

## Lab 4: Diagnostyka modelu.

Ekonometria WNE UW

1. W tym zadaniu będziemy pracować z danymi `firms.csv` i oszacujemy poniższy model:

$$\log(y) = \beta_0 + \beta_k \log(k) + \beta_l \log(l) + \beta_{k2} \log(k)^2 + \beta_{l2} \log(l)^2 + \beta_a age + \beta_{a2} age^2 + \beta_s sektor + \varepsilon \quad (1)$$

Wykorzystaj kod wczytujący dane i wykonujący podstawowe transformacje danych.

- (i) Oszacuj model.
- (ii) Sprawdź czy model cechuje się współliniowością. Jeśli tak, to co jest jej źródłem?
- (iii) Przetestuj czy zmienne podniesione do kwadratu są łącznie istotne statystycznie. Sprawdź czy w modelu bez tych zmiennych występuje współliniowość. Zdecyduj czy zatrzymasz je w modelu.
- (iv) Sprawdź poprawność formy funkcyjnej modelu.
- (v) Dodaj interakcję i przeprowadź test formy funkcyjnej jeszcze raz.
- (vi) Sprawdź czy składnik losowy ma rozkład normalny. Narysuj wykres qq.
- (vii) Sprawdź czy oszacowania są wrażliwe na odstające obserwacje: sprawdź dystans Cook'a, dźwignię i standaryzowane reszty. Oszacuj model na próbie obserwacji z dystansem Cook'a poniżej lub dźwignią poniżej 0.003. Porównaj z modelem szacowanym na całej próbie.
- (viii) Zdefiniuj zmienną *laborshare* jako stosunek kosztu pracy i produkcji (*costofl/y*) Czy parametry modelu różnią się dla firm z *laborshare* większym i mniejszym niż 0.5? Zastosuj test Chowa aby to przetestować. Czy wynik testu ma sens? Oszacuj dwa oddzielne modele i zaprezentuj wyniki. Czy widać różnicę?