Neural networks as an econometric estimator

A simulation study of the properties of neural networks as an estimator of conditional marginal effects.

R. Bjørn

August 17th 2018

Master's thesis defense

University of Copenhagen
rbjoern@outlook.com

Introduction

Argument: Neural networks provide a well-behaved big data estimator, which allows for fully heterogeneous marginal effects.

Relevance: Directly applicable to microeconometric research.

- Marginal effects of continuous regressors often of interest.
- Most estimators do not allow (full) heterogeneity.
- Neural networks can be extended to modern research designs.

Contribution: Illustration of 'novel' estimator and its properties.

Theory: Econometrics vs. data science

Literature review: Pessimists and optimists on causal learning.

Object of interest: Conditional Expectation Function (CEF).

- Causality assigned through assumptions/research design.
- But also optimal predictor in typical cases.

Theory: Econometrics vs. data science

Literature review: Pessimists and optimists on causal learning.

Object of interest: Conditional Expectation Function (CEF).

- Causality assigned through assumptions/research design.
- But also optimal predictor in typical cases.

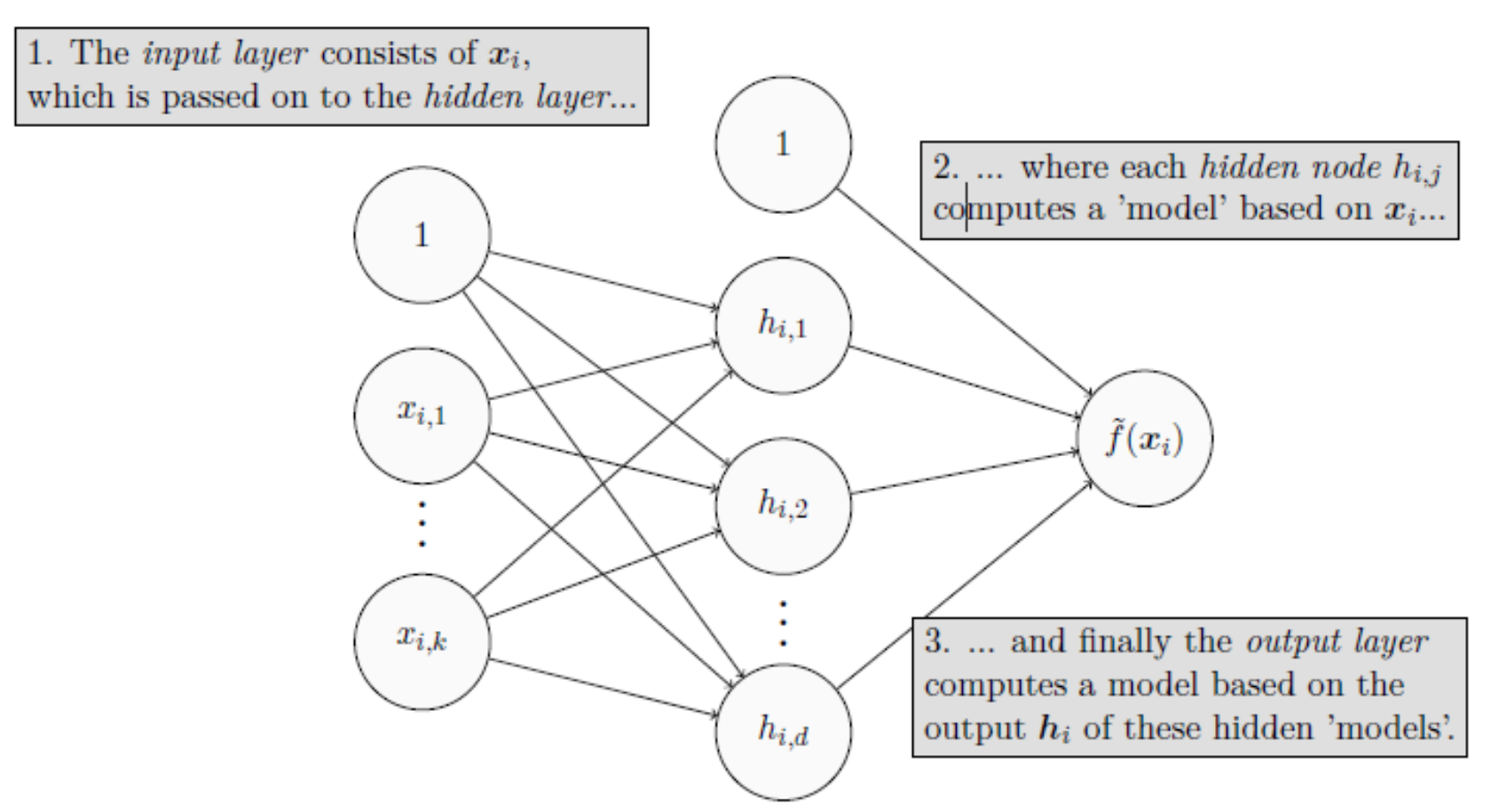
Theoretical framework:

- Assumptions: (i) weak unconfoundedness, (ii) outcome independence, (iii) common functional form.
- Estimators evaluated in general MSE framework.

