UNIVERSITY OF MANNHEIM
School of Social Sciences

Chair of Political Science,
Quantitative Methods
in the Social Sciences

Prof. Thomas Gschwend, Ph.D.
A5, 6, Room B 220
D-68131 Mannheim
0621 181-2087
gschwend@uni-mannheim.de

Assistant: Ursula Eckle
0621 181-2413
methods@uni-mannheim.de

Advanced Quantitative Methods

Spring 2026

Course Details:

Thomas Gschwend (Lecture)	Domantas Undzėnas (Lab)	Lisa-Marie Müller (Lab)
gschwend@uni-mannheim.de	domantas.undzenas@uni-mannheim.de	lisa-marie.mueller@uni-mannheim.de
Wednesday, 8:30–10:00	Thursday, 10:15–11:45	Friday, 10:15–11:45
A5, 6, B244	B6, 23-25, A 103	A5, 6, B143
Tuesday, 13:30–14:30	Tuesday, 14:30–16:00	Tuesday, 15:30–17:00

uni-mannheim-aqm-2026 aqm2026.slack.com aqm-uma.netlify.app

Course Description:

Building on the analytical and theoretical background of the previous course in our MA methods sequence ("Quantitative Methods"), this course on "Advanced Quantitative Methods" introduces interested graduate students to strategies and tools to develop statistical models that are tailored to answer their particular research questions.

You might have noticed by now the linear regression model is often an inappropriate tool for answering substantive questions in political science. This course serves as an introduction to a multitude of probability models that are appropriate when the linear model is inadequate. After introducing the fundamentals from which statistical models are developed, this course will focus on one specific theory of inference, namely on the statistical theory of maximum likelihood. We will also devote considerable time to statistical programming, including the simulation and communication of quantities of substantive interest from these models (using R), in order to encourage students to switch from a consumer-mode to a producer-mode of social science research.

The goal of this course is three-fold: (1) to equip students with the skills needed to conduct research using appropriate statistical models and effectively communicate their findings to a non-technical audience; (2) to establish a solid foundation in the theory of maximum likelihood, enabling students to explore and implement a diverse array of advanced statistical models; and (3) to provide students with the tools necessary to fine-tune existing or to develop new statistical models of political phenomena.

Work through the assigned readings ahead of time. Connect the equations to substantive questions they are meant to answer. We expect everyone to come to class fully prepared. Expect that this will take considerably longer than in a substantive seminar. Do not skip equations! Instead, take notes, prepare questions and team-up with others to answer them, or as last resort, ask them in class. After every class we expect you to go over the lecture notes and your notes again. Furthermore, we additionally offer the possibility to send us questions by Wednesday night. We will try to address them in the lab session on Thursday or Friday. There is no point in getting lost — particularly not in an elective class. Nevertheless, understand that the bulk of learning in this course will take place *outside* the classroom, by reading, practicing using statistical software, and solving problem sets.

Recommended for:

Graduate students in political science in the M.A. Political Science and CDSS PhD students as well as Mannheim Master in Data Science (MMDS) and Mannheim Master in Social Data Science (MMSDS) students. Interested PhD students from other GESS centers can participate subject to the availability of seats.

Prerequisites:

Master students (M.A. Political Science, MMDS, MMSDS) should have successfully passed the previous course in the political science methods sequence "Quantitative Methods" and the accompanying "Tutorial Quantitative Methods", preferably (but not necessarily) with a final grade of 2.0 or better. PhD students should have passed equivalent courses. If you know what $(X'X)^{-1}X'y$ is, you have the necessary background to take this class.

Course Registration:

Students who wish to take the course should register for "Advanced Quantitative Methods" at the student portal. Please note, that the course registration is only complete when you are admitted to the ILIAS group of the course. Furthermore, to be able to receive the grades, students are required to register for the examination in both the lecture and the tutorial during the semester, additionally to the course registration.

Note that this course is highly demanding and entails a substantial work load for students! Students who wish to audit this class should notify the instructor in advance (participation is subject to free room capacity). Please note that only registered students will receive feedback on their written work.

Readings:

We will not use a single textbook for this course. Selected readings are available on the course website (through ILIAS). The following books will be used in the course:

Eliason, Scott R. 1993. *Maximum Likelihood Estimation: Logic and Practice.* Newbury Park: Sage.

King, Gary. 1989. *Unifying Political Methodology.* Ann Arbor: University of Michigan Press.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Newbury Park.: Sage.

