2025

BELGIUM STATA CONFERENCE

PROGRAM

KU Leuven

Campus Brussels

Warmoesberg 26 - 1000 Brussels (BELGIUM)

September 9, 2025

https://www.stata.com/meeting/belgium25/

09:00 am **Registration and Coffee**

09:30 am **Opening**

09:45 am **Keynote** *(1 hour)*

- **Benn Jann (University of Bern)**
  "*Drawing maps using Stata: the geoplot command*"

10:45 am **Coffee Break**

11:00 am **Presentations session 1** *(1h30)*

- **Nicolas Debarsy (University of Lille)**

  "*Efficient Estimation of Regression Models with Spillovers: Flexible Parametric and Semi-Parametric Approaches*"

  This presentation introduces alternative methods for efficiently estimating regression models that include spillover effects, i.e. situations where outcomes are influenced by interactions between individuals. In such settings, the ordinary least squares estimator is generally inappropriate due to simultaneity. While established techniques such as two-stage least squares, generalized method of moments, and (quasi-) maximum likelihood are widely applied, they may fall short of efficiency except under restrictive conditions.
  We propose two estimation strategies grounded in Le Cam's local asymptotic normality (LAN) theory, each differing in how they treat the distribution of the error terms. The first is a semi-parametrically efficient estimator based on residual ranks and signs, requiring only the error distribution to be strongly unimodal. The second is a fully parametric estimator designed to handle skewness and heavy tails, using flexible families of distributions such as Tukey's g-and-h and the sinh-arcsinh distributions.
  Monte Carlo simulations show that both estimators perform well across a range of designs and often outperform more conventional methods. To illustrate their practical value, we revisit the trade regression of Behrens et al. (2012), demonstrating how the results can shift considerably when Gaussian assumptions are relaxed. We also provide a Stata implementation to facilitate replication and application in empirical work.

- **Nicolas Gerardy (UCLouvain)**

  "*Effect of Destination Grammatical Gender on Perceived Destination Masculinity and Femininity*"

  This research explores how grammatical gender - an inherent feature of many languages - influences destination gender perceptions. To investigate this, we designed a survey-based experiment in which participants from different linguistic backgrounds (e.g., French and English) rated destination names on gendered attributes (e.g., perceived masculinity or femininity). The dataset comprises hierarchical responses, with multiple evaluations per respondent, structured to allow for inter-individual variation. We analyzed the data using multiple linear regressions with random intercepts, estimated in Stata via the xtreg command, which accommodates nested data and random effects efficiently. We explain that the grammatical gender of destination names predicts their perceived masculinity or

femininity among speakers of gendered languages such as French. English speakers, whose language lacks grammatical gender, are unaffected by this effect. These findings support the theory of linguistic relativity, showing that language structure can influence cognitive evaluations of non-animate objects. The research contributes to branding and tourism literature by identifying grammatical gender as a key yet overlooked determinant of brand personality. Managerial implications include the strategic importance of considering grammatical gender when naming brands or marketing destinations in gendered-language contexts. Ultimately, the study proposes that international marketers integrate linguistic features into brand positioning strategies, especially when targeting diverse linguistic audiences.

- **Laure Oostenbach (UAntwerpen)**

  "*Profiling users and non-users of meal delivery services in Belgium using latent class analysis*"

  Laura Oostenbach, Matthew Keeble, Thomas Vanoutrive, Maartje Poelman, Carlijn Kamphuis & Lukar Thornton

Background: Ready-to-eat meal delivery services have become increasingly popular in recent years, with potential detrimental health implications as the majority of food promoted and sold is energy-dense and nutrient-poor. However, limited research has examined who uses these services and why. This study explores the profiles of different users and non-users of meal delivery services and further describes the socio-demographic and health characteristics of these profiles.
Methods: Data were from 1086 adults who completed the online 2024 Meal Delivery Survey conducted in the Flanders and Brussels regions of Belgium. Users of meal delivery services reported on usage indicators (e.g., valuing fast delivery, ordering to avoid traffic) and non-users reported on reasons for non-use (e.g., lacking trust in meal hygiene, finding it bad for the environment). Latent class analysis categorised users based on patterns of indicator endorsement, using Stata *gsem* command with *lclass()* option. Amongst non-users, reasons for non-use were examined descriptively for both former users and never users. Socio-demographic and health characteristics were compared across profiles. Analyses were conducted in Stata 18.
Results: Over 66% of the sample (n=720) were users and 34% (n=366) non-users. Latent class analysis identified two user profiles. Efficiency-focused users (48% of user sample) solely prioritised fast and affordable service. Variety- and convenience-driven users (52% of user sample) endorsed a wider range of indicators, including the ability to try different cuisines, avoid supermarkets, and have more leisure time. The latter group included more frequent users. Amongst non-users, both former users (52% of non-user sample) and never users (48% of non-user sample) most commonly reported preferences for in-store food shopping and home cooking as reasons why they did not order meals for delivery. Profiles differed in socio-demographic and health characteristics. For example, variety- and convenience-driven users reported poorer health than efficiency-focused users. Non-users (former and never) were older than users, with never users having the oldest age profile.
Conclusions: This study provides a better understanding of determinants of use and non-use of meal delivery services, informing public health action aimed at improving food behaviours and reducing the burden of diet-related diseases. Results can support the development of targeted interventions addressing main drivers of meal delivery service use.

