



KLEINRÄUMIGE EXTRAPOLATION VON UMFRAGEDATEN

NAMEN:	ALEXANDER LANGE, KAI HUSMANN
MATR. NR.:	21426614
STUDIENGANG:	ANGEWANDTE STATISTIK
MAIL:	ALEXANDER.LANGE@UNI-GOETTINGEN.DE
KURS:	STATISTISCHES PRAKTIKUM
KURSLEITER:	PROF.DR. THOMAS KNEIB
LEHRSTUHL:	STATISTIK
FAKULTÄT:	WIRTSCHAFTSWISSENSCHAFTEN
DATE:	2. SEPTEMBER 2016

INHALTSVERZEICHNIS

1	Einleitung	1
2	Daten	2
2.1	Deskriptive Statistik	2
2.2	Räumliche Effekte	3
3	Methodik	3
3.1	Modell	3
3.2	Modellwahl	3
3.3	Evaluierung	3
4	Ergebnisse	3
5	Ergebnisse	3
	Literatur	4
	Anhang	6

ABBILDUNGSVERZEICHNIS

1	Endogene Variablen	3
---	------------------------------	---

TABELLENVERZEICHNIS

1	Datensatz	2
2	Grundgesamtheit Bürgerumfrage	6
3	Grundgesamtheit Zensus	6

1 EINLEITUNG

2 DATEN

2.1 DESKRIPTIVE STATISTIK

In diesem Kapitel soll der Stichprobendatensatz vorgestellt werden, sowie ein kurzer Einblick in die beiden Datensätze zur Grundgesamtheit von Stuttgart gegeben werden. Anhand des Stichprobendatensatzes soll im Verlauf der Arbeit das Modell zur Extrapolation geschätzt werden und mit Hilfe der Datensätze zur Grundgesamtheit die Häufigkeiten der abhängigen Variablen prognostiziert werden.

Tabelle 1 gibt einen Überblick über den Inhalt der Stichprobe. Wie zu erkennen ist, handelt es sich dabei um sozioökonomische Variablen, welche sich auf nominalem oder ordinalen Skalenniveau befinden. Zudem enthält der Datensatz drei verschiedene räumliche Informationen zu den Beobachtungen, auf die im nächsten Kapitel ausführlicher eingegangen wird.

TABELLE 1: DATENSATZ

Anzahl Beobachtungen Stichprobe: 3.143

Variable	Anzahl Klassen	Modellierung
Bewertung Wohngegend	6	Geordnet Kategorial
Meinung Stuttgart 21	6	Geordnet Kategorial
Personenanzahl im Haushalt	5	Nicht Parametrisch
Monatliches Netto Haushaltseinkommen	6	Nicht Parametrisch
Altersklasse Befragter	6	Nicht Parametrisch
Geschlecht	2	Parametrisch
Familienstand	4	Parametrisch
Nationalität	2	Parametrisch
Stadtbezirk	23	Diskret Räumlich
Stadtteil	142	Diskret Räumlich
Gauß-Krüger		Stetig Räumlich

Die rechte Spalte zeigt an, wie die einzelnen Variablen in das zu schätzende Modell einfließen sollen. Dabei ist zu erkennen, dass nominal skalierte Variablen, wie z.B. Nationalität als parametrisch und ordinal skalierte Variablen wie z.B. die Altersklasse der Befragten als nicht parametrisch modelliert werden sollen. Es ist plausibel den Effekt von ordinalen Variablen als nicht linear anzunehmen, weil [...]. Die beiden Datensätze zur Grundgesamtheit stammen aus einer Bürgerumfrage mit 470.190 Beobachtungen und dem Zensus mit 380.238 Beobachtungen. Im Verhältnis zu den Grundgesamtheiten dieser Größenordnung sind 3143 Beobachtungen in der Stichprobe relativ gering, was eine gewisse Unsicherheit für die Extrapolation mit sich bringt [...].

Weiterhin ist zu beachten, dass Informationen zu dem monatlichen Netto Haushaltseinkommen in beiden Grundgesamtheiten fehlen und somit die Variable nicht für die Prognose verwendet werden kann. Auch war eine denkbare Erstellung von Proxy-Variablen nicht möglich. Eine genaue Auflistung der enthaltenen Variablen aus den Grundgesamtheiten sind im Anhang verfügbar. Die Arbeit zielt darauf ab, die Meinung der Befragten zu dem Projekt Stuttgart 21 und die Zufriedenheit mit der Wohngegend der Befragten auf die Grundgesamtheit zu extrapolieren. Daher ist es sinnvoll die Ausprägungen dieser Variablen genauer zu untersuchen.