Software:

Students need to bring their own computers to lab sessions. R will be the software package of choice. There will be homework problems that require you to edit and write some R-code. The open-source statistical programming language R is particularly suited for carrying out state-of-the-art computer-based simulations and programming advanced statistical models. It also generates really nice publication-quality graphics. The software runs under a wide array of operating systems. R can be downloaded for free at `http://www.r-project.org/`.

A very good graphical user interface for R (which we will also use during the lab sessions) is RStudio. In recent years a growing number of features have been added to this graphical user interface, which makes it the preferred choice for learning R – also for beginners. It is cross-platform and open-source. RStudio can be downloaded for free at `https://posit.co`. A style guide to make your code easier to read, share, and verify can be found at `http://adv-r.had.co.nz/Style.html`. Please make sure to install the latest versions of R and RStudio before the first lab session.

To facilitate an efficient workflow, we integrate Github into the course. git is a version control system that makes it easy to track changes and work on code collaboratively. GitHub is a hosting service for git. You can think of it like a public Dropbox for code. We will use it to distribute code and assignments to you. And you will use it to keep track of your code and collaborate in teams. You can find the course on GitHub here.

Course Requirements:

Grading will be based on the following components:

- **Homework Assignments (25%)**

  There will be six graded mandatory homework assignments in the form of problem sets, replications, simulations, or extensions of the analyses conducted in class and in the lab. Assignments will be made available after class on Wednesday, and solutions must be uploaded by 23:59 on the following Wednesday (unless noted otherwise). You are required to complete every task in each assignment, and the contents of your GitHub repositories must be sufficient to replicate the analysis *by the deadline*. Late submissions and non-reproducible files will not be accepted. Moreover, note our policy below that clarifies how AI-based tools may and may not be used when working on class assignments.

You will work in small groups on the assignments. Usually 2–3 people per group works best. Please indicate with whom you worked on the assignment on your homework. Moreover, you are strongly encouraged to seek advice from the instructors during office hours or via Slack, also *before* the submission deadline. Note that instructive discussions about the material are best done during office hours rather than via Slack.

- **Final Paper (75%)**

  There will be a final draft paper but no final exam. Each student will produce a co-authored manuscript (or a solo-authored manuscript, with permission of the instructor). This manuscript should apply or develop an appropriate statistical model to an important substantive problem. Students will choose their own topics. What works particularly well is to start replicating an already published article to develop it into a different manuscript using your own argument. We recommend selecting an article of personal interest, published in a top journal within the last few years, and employing methods discussed or to be covered in the course (or alternative methods at about the same level of sophistication).

  The draft paper must include all analyses, tables, figures, and descriptions of the results. A good write-up of the draft paper should read like the third quarter of a journal article. The rest of the draft may be in detailed outline form, although it would be better to have it fully written.

  You also need to provide all necessary information to replicate your analysis. The replication material must include your data and computer code to be able to reproduce all tables and figures that make it in the paper. We expect you to comment your computer code heavily to explain what you are doing. Your code must be neatly formatted and run without problems. To that end, please avoid writing computer-specific lines into your code that will prevent it from running on other machines. We will award partial credit if necessary.

  The final draft paper together with all replication material are due on **3 June 2026**. Please submit all files electronically via Github by **10am CEST** that day. Late submissions will not be accepted.

Policy Regarding the Use of Large Language Models (LLMs).

Large Language Models (LLMs), such as ChatGPT, GitHub Copilot, and similar AI-based tools, are increasingly used in academic and professional contexts. In this course, their use is not categorically prohibited, but it is subject to important limitations and expectations.

LLMs can be a legitimate and productive aid for certain tasks, particularly in coding and statistical analysis. For example, they can help identify syntax errors, explain error messages, and suggest debugging strategies. Used in this way, LLMs can save time, especially when working with unfamiliar software or programming languages.

However, engaging with the homework assignments required to pass this course is a central part of how you learn in this course. The purpose of these assignments is not merely to produce correct output, but to develop the ability to reason through problems, design solutions, understand underlying assumptions, and interpret results. If you rely on LLMs to generate solutions for you, rather than engaging directly with the material, you bypass this learning process. As a result:

- You will not develop the conceptual understanding required to adapt methods to new problems.
- You will struggle when asked to explain, modify, or extend your own code or analysis.
- You are likely to perform poorly on the final paper, which requires developing your own idea, adapting methods to your case, justifying model selection, and meaningfully interpreting results. These tasks cannot be outsourced to an LLM.

You are expected to complete the homework assignments on your own, actively engage with the challenges they pose, and discuss ideas with other students. Struggling, making mistakes, debugging, and revising are essential parts of the learning process. Using an LLM as a shortcut undermines this process and ultimately harms you, not the course staff.