$\Rightarrow$ Neural networks may work well (as *universal approximators*).

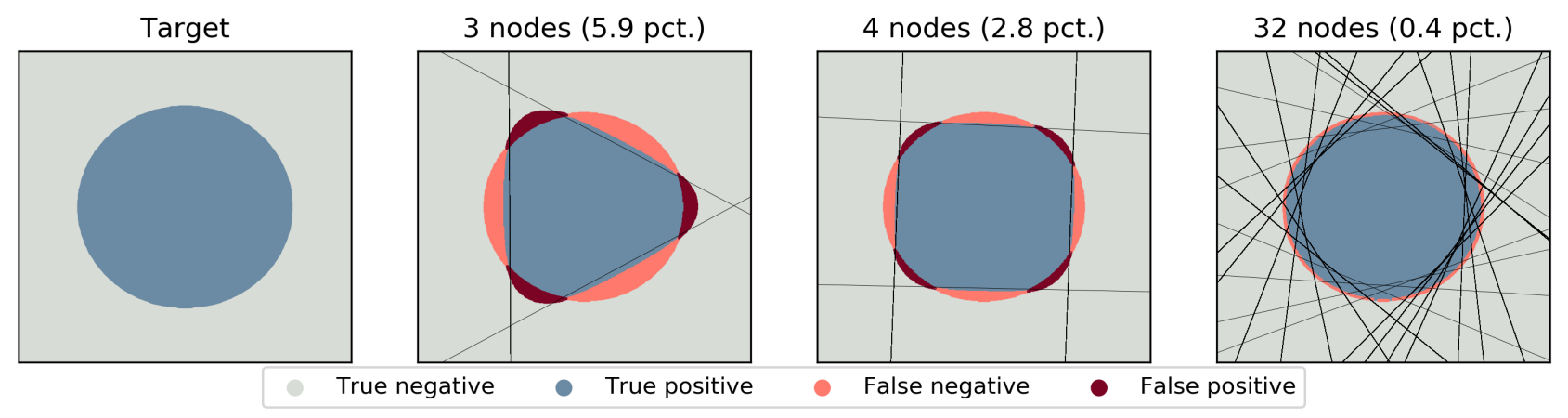
Model: Illustration of neural networks

Figure 3.1: A neural network with a single hidden layer



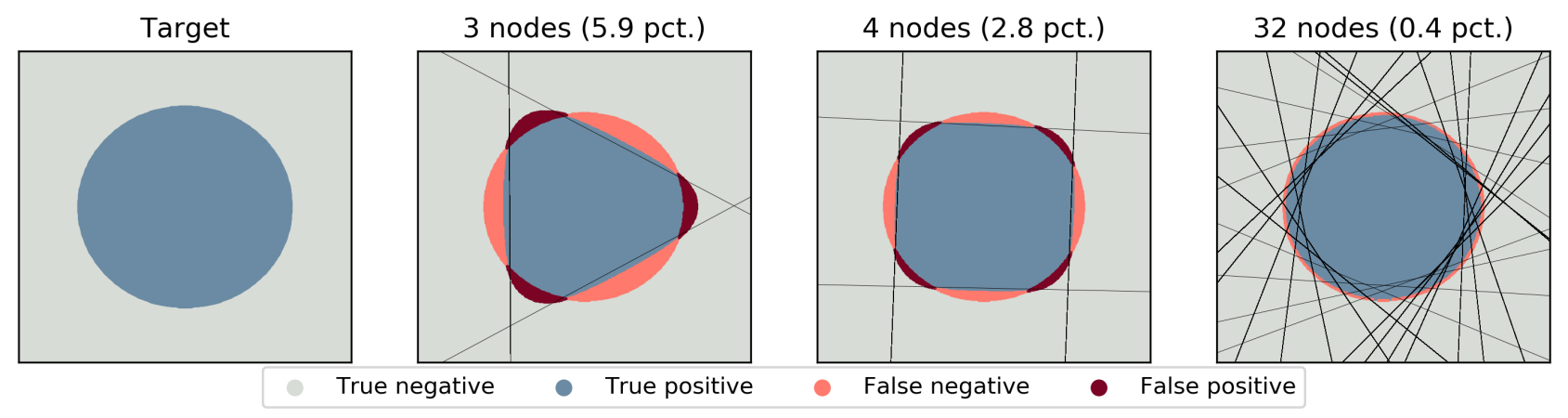
Model: Illustration of neural network

Figure 3.2: Alas, witness the approximation capacity of the mighty neural network!



