

Prof. Dr. Christoph Hanck
Lehrstuhl für Ökonometrie
Universität Duisburg-Essen

Abschlussbericht zum Projekt

Reproducible Research in der ökonometrischen Grundausbildung

Gefördert durch den Stifterverband und das Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen im Rahmen der Fellowship für Digitale Hochschullehre 2016

1. Motivation, Problemstellung und Zielsetzung

In Zeiten von Digitalisierung und Big Data wird funktionales Wissen in computergestützter Statistik sowie die Fähigkeit, entsprechende Ergebnisse in geeigneter Weise zu kommunizieren, immer bedeutender. Daher wird bereits in einigen Lehrveranstaltungen des Bachelorstudiengangs *Volkswirtschaftslehre* der Universität Duisburg-Essen neben theoretischem Wissen auch Kompetenz im Umgang mit *R* (R Core Team, 2018) vermittelt. *R* ist eine Open Source Softwareumgebung für statistische Programmierung, die sich als maßgeblich für die Qualität empirischer Seminar- und Abschlussarbeiten erweist und zunehmend auch von Unternehmen eingesetzt wird. Der Erwerb dieser methodischen Schlüsselkompetenzen ist somit ein wesentliches Qualifikationsziel dieser Lehrveranstaltungen.

Insbesondere in der Veranstaltung *Einführung in die Ökonometrie* kommt der statistischen Programmierung mit *R* eine Schlüsselrolle zu: Die Studierenden erlernen für die empirische Wirtschaftsforschung relevante Methoden zur Schätzung kausaler Effekte anhand nichtexperimenteller Daten und deren Anwendung mit *R*. Aufgrund großer Teilnehmeranzahlen und oftmals fehlenden Vorkenntnissen in Programmierung seitens der Studierenden ist so bspw. eine didaktisch sinnvolle Integration von Kleingruppenarbeit in die Lehrveranstaltungen nicht umsetzbar, sodass überwiegend ineffizienter Frontalunterricht erfolgte und *R*-Kompetenzen nicht im wünschenswerten Umfang vermittelt werden konnten. Mangels eines auf die empirischen Inhalte der Veranstaltung ausgerichteten E-Learning-Angebots ist ein klassisches Selbststudium für die Studierenden mit vergleichsweise großem Aufwand verbunden, zumal es hinsichtlich *R*-Programmierung kaum einsteigerfreundliche Literatur mit ökonometrischem Bezug gibt und die Lernkurve beim Einstieg in die Programmierung erfahrungsgemäß flach ist.

Ziel der im Rahmen des Projekts eingeführten Lehrinnovation ist daher die Stärkung studentischer Kompetenzen in der Nutzung von *R*. Hierbei wird autodidaktisches Engagement der Studierenden durch einen Online-Companion gefördert, der zudem Synergie-Effekte beim Erarbeiten mathematisch-formaler und empirisch orientierter Lehrinhalte gewährleistet. Die Implementierung dieses E-Learning-Angebots ermöglicht eine didaktische Neuorganisation des Moduls *Einführung in die Ökonometrie*, wobei bestimmte Lern- und Lehrelemente gezielt in den inverted-classroom verlagert bzw. in die interaktiven Präsenzübungen integriert werden. Die Nutzung der Lehrinnovation durch die Studierenden wird durch ein Bonuspunktesystem gezielt unterstützt.

2. Über die implementierte Lehrinnovation

Kern der Lehrinnovation *Introduction to Econometrics with R* ist ein interaktiver, speziell für das Modul *Einführung in die Ökonometrie* entwickelter Online-Companion. Dieser fungiert einerseits als Bindeglied zwischen den theoretischen Lehrinhalten klassischer Präsenzveranstaltungen (Vorlesung und Übung) und den in *R*-Übungen vermittelten empirischen Kompetenzen und andererseits als Medium für Blended Learning. Der Online-Companion ist wesentlich an dem in den Lehrveranstaltungen verwendeten Lehrbuch *Introduction to Econometrics* (Stock & Watson, 2015) ausgerichtet und zeigt, wie zentrale Ergebnisse empirischer Studien mit *R* reproduziert werden können und wie sich grundlegende Resultate der mathematischen Statistik mit Hilfe von Simulationsstudien validieren lassen. Hierfür wurde die Form eines reproduzierbaren Forschungsberichts gewählt, welcher die Studierenden mit Hilfe von kommentierten Code-Beispielen und feedback-gestützten Programmieraufgaben anleitet, eigenständig programmierbezogene Lerninhalte zu erarbeiten und diese auf neue Fragestellungen anzuwenden. Durch die Einbindung der interaktiven *R*-Aufgaben wird ein kontinuierlicher Transfer zwischen der mathematischen Formulierung von Modellen und deren softwareseitiger Umsetzung erreicht. Um den mathematisch-formalen Anteil des Vermittelten gering zu halten, werden zudem interaktive Visualisierungen eingesetzt, welche komplexe Konzepte in Statistik und Wahrscheinlichkeitstheorie anschaulich darstellen (s. Abbildung 1).