We have repeatedly observed this pattern in our own first-hand teaching experience with previous cohorts: students who use LLMs excessively or uncritically tend to perform significantly worse by the end of the course. Students who attempt to solve the assignments on their own may struggle at times, but typically develop much stronger skills, which is reflected in their final grades.

Be aware that misuse of LLMs is often detectable. Inconsistencies between assignments, lack of understanding in discussions, and an inability to explain submitted work are clear indicators. Claiming authorship of work you did not meaningfully produce or understand is considered plagiarism. Never do it!

In short, while LLMs can be a helpful support tool, you should not rely on them to do your thinking for you. The skills you are meant to develop in this course, including problem-solving, analytical reasoning, and independent judgment, are precisely the skills you forfeit when you let an LLM take over.

What to do today?

Find a coauthor and start working on the draft paper very soon.

Other Considerations:

A great website with many R code examples is the UCLA Stat Consulting Site. Another good site that introduces R to SPSS or Stata users is Quick-R. The standard site to search for R (code, problems etc.) on the internet is Stack Overflow.

Within your assignments, you will be required to write mathematical expressions and formulas. A great typesetting software package to do that is LaTeX. Rstudio supports LaTeX-style math formulas and we strongly encourage you to make use of it. You can find an introduction to mathematical expressions in LaTeX here and a list of mathematical symbols here.

Mental Health and Wellness

We find ourselves in unprecedented times, and the disruptions, challenges, and stressors have affected everyone, often straining both academic performance and overall well-being. If you encounter substantial stress, worry, alterations in mood, or difficulties with eating or sleeping this semester -— whether due to this course, others, or external factors -— please feel free to reach out to any of the course instructors for a discussion. Support during challenging times is beneficial for everyone. We are not only willing to lend an understanding ear and make accommodations with deadlines as necessary, but we can also guide you to additional support resources available on campus.

Accessibility and Accommodations

It is our goal that this class is an accessible and welcoming experience for all students, including those with disabilities that may impact learning in this class. If you anticipate or experience academic barriers based on your disability (including mental health, chronic or temporary medical conditions), please let us know immediately so that we can privately discuss options. After registration, make arrangements with us as soon as possible to discuss your accommodations so that they may be implemented in a timely fashion.

Detailed Course Outline:

Week 1 (11 February 2026): Introduction. OLS Recap.

> Cunningham, Scott. *Causal Inference. The Mixtape.* Chapter 2:Probability an Regression Review. https://mixtape.scunning.com/02-probability_and_regression

Week 2 (12 February 2026): OLS in Matrix Form.

> Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage. Chapter 1-2.
>
> Wooldridge, Jeffrey, M. 2009. *Introductory Econometrics. A Modern Approach.* Appendix D & E.

Work through Wooldridge's appendix D first, because it will be heavily used in appendix E. Try to test yourself by doing some concrete examples (2×2- or 2×3-matrices are totally fine) to make sure you understand what's going on. Moreover, I suggest to closely read section E.1 – E.3 and skim the rest of this appendix. Also, in order review the linear model in matrix form, take a look at Scott Long's chapter 1–2, in particularly sections 2.1–2.5.

Homework 1 will be assigned.

Week 3 (25 February 2026): OLS in Matrix Form and Probability Theory

> King, Gary. 1989. *Unifying Political Methodology.* Ann Arbor: University of Michigan Press. Chapter $1 + 3$.
>
> Moore, Will H., and David A. Siegel. 2013. *A Mathematics Course for Political and Social Research.* Ann Arbor: Princeton University Press. Chapter $9 - 11$.
>
> Fox, John. 2007. *Applied Regression, Generalized Linear Models, and Related Methods, 2^{nd} Edition, Thousand Oaks, CA: Sage. Appendix $B + D$.*

We will wrap-up our discussion of the linear model in matrix form and do some applications with probability distributions to get more familiarity with them. The core reading will be chapter 1 and 3 from King (1989). Furthermore, a nice overview about probability theory and particular probability distributions (which we also covered last semester) is provided by Moore and Segal (2013).

Note for further reading I also provide John Fox's appendix in its entirety. If you want to read more about linear algebra take a look at his treatment in Appendix B. It includes nice examples and graphs that provide some more intuition. Also check out sections D.1–D.5 if you prefer additional reading on probability distributions.

Homework 1 is due and Homework 2 will be assigned.

Week 4 (4 March 2026): A first peek at Maximum Likelihood

King, Gary. 1989. *Unifying Political Methodology.* Ann Arbor: University of Michigan Press. Chapter 4-4.3.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage. Chapter 2.6.