Dazu wurde Abbildung 1 erstellt. Sie zeigt die Häufigkeiten der einzelnen Ausprägungen der beiden endogenen Variablen. Wie schon zuvor aus Tabelle 1 ersichtlich, besitzen beide Variablen sechs mögliche Realisationen, wobei die Klasse *Keine Angabe* keine Informationen über die Meinung der Befragten liefert und diese Beobachtungen daher aus dem Datensatz entfernt werden müssen. Damit bleiben fünf mögliche Klassen zur Modellierung übrig.

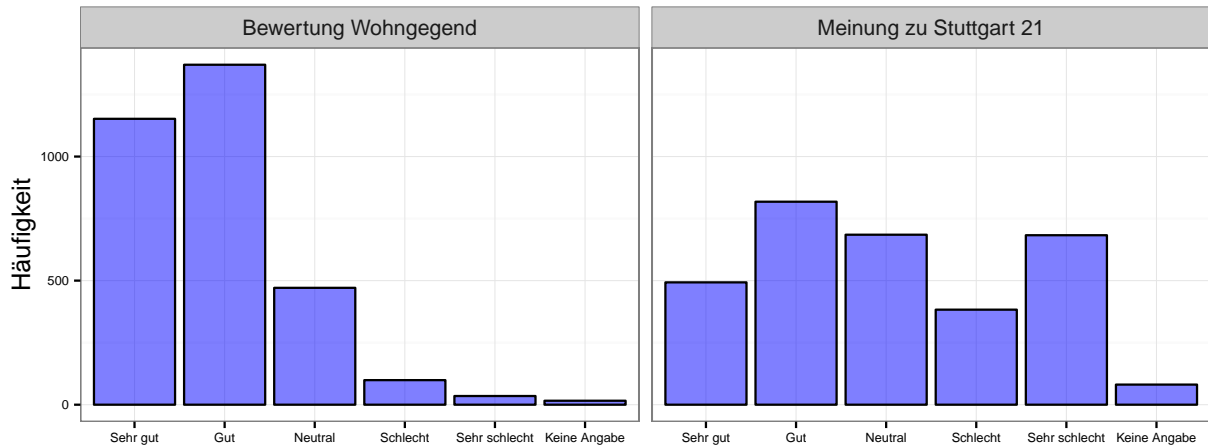


ABBILDUNG 1: ENDOGENE VARIABLEN

Aus Abbildung 1 ist zudem ersichtlich, dass die Verteilungen der Beiden Variablen sehr verschieden sind für die gleichen Antwortmöglichkeiten. Während die meisten befragten Personen ihre Wohngegend mit *Gut* oder *Sehr gut* bewertet haben, sind die Antworten zum Projekt Stuttgart 21 eher gleichmäßig verteilt. Mögliche Ursachen für dieses Phänomen könnten sein, dass [...]. Für die exogen zu modellierenden Variablen sind detailliertere Informationen zu den Häufigkeiten der Ausprägungen im Anhang verfügbar.

