearly different

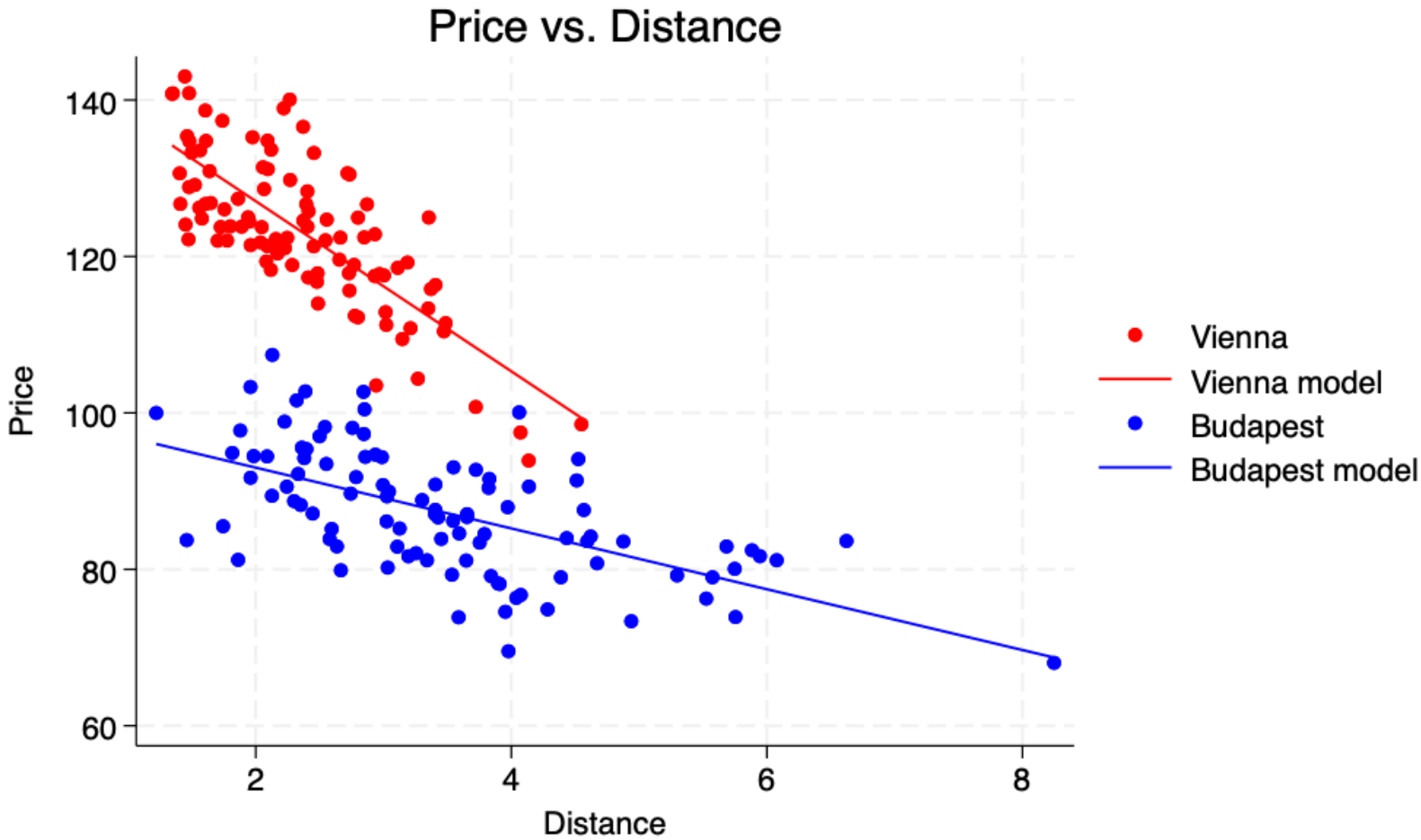

Figure 2: The models are clearly different

Oaxaca-Blinder decomposition

Predict the price of a hotel in Budapest using the Vienna model.

Subtact from the actual price in Budapest. This is the part explained by the characteristics ("hotels are different").

The remaining part is explained by the economic returns to these characteristics ("Budapest is cheaper").

Predicted prices

City/Model	Vienna	Budapest	Pricing difference
Vienna	**123**	92	31
Budapest	111	**87**	24
Characteristics	12	5	Total: 36

Oaxaca-Blinder decomposition

Of the $36 price difference,

$12 is due to the characteristics of the hotels

$24 is due to the economic returns to these characteristics.

Discrimination

The Oaxaca-Blinder decomposition is also used to measure discrimination.

If the difference

$$\text{wage}_A - \text{wage}_B$$

is due to differences in **legitimate** business practices, it is not discrimination.

It can help put a dollar value on discrimination.