2.1. Softwareseitige Umsetzung

Der Companion ist im Wesentlichen eine modular erweiterbare Webseite, die den Studierenden über das Learning Management System *Moodle* bereitgestellt wird. Die Seite wird überwiegend mit *R* und der Auszeichnungssprache *Rmarkdown* programmiert und durch das auf *Knitr* (Xie, 2018a) und *Pandoc* basierende Paket *bookdown* (Xie, 2018b), generiert.¹ Wesentliche Erweiterungen sind die oben genannten Visualisierungen, welche auf der frei verfügbaren JavaScript Bibliothek *D3.js* basieren sowie feedbackgestützte Programmieraufgaben, die mit *datacamp-light.js* realisiert werden. Diese ebenfalls frei nutzbare Bibliothek der

¹ Der Companion kann unter <https://www.econometrics-with-r.org> eingesehen werden.



Abbildung 1: Interaktive Visualisierung einer Simulationsstudie im Online-Companion

Plattform *DataCamp* ermöglicht die Integration einer vollfunktionsfähigen *R*-Konsole mit übersichtlicher Benutzeroberfläche innerhalb der Lektionen (s. Abbildung 2).² So werden die Studierenden in die Lage versetzt, neu erlernte *R*-Befehle direkt auszuführen oder sich mit den beschriebenen Funktionen vertraut zu machen. Weiterhin erlaubt die Integration der *R*-Funktionssammlung *testwhat* den Einsatz von submission correctness tests (SCTs). Die SCTs ermöglichen ein individuelles, dynamisches Feedback. So werden bspw. bei Syntaxfehlern automatisch die betreffenden Stellen im Code markiert und entsprechende Warnungen ausgegeben. Außerdem können automatisch Hinweise generiert werden, um dem Nutzer eingabeabhängig eine bestimmte Vorgehensweise zu empfehlen.

Um die Umsetzung von Programmieranwendungen aus dem Online-Companion mit integrierten Entwicklungsumgebungen wie *RStudio* zu erleichtern, wurde außerdem eine *R*-Paket entwickelt, welches es den Studierenden erlaubt, Inhalte des Online-Companions offline verfügbar zu machen. So können bspw. sämtliche Datensätze und Code-Beispiele einer Lektion automatisch in gängige Formate konvertiert und heruntergeladen werden.³

Der Quellcode des Companions ist auf der Plattform *Github* veröffentlicht und steht Dritten zur nicht-kommerziellen Nutzung frei zur Verfügung.⁴ Hinsichtlich der Verstetigung der Lehrinnovation ist dieses Setup besonders nützlich, da *R*-basierte Lerninhalte mit geringem Aufwand aktualisiert bzw. hinzugefügt werden können. Hierzu sind dank der Nutzung o. g. Softwarepakete neben *R* und *Rmarkdown* keinerlei Kenntnisse in der Programmierung von Webseiten erforderlich, sodass auch engagierte Studierende konkrete Verbesserungsvorschläge unterbreiten und somit zum Companion beitragen können (s. Kapitel 4).

² Siehe <https://www.datacamp.com/>

³ Das *R*-Paket ist auch für Dritte in einem [Repository](#) auf der Plattform *Github* frei verfügbar.

⁴ Der Quellcode zum Companion ist [hier](#) einsehbar.