2.2 RÄUMLICHE EFFEKTE

3 METHODIK

3.1 MODELL

3.2 MODELLWAHL

3.3 EVALUIERUNG

4 ERGEBNISSE

5 ERGEBNISSE

LITERATUR

- [Bacchiocchi and Fanelli, 2015] Bacchiocchi, E. and Fanelli, L. (2015). Identification in structural vector autoregressive models with structural changes, with an application to us monetary policy. *Oxford Bulletin of Economics and Statistics*, 77(6):761–779.
- [Berg, 2010] Berg, T. O. (2010). *VAR models on the relation between stock prices and the macroeconomy*. PhD thesis, Johann Wolfgang Goethe-University Frankfurt am Main.
- [Box and Draper, 1987] Box, G. E. P. and Draper, N. R. (1987). *Empirical Model-building and Response Surfaces*. Wiley.
- [Hall and Lieberman, 2009] Hall, R. and Lieberman, M. (2009). *Economics: Principles and Applications*. Cengage Learning.
- [Herwartz, 2014] Herwartz, H. (2014). Structural analysis with independent innovations. (208).
- [Herwartz and Plödt, 2014] Herwartz, H. and Plödt, M. (2014). Sign restrictions and statistical identification under volatility breaks – simulation based evidence and an empirical application to monetary policy analysis. (C13-V1).
- [Klaeßling, 2003] Klaeßling, M. (2003). Monetary policy shocks - a nonfundamental look at the data. Working Paper Series 0228, European Central Bank.
- [Lanne and Lütkepohl, 2008a] Lanne, M. and Lütkepohl, H. (2008a). Identifying monetary policy shocks via changes in volatility. *Journal of Money, Credit and Banking*, 40(6):1131–1149.
- [Lanne and Lütkepohl, 2008b] Lanne, M. and Lütkepohl, H. (2008b). A statistical comparison of alternative identification schemes for monetary policy shocks. (ECO2008/23).
- [Lütkepohl, 2005] Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer, Berlin [u.a.].
- [Lütkepohl and Krätzig, 2004] Lütkepohl, H. and Krätzig, M. (2004). *Applied Time Series Econometrics*. Cambridge University Press.
- [Pfaff, 2008a] Pfaff, B. (2008a). *Analysis of Integrated and Cointegrated Time Series with R*. Springer, New York, second edition. ISBN 0-387-27960-1.
- [Pfaff, 2008b] Pfaff, B. (2008b). *VAR, SVAR and SVEC Models: Implementation Within R Package vars*. Journal of Statistical Software 27(4).
- [R Core Team, 2015] R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- [Ravn, 2013] Ravn, S. H. (2013). *Essays on the Interactions between Financial Markets, the Macroeconomy, and Economic Policy*. PhD thesis, University of Copenhagen.
- [Sørensen and Whitta-Jacobsen, 2010] Sørensen, P. and Whitta-Jacobsen, H. (2010). *Introducing Advanced Macroeconomics: Growth and Business Cycles*. Economia e discipline aziendali. McGraw-Hill Higher Education.

- [Spahn, 2012] Spahn, H.-P. (2012). *Geldpolitik: Finanzmärkte, neue Makroökonomie und zinspolitische Strategien*. Vahlen, München.
- [Wickham, 2009] Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- [Wooldridge, 2008] Wooldridge, J. (2008). *Introductory Econometrics: A Modern Approach (with Economic Applications, Data Sets, Student Solutions Manual Printed Access Card)*. South-Western College Pub, 4 edition.

ANHANG

TABELLE 2: GRUNDGESAMTHEIT BÜRGERUMFRAGE

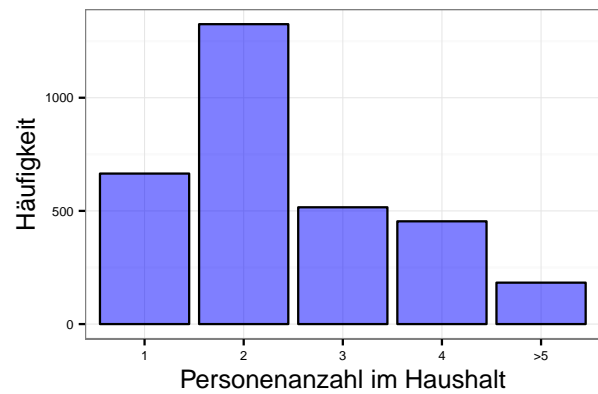
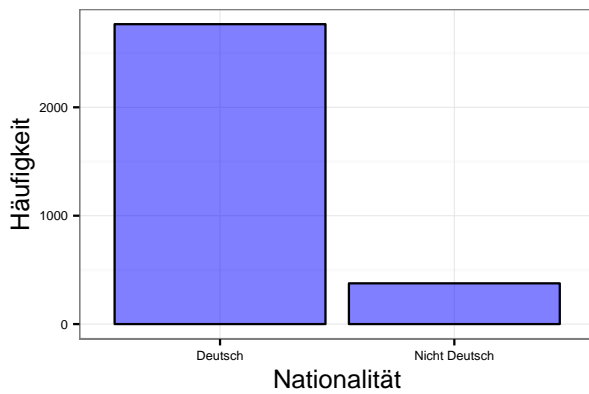
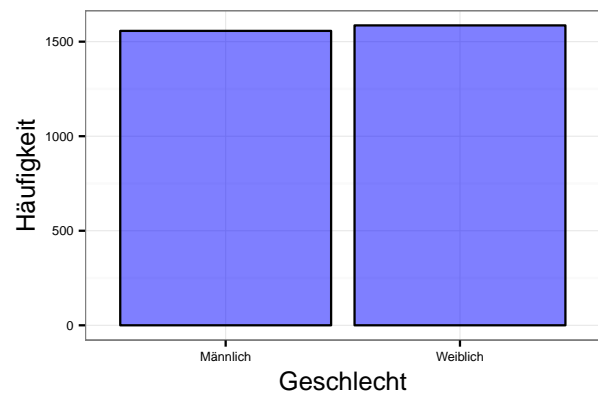
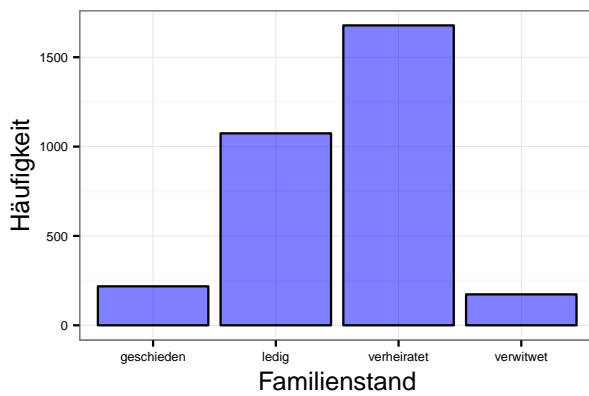
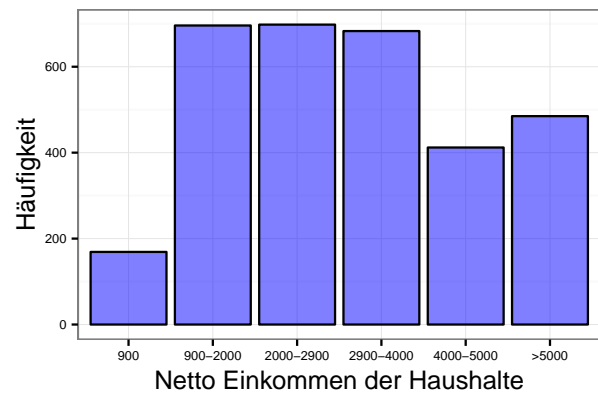
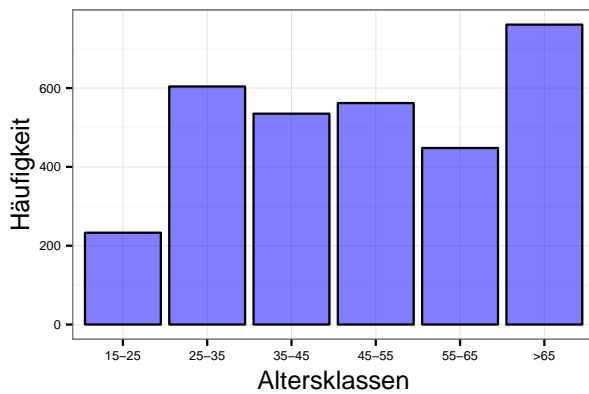
Anzahl Beobachtungen: 470.190