We will finally start with an introduction of the likelihood theory of inference. Please read chapter 4 (only until section 4.3) of King's UPM book. For a quick peek at MLE I'd like you to refer to Long's chapter 2.6.

Homework 2 is due.

Week 5 (11 March 2026): Maximum Likelihood Estimation and Heteroskedastic Regression

Eliason, Scott R. 1993 *Maximum Likelihood Estimation: Logic and Practice.* Newbury Park: Sage. Chapter 1-4.

Franklin, Charles H. 1991. "Eschewing Obfuscation? Campaigns and the Perception of Senate Incumbents". *American Political Science Review* 85(4): 1193–1214.

Golder, Matt, and Gabriella Lloyd. 2014. "Re-Evaluating the Relationship between Electoral Rules and Ideological Congruence." *European Journal of Political Research* 53(1): 200–212.

King, Gary. 1989. *Unifying Political Methodology.* Ann Arbor: University of Michigan Press. Chapter 4.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage. Chapter 3.6.1 – 3.6.2.

Make sure you closely (re)-read the entire King's UPM, chapter 4. For those of you who appreciate a slightly different take on MLE take a look at Eliason (1993). Please also read a short section in Long (1997) chapter 3.6.1 and 3.6.2 in order to get a sense of how to actually estimate standard errors using maximum likelihood. For an nice application on how to set-up a heteroskedastic regression model take a look at the "classic" Franklin (1991) paper. Alternatively, another interesting application of a heteroskedastic regression model is found in Golder and Lloyd (2014).

Homework 3 will be assigned.

Week 6 (18 March 2026): Models for Binary Dependent Variables & Model Fit

Esarey, Justin, and Andrew Pierce. 2012. "Assessing Fit Quality and Testing for Misspecification in Binary-Dependent Variable Models." *Political Analysis* 20(4), 480–500.

Greenhill, Brian, Michael D. Ward, and Audrey Sacks. 2011 "The Separation

Plot: A New Visual Method for Evaluating the Fit of Binary Models." *American Journal of Political Science* 55(4), 991–1002.

King, Gary. 1989. *Unifying Political Methodology.* Ann Arbor: University of Michigan Press. Chapter 5.1–5.3.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage. Chapter 3.

Giger, Nathalie, Simon Lanz and Catherine de Vries. 2020. "The motivational basis of constituency work: how intrinsic and extrinsic motivations interact". *Political Science Research and Methods*, 8(3), 493–508.

Neunhoeffer, Marcel, and Sternberg, Sebastian (2019). "How Cross-Validation Can Go Wrong and What to Do About It". *Political Analysis*, 27(1), 101–106

We will take a closer look at models for dichotomous dependent variables. Please take a close look at Chapter 5.1-5.3 of King (1989) and Chapter 3 in Long (1997). Also skim the Esarey/Pierce (2012) as well as the "separation plot" paper of Greenhill et al. (2011) for new strategies of how you could evaluate your model in terms of model fit. For an applied example of logistic regression models, have a look at the recent paper by Giger et al. (2020). They also present nice quantities of interest which we will cover next week. Finally, former students of our AQM class now contribute to this literature themselves! Take a look at Marcel and Sebastian's work on *cross-validation* to assess model fit. They both took AQM a few years ago as well.

Homework 3 is due.

Week 7 (25 March 2026): Interpretation and Simulation

Abrajano, Marisa A., R. Michael Alvarez, and Jonathan Nagler. 2008. "The Hispanic Vote in the 2004 Presidential Election: Insecurity and Moral Concerns." *The Journal of Politics* 70(2): 368—82.

Hanmer, Michael J., and Kerem Ozan Kalkan. 2013. "Behind the Curve: Clarifying the Best Approach to Calculating Predicted Probabilities and Marginal Effects from Limited Dependent Variable Models." *American Journal of Political Science* 57(1), 263–277.

King, Gary, Michael Tomz and Jason Wittenberg. 2000 "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44(2): 347–361.

Please read closely King et al (2000). If you have read it before - read it again! I read it several times myself after I digested it first (it was a prominent working paper at that time). This piece is definitely on the "Top 10" list of papers every MA student has to digest. Also take a look at the Hanmer & Kalkan paper to understand the difference between average-case and observed-value approaches. Which one do you prefer? Finally, to see an example of how much substance you can convey through simulating quantities of interest, take a look at Abrajano et al. instructive article.

Homework 4 will be assigned.

Easter Recess: No class on 1 April 2026 & 8 April 2026!

