

# Swiss Leading House

## Economics of Education, Firm Behaviour and Training Policies



Co-Director  
Prof. Dr. Uschi Backes-Gellner



Co-Director  
Prof. Dr. Stefan C. Wolter

---

Spring Term 2020

<p>Leading House Ph.D. Course “Machine Learning with Python for Education and Personnel Economists” - Syllabus -</p>
--

<b>Instructor</b>	<b>Michael E. Rose</b> , PhD, Post Doctoral Researcher at Max Planck Institute for Innovation and Competition, Munich, Germany
<b>Workshop dates</b>	March 9-13, 2020
<b>Location</b>	University of Zurich, Room tba
<b>Preliminary Schedule</b>	The course takes place in the form of an intensive 5-day course. Monday, March 9: 13:30-17:30 Tuesday, March 10: 09:00-17:30 Wednesday, March 11: 09:00-17:30 Thursday, March 12: 09:00-17:30 Friday, March 13: 09:00-12:00
<b>Module Number, ECTS</b>	tba; 3 ECTS
<b>Course Webpage</b>	<a href="https://www.educationeconomics.uzh.ch/en/Lehre/Courses.html">https://www.educationeconomics.uzh.ch/en/Lehre/Courses.html</a>

### Course Description

The course allows PhD students in Economics of Education and/or Personnel Economics to start their own empirical projects with Python, introduces the most important packages for empirical researchers and enables students to apply the most important Machine Learning models.

[www.economics-of-education.ch](http://www.economics-of-education.ch)

The Swiss Leading House on “Economics of Education, Firm Behaviour and Training Policies”  
A Research Programme of the State Secretariat for Education, Research and Innovation SERI

**Prof. Dr. Uschi Backes-Gellner**

Universität Zürich, Institut für Betriebswirtschaftslehre, Plattenstr. 14, CH-8032 Zürich  
Tel. +41 44 634 42 81, Fax +41 44 634 43 70, [backes-gellner@business.uzh.ch](mailto:backes-gellner@business.uzh.ch)

**Prof. Dr. Stefan C. Wolter**

Universität Bern, Volkswirt. Institut, Forschungsstelle für Bildungsökonomie, Schanzeneckstr. 1, Postfach 8573, CH-3001 Bern  
Tel. +41 31 631 36 56, Fax +41 31 631 39 92, [stefan.wolter@vwi.unibe.ch](mailto:stefan.wolter@vwi.unibe.ch)

# Swiss Leading House

## Economics of Education, Firm Behaviour and Training Policies



Co-Director  
Prof. Dr. Uschi Backes-Gellner



Co-Director  
Prof. Dr. Stefan C. Wolter

---

### Topics / Content

#### Unit 1: Empirical research using Python

- pandas as key module for data wrangling in Python
- Data formats: csv, json, xml, pdf, html, plain text
- Plotting with Matplotlib and seaborn

#### Unit 2: Project management

- Coding workflow (script mode and terminal vs. notebook vs. IDE)
- Debugging and scaffolding
- Version control with Git and GitHub
- Script design and PEP8

#### Unit 3: Supervised Machine Learning

- What is Machine Learning?
- Workflow I: Split, train, evaluate
- Distances and evaluation metrics
- Linear models, regularization
- [Decision trees and random forests]
- Neural networks
- Feature engineering
- Workflow II: Model selection, Pipelines and Cross-Validation

#### Unit 4: Unsupervised Machine Learning

- Difference between supervised and unsupervised learning
- Scaling
- [Principle Component Analysis]
- Agglomerative clustering
- Hierarchical clustering
- [DBSCAN]
- Applied software: dedupe

#### Unit 5: Natural Language Processing

- Sentiment analysis with TextBlob, Readability, regular expression
- Preprocessing text: Cleaning, Stemming, Tokenizing, Vectorization
- [Clustering text]
- Topic modelling with Latent Dirichlet Analysis
- [Word clouds]

# Swiss Leading House

## Economics of Education, Firm Behaviour and Training Policies



Co-Director  
Prof. Dr. Uschi Backes-Gellner



Co-Director  
Prof. Dr. Stefan C. Wolter

### Prerequisites to participate in the course

Basic knowledge in Python is a requirement to participate in the course. To ensure sufficient knowledge students have to successfully finish the free UdeMy online course "Python for Absolute Beginners" (about 5 hours) before Sunday, March 1<sup>st</sup>, 2020, 20:00h CET: <https://www.udemy.com/course/learn-python-3-from-scratch-python-for-absolute-beginners/>. Students may apply for a waiver if they submit a proof of sufficient knowledge in Python.

### Required Literature

- Shapiro, J. and M. Gentzkow: "Code and Data for the Social Sciences: A Practitioners Guide"
- Gentzkow, M., B. Kelly and M. Taddy (JEL 2019): "Text as Data"

### Further readings

- Downey, Allen B.: "How to think like a Computer Scientist"
- Mueller, A. and Sarah Guido: "Introduction to Machine Learning with Python"

### Teacher

- Michael E. Rose, PhD, Post Doctoral Researcher at MPI for Innovation and Competition
- Experience in teaching ML and Big Data in Python to PhD students (LMU Munich, ifo institute, and Georgia Institute of Technology), Computational Mathematics in Matlab to Master's students (U Cape Town), Time Series Econometrics in Python to Master's students (U Cape Town), SQL, Excel and VBA to Master's students (U Cape Town)
- Daily usage of Python, ML in own research
- Lead development of two python packages (pybliometrics and sosia)

### Target audience

The course is particularly designed for doctoral students in the course programme on economics of education of the Swiss Leading House. Doctoral students in economics or business economics with a strong interest in economics of education or personnel economics and machine learning are welcome as well. The seminar will take place en bloc in order to enable external Ph.D. students to attend.

### Credit Requirements and Grading

1. Full course attendance. Students are expected to come prepared to class and exercises. It will facilitate discussion and improve overall learning.
2. Each course unit is complemented with a training session that takes place in the afternoons. Each student has to finish all the exercises presented during the training sessions at the end of the week (Sunday, March 15<sup>th</sup>, 20:00 CET). The work is to be done in groups of two (or three if there is an odd number of participants).

### Application

The number of participants is limited. Course registration until **January 13, 2020** is mandatory. Please send your application (including a one-page CV) to [fabienne.kiener@business.uzh.ch](mailto:fabienne.kiener@business.uzh.ch).

# Swiss Leading House

## Economics of Education, Firm Behaviour and Training Policies



Co-Director  
Prof. Dr. Uschi Backes-Gellner



Co-Director  
Prof. Dr. Stefan C. Wolter

---

### WWF Statutory Course Policies

According to WWF study regulations, all exam dates are final as published in the VVZ and syllabus. This means that the final exam date is not negotiable. It will not be possible to take any exams on different dates.

Academic dishonesty in any form will not be tolerated. Anyone caught cheating or engaging in unethical behavior will be reported to the Dean's office according to the guidelines on academic dishonesty set forth by the University of Zurich.

The information in this syllabus supports the official information in the electronic university registration tool (VVZ – Vorlesungsverzeichnis). In cases of doubt, the official information at the VVZ is decisive.

For UZH students: Don't forget to officially register using the registration tool of the University of Zurich.