3. Restrukturierung der Lehrveranstaltung

Vor dem Hintergrund einer im Wintersemester 2017/2018 umgesetzten didaktischen Neuorganisation der Bachelorveranstaltung *Einführung in die Ökonometrie* und des erstmaligen Einsatzes des Online-Companions im Lehrbetrieb wurde insbesondere die begleitende Übung zur Vorlesung umstrukturiert. Diese Umstrukturierung wird nachfolgend erläutert und umfasst die folgenden Maßnahmen:

1. *Konzeption und Erstellung eines Katalogs interaktiver R-Programmieraufgaben*

Mit Unterstützung des wissenschaftlichen Mitarbeiters im Projekt erfolgte die Konzeption und Erstellung eines Katalogs von begleitenden interaktiven *R*-Programmieraufgaben. Die Aufgaben wurden zu Beginn in Form eines dedizierten Kurses auf der Plattform *DataCamp* bereitgestellt und im weiteren Verlauf des Projekts direkt in den Online-Companion integriert. Die Aufgaben können somit direkt im Online-Companion von den Studierenden bearbeitet werden, bspw. in Form separater Kapitel oder an geeigneter Stelle im Fließtext (vgl. Abbildung 2).

2. *Obligatorische Lektionen mit einführenden R-Programmieraufgaben*

Zur Vorbereitung auf die Veranstaltung konnten so bereits vor Beginn der Präsenzübungen grundlegende Funktion von *R* von den Studierenden im Selbststudium erlernt werden. Diese Verlagerung elementarer Lehrinhalte ermöglichte einen zügigen Einstieg in die Programmierung ökonomischer Anwendungen im Übungsbetrieb.

3. *Integration von Online-Übungsaufgaben in die regulären Übungen.*

Die in 1. genannten Aufgaben werden zur gezielten Vor- und Nachbereitung der Einzeltermine sowie für die Vorbereitung auf Testate, Hausaufgaben und die Abschlussklausur genutzt.⁵ Insbesondere das automatische Feedbacksystem der interaktiven *R*-Übungsaufgaben in Kombination mit dokumentierten Code-Beispielen hat sich hier als hilfreich erwiesen, da so eine Verlagerung von Kursbestandteilen in den inverted classroom ermöglicht wird, die bisher zeitaufwändig und ausschließlich in Form von Frontalunterricht in den gesonderten Computer-Übungen unterrichtet wurden. Dies konnte für eine verstärkte Einbindung der Teilnehmer in die Präsenzveranstaltungen genutzt werden, etwa für das gemeinsame Lösen von *R*-Aufgabenstellungen oder die Präsentation von Programmier-Hausaufgaben durch die

⁵ 90% aller Teilnehmer der Abschlussklausuren im Wintersemester 2017/2018 haben dieses Angebot wahrgenommen (s. Punkt 4).

8. Regression on a Dummy Variable III

In this exercise, you have to visualize some of the results from the dummy regression model

$$\hat{Y}_i = -0.66 + 1.43 \times D_i$$

estimated in the previous exercise.

A data frame `DF` with 100 observations on `X` and the binary variable `dummy` as well as the model object `dummy_mod` from the previous exercise are available in your working environment.

Instructions:

- Start by drawing a visually appealing plot of the observations on `Y` and `D` based on the code chunk provided in Script.R. Replace the `???` by the correct expressions!
- Add the regression line to the plot.

script.R	R Console
<pre> 1 # Replace the `???' by the correct values 2 plot(x = ??? , y = ???, 3 pch = 20, 4 cex = 1, 5 col = "Steelblue", 6 xlab = expression(D[i]), ylab = "Test Score", 7 main = "Dummy Regression" 8) 9 10 # add the regression line </pre>	<pre> > </pre>

Solution
Submit

Abbildung 2: Interaktive R-Übungsaufgabe im Online-Companion

Studierenden. Letzteres sowie das Absolvieren vertiefender R-Lektionen wird mit Bonuspunkten belohnt.

4. Bonuspunktesystem und Prüfungsform

Während der Vorlesungszeit besteht für die Studierenden die Möglichkeit für das Absolvieren von Testaten in R-Programmierung bzw. für Bearbeitung von Hausaufgaben Bonuspunkte für die Abschlussklausur zu erwerben. Ersteres wird ebenfalls mit webbasierten R-Aufgaben realisiert und von der Mehrheit der Studierenden in Anspruch genommen.⁶ Diese Hausaufgaben kombinieren Theorie- mit fortgeschrittenen Programmieraufgaben und sind unter Zuhilfenahme entsprechender Lektionen des Online-Companions einfacher zu lösen. So wird ein weiterer Anreiz geschaffen, sich intensiver mit statistischer Programmierung zu befassen.