12:30 am **Lunch**

01:30 pm **Training seminar** *(45 mn)*

- **Nick Deschacht (KU Leuven)**
  "*Producing automated tables in Stata using collect*"

02:15 pm **Presentations session 2** *(1h30)*

- **Philippe Van Kerm (University of Luxembourg)**
  "Shapley value calculations: Implementation and illustrations"

  This talk will illustrate the use of the Shapley value in regression and various decomposition analyses. It will first introduce the concept of the Shapely value and related measures. It will then describe its use in regression and different types of decomposition analyses. It will introduce a prefix command to facilitate implementation of calculations of the Shapley value in Stata.

- **Diogo Conceição (KU Leuven)**
  "Anchors Aweigh? Early Cognitive Biases in Flemish Classrooms"

  This study investigates the impact of anchoring on primary school students' decisions regarding summer school participation in Flanders, Belgium. Making use of a randomized controlled trial, it is shown that students exposed to a low irrelevant anchor recalled, on average, two fewer words from a list of 15 compared to their peers. This effect was particularly pronounced among students with high SES, low achievement, negative perceptions of school, or high achievement (between 2.136 and 3.009 fewer words). A theoretical framework is also proposed in order to understand the mechanism behind this effect, positing that it is based on the subjective beliefs of individuals about their probability of being right or wrong in their prediction, conditional on having been exposed to the anchor. The results suggest that anchoring influences cognitive tasks beyond alternative explanations such as ability, self-confidence, or motivation. It is also shown that taking into account the proposed mechanism essentially erases all effects. The findings have implications for public policy, particularly in the context of taking advantage of these mental shortcuts to increase effectiveness.

- **Vincenzo Verardi (UCLouvain)**
  "Skellam Regression in Stata"

  The Skellam distribution models the difference between two Poisson random variables that may have different means. In its classical definition, the two Poisson variables are assumed to be independent. However, the distribution also covers cases where dependence arises through an additive common component that cancels out when taking the difference.

  It is named after the British statistician and ecologist John Gordon Skellam (1914–1979).

  This distribution generalises the Irwin distribution (Irwin, 1937), which models the difference between two independent Poisson random variables sharing the same mean.

A Skellam regression relies on Maximum Likelihood estimation to link the conditional means of the underlying Poisson processes to a set of covariates.

In this talk, we show how to formulate the likelihood, derive the gradient and the Hessian for numerical optimisation, and implement the model in Stata. We also present an example based on football (soccer) scores, where the difference in goals between two teams is naturally modelled with a Skellam distribution.

03:45 pm **Coffee Break**

04:15 pm **Keynote** *(1 hour)*

- **Di Liu (StataCorp)**
  "*Conditional average treatment-effects estimation using Stata*"

Treatment effects estimate the causal effects of a treatment on an outcome. The effect may be heterogeneous. Average treatment effects conditional on a set of variables (CATEs) help us understand heterogeneous treatment effects. By construction, they are useful to evaluate how different treatment-assignment policies affect different groups in the population.

In this talk, we will show how to use Stata 19's new command **cate** to answer questions such as the following:

1. Are the treatment effects heterogeneous?
2. How do the treatment effects vary with some variables?
3. Do the treatment effects vary across prespecified groups?
4. Are there unknown groups in the data for which treatment effects differ?
5. Which is best among possible treatment-assignment rules?

05:15 pm **Open Panel Discussion with Stata Developers** *(15 min)*

Contribute to the Stata community by sharing your feedback with StataCorp's developers. From feature improvements to bug fixes and new ways to analyze data, we want to hear how Stata can be made better for you.

05:30 pm **End of the conference & Drink**

SCIENTIFIC COMMITTEE

Guido Pepermans *(KU Leuven)*, **Ilse Ruyssen** *(UGent)*, **Jef Hendrickx** *(KU Leuven)*, **Kristien Coucke** *(KU Leuven)*, **Nick Deschacht** *(KU Leuven)*, **Sébastien Fontenay** *(Universitat Pompeu Fabra Barcelona)*, **Simon De Jaeger** *(KU Leuven)*, **Stijn Vanormelingen** *(KU Leuven)*, **Sunčica Vujić** *(Universiteit Antwerpen)*, **Vincent Vandenberghe** *(UCLouvain)*, **Vincenzo Verardi** *(UCLouvain)*.

Rue des anciens étangs 40
1190 Brussels Forest (BELGIUM)
www.ritme.com/en

4905 Lakeway Drive College Station
Texas 77845-4512 (USA)
www.stata.com

Warmoesberg 26
1000 Brussels (BELGIUM)
www.kuleuven.be