Model: Illustration of neural network

Figure 3.2: Alas, witness the approximation capacity of the mighty neural network!



- Optimal network is found using MLE (which leads to CEF).
- Challenges: (i) optimization, (ii) overfitting, (iii) architecture.
- Continuous, so partial derivatives are well-defined.

Methodology: Summary of study

Research question: Can networks estimate marginal effects?

Research strategy: Monte Carlo simulation

- Flexible set-up for regression and classification scenarios.
- Nine data-generating processes, from linear to rather complex.
- Weakness: Results are confined to investigated cases.

Code module available on GitHub.

Far more detail in paper and appendices.

Results I: Well-behaved estimator in 'clean' case

Well-behaved big data estimator for marginal effects in 'clean' case:

- Estimates are consistent (for a low approximation error)...
- ... but exhibit high variance so large samples are needed.

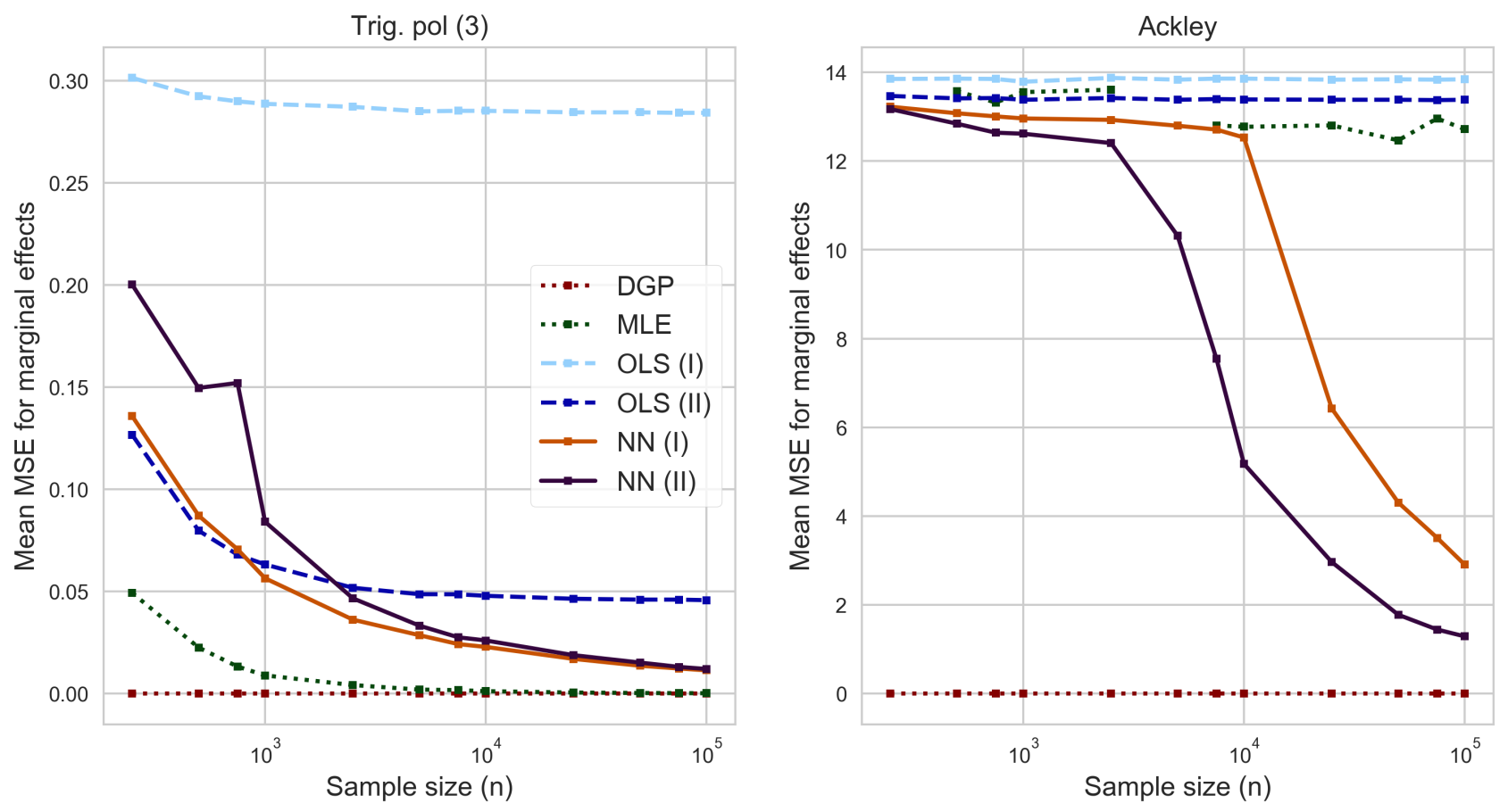
This is mainly interesting because it allows heterogenous effects:

- Neural networks are inefficient for average effects...
- ... but correctly reflects differences in individual effects.

Valid statistical inference is possible (although computationally expensive) through bootstrapping.

Results I: Well-behaved estimator in 'clean' case

Figure 5.1: Neural networks remain consistent for complex scenarios



Results II: Extension to instrumental variables

Neural networks are subject to traditional issues, e.g.:

- Measurement error leads to attenuation bias ($\rightarrow 0$).
- Confounders leads to omitted variable bias.

Fortunately, traditional solutions can also be extended to networks:

- I illustrate an extension for instrumental variables...
- ... which is valid, but requires a large effective sample.
- Other modern research designs can also likely be extended.

Replication of study

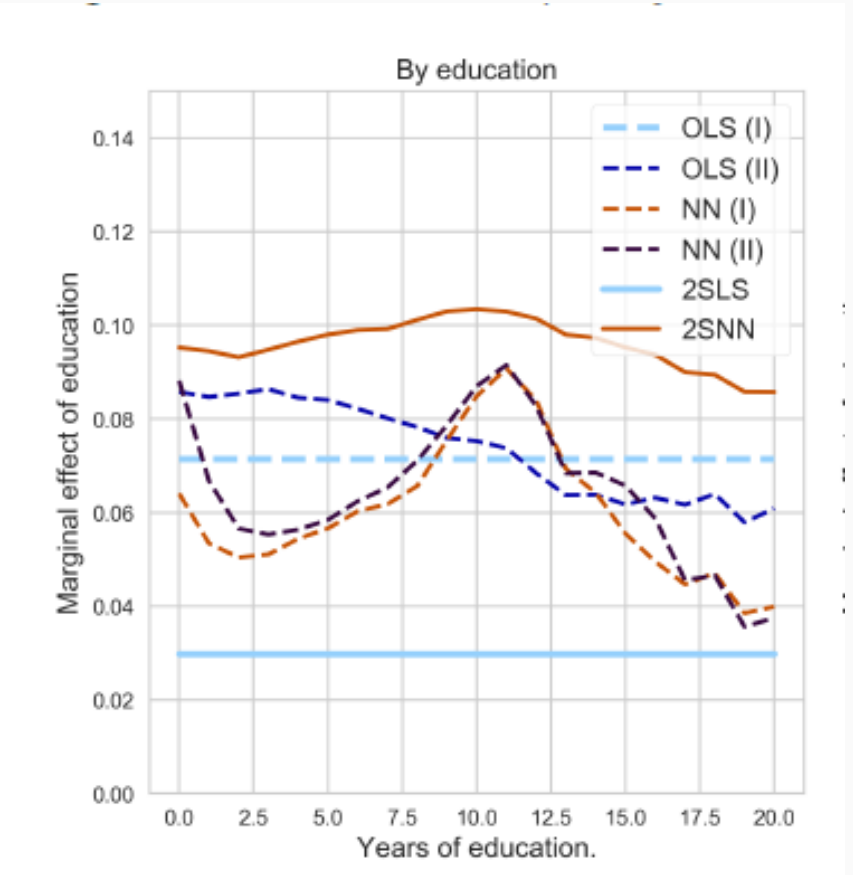
Angrist & Kruger (1991):

- Returns to schooling.
- IV: Compulsory schooling.

Neural network replication:

- Comparable average effects..
- ... although 2SNN is volatile.
- But effects appear to be heterogeneous!

Figure 6.1: Replication of study



Conclusion: Exciting econometric estimator

The question is far from closed...

- Do results hold beyond the simulation study?
- Are networks preferable compared to realistic alternatives?
- How would tweaking the networks affect estimates?

... but neural networks appear to provide an exciting estimator.

- Consistent estimates of fully heterogenous marginal effects.
- Volatile, but efficient enough for medium/big data.
- Bootstrapping allows statistical inference.
- Extendable to modern research designs.

Many studies could likely be enriched using neural networks.