Die abschließende Modulklausur ist an dem neuen didaktischen Konzept der Lehrveranstaltungen ausgerichtet und prüft Theorie- und R-Kenntnisse in etwa gleichem Umfang.

4. Evaluation und Verstetigung der Lehrinnovation

Der erstmalige Einsatz der Lehrinnovation und das adjustierte Konzept der Präsenzübungen im Wintersemester 2017/2018 wurde durch die Studierenden evaluiert. Hierbei wurde insbesondere das Arrangement von R-Anwendungen und interaktiven Übungsaufgaben, der einführende R-Kurs und der bereitgestellte Aufgabenkatalog positiv bewertet. Eine Mehrheit der Studierenden befürwortet zudem eine stärkere Integration des Online-Companions in den Übungsbetrieb. Ein ähnliches Stimmungsbild ergab eine vergleichbare

⁶ 60% aller Kursteilnehmer im Wintersemester 2017/2018 haben mindestens eines der beiden Testate absolviert.

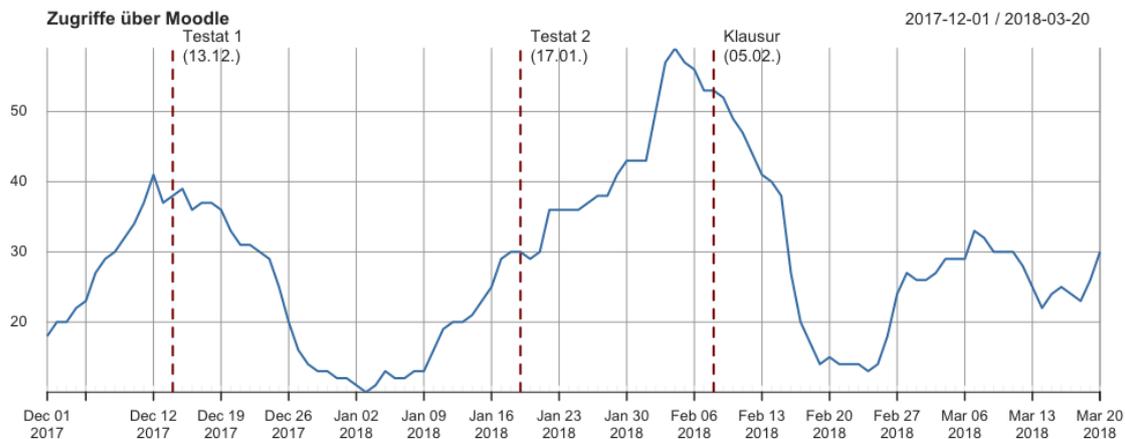


Abbildung 3: Nutzerzugriffe im Zeitverlauf am Beispiel des Wintersemesters 2017/2018

Evaluation im laufenden Wintersemester 2018/2019. 70% der befragten Studierenden geben, an die neuen Materialien häufig zu nutzen und sehen darin eine Bereicherung ihres Selbststudiums.

Nutzungsstatistiken zeigen, dass die Studierenden in den Zeiträumen unmittelbar vor den Präsenztastaten, während der Bearbeitungszeiten der Hausaufgaben, sowie vor der Abschlussklausur verstärkt auf den Online-Companion zugreifen und das Angebot demnach auch zur Prüfungsvorbereitung nutzen (s. Abbildung 3). Weiterhin haben 90% der KlausurteilnehmerInnen den in Umstrukturierungsmaßnahme 2 (s. Kapitel 3, Punkt 2) beschriebenen obligatorischen *R*-Einführungskurs bereits vor Beginn der Präsenzübung im Selbststudium absolviert, was Inhalte einer bisher zweistündigen Einführung in der Übung redundant macht und so wie gewünscht einen schnelleren Einstieg in vertiefende Inhalte ermöglichte.

