# Noah Stegehuis

**Address:** President Steynstraat, 7-1, Amsterdam, The Netherlands
**Tel:** +31628136276
**Email:** noahjstegehuis@gmail.com **Personal website**
**Citizenship:** Dutch
**Date of Birth:** 03/08/1997

**Placement director:** Prof. Eric Bartelsman e.j.bartelsman@vu.nl
**Placement assistant:** Christina Månsson c.mansson@tinbergen.nl

## EDUCATION

2020-2025 **Vrije Universiteit Amsterdam and Tinbergen Institute**
Ph.D. Econometrics and Data Science.

- Contributor to project of F. Blasques "Econometric Methods for Incorrect Models" (Grant nr: Vidi.195.099).
- Supervisors: Prof. Francisco Blasques, Prof. Siem Jan Koopman and Assoc. Prof. Paolo Gorgi.
- Research: Theoretical developments for causal inference with experimental and observational data, time varying instrumental variable parameter models, synthetic control methods

2019-2020 **Vrije Universiteit Amsterdam**
MSc Econometric Theory (Honours Programme), *Cum Laude.*

2015-2018 **Maastricht University**
BSc Econometrics & Operations Research (Honours Programme), *Cum Laude.* Exchange Semester Econometrics at Hong Kong University of Science and Technology (HKUST).

## RESEARCH FIELDS

Primary fields: Econometrics, Causal Inference, Data Science
Secondary fields: Time series analysis, Macroeconometrics, Machine Learning

## REFERENCES

Prof. Francisco Blasques, Vrije Universiteit Amsterdam, f.blasques@vu.nl
Prof. Siem Jan Koopman, Vrije Universiteit Amsterdam, s.j.koopman@vu.nl
Assoc. Prof. Paolo Gorgi, Vrije Universiteit Amsterdam, p.gorgi@vu.nl

## TEACHING EXPERIENCE

2020-2024 **Teaching Assistant**, Vrije Universiteit Amsterdam, Advanced Econometrics (Msc Econometrics & Data Science)

2020-2024 **Teaching Assistant**, Vrije Universiteit Amsterdam, Business Statistics (BSc International Business Administration)

2020-2024 **Thesis Supervisor** (BSc, MSc Econometrics & Data Science)

## OTHER PROFESSIONAL EXPERIENCE

2021-2022 **Churned AI, Research Data Scientist**

- Designed and developed a scalable data pipeline for automating machine learning model estimation, regularisation and prediction of client datasets
- Developed a Python package that is integrated into the dashboards of client companies, providing real-time insights into their customer churn rates

2018-2019 **Logex Healthcare Analytics, Data Science Work Student**

- Contributed to data science projects with large-scale societal impact, providing benchmark reports for medical specialists and healthcare stakeholders to improve healthcare policy
- Created and visualized reports for the Dutch National Healthcare Authority and the first international clients

2018 **Accenture, Bachelor Thesis Intern**

- Developed sales forecasting models using Vector Auto-Regressive techniques for a Dutch football club, delivering actionable recommendations.

## RESEARCH PAPERS

*A Score-Driven Filter for Causal Regression Models with Time-Varying Parameters and Endogenous Regressors* (with F. Blasques) **[Job market paper]**

Abstract: This paper proposes a score-driven model for filtering time-varying causal parameters using instrumental variables. In the presence of suitable instruments, we show that we can uncover dynamic causal relations between variables, even in the presence of regressor endogeneity which may arise because of simultaneity, omitted variables, or measurement errors. Due to the observation-driven nature of score models, the filtering method is simple and practical to implement. We establish the asymptotic properties of the maximum likelihood estimator and show that the instrumental-variable score-driven filter converges to the unique unknown causal path of the true parameter. We further analyze the finite sample properties of the filtered causal parameter in a comprehensive Monte Carlo exercise. Finally, we reveal the empirical relevance of this method in an application to aggregate consumption in macroeconomic data and we provide a time-varying estimate of price elasticity of demand for a dataset on recorded market prices.

*Mitigating Estimation Risk: A Data-Driven Fusion of Experimental and Observational Data* (with F. Blasques, S.J. Koopman, P. Gorgi) [submitted]

Abstract: The identification of causal effects of marketing campaigns (advertisements, discounts, promotions, loyalty programs) require the collection of experimental data. Such data sets frequently suffer from limited sample sizes due to constraints (time, budget) which can result in imprecise estimators and inconclusive outcomes. At the same time, companies passively

accumulate observational data which oftentimes cannot be used to measure causal effects of marketing campaigns due to endogeneity issues. In this paper we show how estimation uncertainty of causal effects can be reduced by combining the two data sources by employing a self-regulatory weighting scheme that adapts to the underlying bias and variance. We also introduce an instrument-free exogeneity test designed to assess whether the observational data is significantly endogenous, and experimentation is necessary. To demonstrate the effectiveness of our approach, we implement the combined estimator for a real-life data set in which returning customers were awarded with a discount. We demonstrate how the indecisive result of the experimental data alone can be improved by our weighted estimator and arrive to the conclusion that the loyalty discount has a notably negative effect on net sales.

## RESEARCH IN PROGRESS

*Extending Generalised Synthetic Control: Total Effect Estimation in the Presence of Mediators*
(with F. Blasques)

Abstract: Synthetic control methods have recently gained significant popularity for the identification of causal effects of policy decisions. Generalised synthetic control (Xu, 2017) offers an alternative to traditional synthetic control methods, by directly modelling the time-varying heterogeneity using an interactive fixed effects panel data model, which allows for the incorporation of time varying covariates. The counterfactual outcome is then constructed by simply taking the fitted values, using the treated unit's own covariates and the estimated parameters, factors and loadings. However, when a treated unit's covariates are also affected by treatment (mediators), then this counterfactual does not form a reliable basis for what would have truly happened in absence of the intervention. With generalised synthetic control (GSC) one would only estimate the direct effect, and not the total effect that includes the change in regressors. In this paper we propose an extension of the GSC estimator, that allows for the estimation of the total effect of an intervention.

References:
Xu, Y. (2017). Generalized synthetic control method: Causal inference with interactive fixed effects models. Political Analysis, 25 (1), 57–76.

## CONFERENCES AND SEMINARS

| | |
|---|---|
| 2024 | VU seminar, Econometrics department, Amsterdam |
| 2024 | Netherlands Econometrics Study Group, Maastricht |
| 2023 | International Association for Applied Econometrics (IAAE) in Oslo |
| 2023 | International Conference on Econometrics and Statistics (EcoSta) in Tokyo, Japan |
| 2023 | Netherlands Econometrics Study Group, Rotterdam |
| 2021 | VU seminar, Econometrics department, Amsterdam |

## MISCELLANEOUS

**Programming skills:** Python, R, SQL, Matlab, LaTeX, MS Excel (VBA), Git
**AI & Data Science Tools:** causal inference, deep learning architectures (CNNs, RNNs), scikit-learn, TensorFlow, PyTorch, ensemble methods (RF, gradient boosting), XGBoost, Matplotlib, Seaborn, NumPy, Pandas, Statsmodels, joblib, pickle
**Languages:** English (fluent), Dutch (native), French (intermediate), Spanish (pre-intermediate)
**Hobbies and interests:** Reading, ballet, film, art, architecture, travelling, baking