Week 8 (15 April 2026): Ordered Choice Models & How to write a publishable Paper

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage. Chapter 5.

Jackman, Simon. 2000. *Models for Ordered Outcomes.* Lecture Notes.

In the first part we will discuss ordered choice models. Focus on Long's chapter 5. For a discussion of nifty applications of the ordered choice model, take a look at Jackman's lecture notes. In the second part of today's lecture I will provide you with some strategies on how you can write a publishable paper in this class.

Week 9 (22 April 2026): Multinomial Choice Models

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage. Chapter 6.

Gschwend, Thomas, and Leuffen, Dirk. 2005. "Divided We Stand – Unified We Govern? Cohabitation and Regime Voting in the 2002 French Elections." *British Journal of Political Science* 35(4), 691–712.

We will cover "multinomial choice models". Please take a look at Long's Chapter 6. In case you wanna see an application of this model, take a look and Gschwend & Leuffen's *BJPolS*.

Homework 4 is due.

Week 10 (29 April 2026): Conditional Logit Model

Alvarez, R. Michael, and Jonathan Nagler. 1998 "When Politics and Models Collide: Estimating Models of Multiparty Elections." *American Journal of Political Science* 42(1): 55-96.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage. Chapter 6.

We will discuss conditional logit models. Please closely read Alvarez and Nagler's 1998 AJPS piece and re-read Long's Chapter 6.

Homework 5 will be assigned.

Week 11 (6 May 2026): Selection Bias and Multi-Equation Models

Dellmuth, Lisa Maria, and Michael F. Stoffel. 2012 "Distributive politics and intergovernmental transfers: The local allocation of European Union structural funds." *European Union Politics* 13(3): 413-433.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage. Chapter 7.

Timpone, Richard J. 1998. "Structure, Behavior, and Voter Turnout in the United States". *American Political Science Review* 92(1): 145-158.

We will discuss selection bias models, particularly tobit and heckman models. For tobit models please closely read Long's chapter 7 and for a nice application take a look at the recent award-winning Dellmuth and Stoffel paper. In order to better understand the Heckman model browse through Rich Timpone's APSR piece.

Homework 6 will be assigned.

Week 12 (13 May 2026): Multi-level Models

Bell, Andrew, and Kelvyn Jones. 2015. "Explaining Fixed Effects: Random Effects modeling of Time-Series Cross-Sectional and Panel Data". *Political Science Research and Methods* 3(1): 133-–153.

Clark, Tom S., and Drew A. Linzer. 2015. "Should I Use Fixed or Random Effects?" *Political Science Research and Methods* 3(2): 399-–408.

Gelman, Andrew, and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press. Chapter 11–13

Plümper, Thomas, and Troeger Vera E. 2019. "Not so Harmless After All: The Fixed-Effects Model". *Political Analysis* 27(1): 21-–45.

Steenbergen, Marco R., and Bradford S. Jones. 2002. "Modeling Multilevel Data Structures." *American Journal of Political Science* 46(1): 218-–37.

Make sure to closely read all three chapters of Gelman & Hill and skim the Clark & Linzer article. We will discuss in class as an application the model of Steenbergen & Jones. For those who deeply care about TSCS data, take also a close look at the Bell & Jones (2015) paper. Additionally, Plümper & Troeger (2019) provide a nice article about how fixed effects regression can be more biased than simpply pooling models under certain circumstances.

Homework 5 is due

Week 13 (20 May 2026): Baby Bayes – a primer

Martin, Andrew D. 2008. "Bayesian Analysis." In *Oxford Handbook of Political Methodology*, eds. Janet M. Box-Steffensmeier, Henry E. Brady, and David

Collier. Oxford: Oxford University Press, 494–510.

Stegmueller, Daniel. 2013. "How Many Countries for Multilevel Modeling? A Comparison of Frequentist and Bayesian Approaches." *American Journal of Political Science* 57(3): 748—61.

Finally, we will cover some ground in terms of Bayesian Analysis. Please read the intro article by Andrew Martin. People interested in Bayesian multi-level models should also consult the now "classic" Stegmueller piece. For those who are interested in how to apply models you have learnt so far from a Bayesian perspective, please consider Statistical Rethinking by Richard McElreath. It is a nice introduction to the Bayesian philosophy and models. The book is also accompanied by free Youtube lectures.

Homework 6 is due

Week 14 (27 May 2026): Student Presentations

In order to provide you with feedback on your final papers we will have short (< 5 min.) presentations of your hypothesis and the key results of your paper. Please email me your presentations until **8am** that day.