Wenngleich die Lehrinnovation insgesamt als sinnvolle Ergänzung zu den bisher genutzten Lern- und Lehrmaterialien wahrgenommen wird, äußern sich die Studierenden in beiden Evaluationen kontrovers zu den Bewertungskriterien Sprache, Ausführlichkeit und Quantität der neuen Inhalte. So wird teilweise die Bilingualität der Lehrmaterialien bemängelt und die (englischen) Erläuterungen der Programmieranwendungen im Online-Companion teils als Hürde wahrgenommen.⁷ Wenngleich der Detailgrad dieser Erläuterungen meist positiv aufgenommen wird, findet sich eine Teilmenge von Studierenden, welche die Lernziele und Klausuranforderungen hinsichtlich *R* als nicht hinreichend deutlich erachtet und vom Umfang des Online-Companions eingeschüchtert ist.⁸ In dieser Hinsicht besteht ein Anpassungsbedarf. Im Besonderen betrifft dies die Konzeption von in den Präsenzveranstaltungen diskutierten Übungsaufgaben, die sich unmittelbar auf den Online-Companion beziehen.

⁷ Die Inhalte des Online-Companions sind ausschließlich englischsprachig, die verknüpften klassischen Übungsblätter und Folien hingegen deutsch. Die Modulklausur wird ebenfalls auf Deutsch abgelegt.

⁸ Der Online-Companion umfasst sämtliche Kapitel von Stock & Watson (2015), die jedoch in den Lehrveranstaltungen nicht alle bzw. nicht vollständig behandelt werden.

Weiterhin besteht aus Sicht der Studierenden und des Übungsleiters noch Verbesserungspotential bei der proaktiven Förderung eigenständiger Arbeit in den Präsenzübungen. Als Reaktion hierauf wird im Wintersemester 2018/2019 der Einsatz studentischer Hilfskräfte zur Unterstützung im Übungsbetrieb erprobt.

Ein nicht intendierter, ambivalenter Effekt der Umstrukturierungen und der Lehrinnovation ist eine starke Fokussierung der Studierenden auf den Übungsbetrieb, was sich mutmaßlich negativ auf die Teilnehmeranzahl in der Vorlesung auswirkt. Dieses Verhalten reflektiert die Wahrnehmung der zu den empirischen Kompetenzen hin verlagerten Prüfungsanforderungen seitens der Studierenden (s. Kapitel 3, Punkt 4). Beurteilt anhand von Prüfungsleistungen wie Hausaufgaben und der Abschlussklausur impliziert dies jedoch nicht, dass die Studierenden sich weniger intensiv auf mathematische Prüfungsinhalte vorbereiten. So wird der Online-Companion im Selbststudium überwiegend wie intendiert als Komplement zum Vorlesungsskript eingesetzt.

Unter Berücksichtigung der Evaluationsergebnisse wird der Online-Companion im Verstetigungsprozess fortlaufend verbessert. Dies gelingt dank der genutzten *R*-basierten Entwicklungsumgebung und Software auch mit unmittelbarer Unterstützung von studentischen Hilfskräften, deren Hinweise und Vorschläge sich regelmäßig als hilfreich erweisen. Außerdem wird der Companion auch jenseits der Veranstaltung *Einführung in die Ökonometrie* eingesetzt, etwa zur Unterstützung der Studierenden bei der Erstellung empirischer Seminar- und Bachelorarbeiten mit *R* oder als Zusatzangebot in freiwilligen Propädeutika. Hiervon profitieren Studierende in der Einstiegsphase der Masterstudiengänge *Energy and Finance*, *Gesundheitsökonomik* und *Volkswirtschaftslehre* mit wenig Vorkenntnissen in *R*, da die Pflicht- und Wahlpflichtcurricula dieser Studiengänge ebenfalls Lehrveranstaltungen mit einem Fokus auf Ökonometrie und statistische Programmierung mit *R* beinhalten.

Im Zuge einer Neuauflage des der Veranstaltung zugrundeliegenden Lehrbuchs von Stock & Watson (2015) wird im Jahr 2019 eine Aktualisierung und eine entsprechende Erweiterung des Online-Companions erfolgen. Dies wird die Lehrinnovation auch weiterhin für externe Studierende bzw. Lehrende gleichgelagerter Lehrveranstaltungen attraktiv machen.

5. Literaturangaben

James H. Stock und Mark W. Watson (2015), "Introduction to Econometrics", Pearson, 3. aktualisierte Auflage.

R Core Team (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Yihui Xie (2018a). knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.21

Yihui Xie (2018b). bookdown: Authoring Books and Technical Documents with *RMarkdown*. R package version 0.7.24.