Variable	Skalenniveau	Anzahl Klassen
Altersklasse	Ordinal	14
Geschlecht	Nominal	2
Nationalität	Nominal	2
Familienstand	Nominal	4
Haushaltsgröße	Ordinal	–
Wohndauer	Ordinal	–
ALG II Quote	Ordinal	9
Ein/Zweifamilienhäuser	Ordinal	–
Gauß-Krüger	Stetig Räumlich	

TABELLE 3: GRUNDGESAMTHEIT ZENSUS

Anzahl Beobachtungen: 380.238

Variable	Skalenniveau	Anzahl Klassen
Altersklasse	Ordinal	9
Geschlecht	Nominal	2
Nationalität	Nominal	2
Familienstand	Nominal	4
Haushaltsgröße	Ordinal	–
Wohnfläche	Ordinal	24
Stellung Beruf	Nominal	9
Beamter	Nominal	2
Gebäudetyp	Nominal	10
Gebäudenutzung	Nominal	–
Gauß-Krüger	Stetig Räumlich	



Hiermit versichere ich, dass ich die vorliegende Hausarbeit selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Alle wörtlich oder sinngemäß den Schriften anderer entnommenen Stellen habe ich unter Angabe der Quellen kenntlich gemacht. Dies gilt auch für beigefügte Zeichnungen, Skizzen, bildliche Darstellungen und dergleichen.

Mir ist bewusst, dass ich mich im Falle einer unbeabsichtigten oder vorsätzlichen Missachtung durch den fehlerhaften Umgang mit Quellen unter Umständen strafbar mache und die vorliegende Hausarbeit mit nicht ausreichend bewertet wird.

Göttingen, den
Unterschrift

Hiermit erlaube ich, dass meine Arbeit auf Betrug und falsche, sowie fehlende Zitate auch online geprüft wird.

Mir ist bewusst, dass ich mich im Falle einer unbeabsichtigten oder vorsätzlichen Missachtung durch den fehlerhaften Umgang mit Quellen unter Umständen strafbar mache und die vorliegende Hausarbeit mit nicht ausreichend bewertet wird.

Göttingen, den
Unterschrift