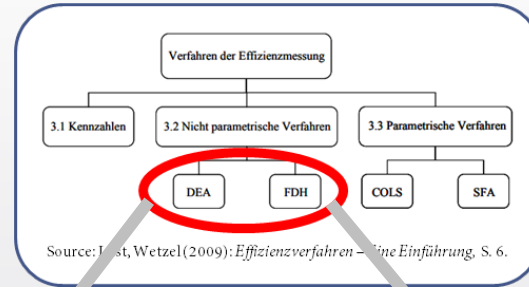




## Ablauf des Seminars

Termin Dienstag: 14:45	Inhalt
15.01.2019	Theoretische Einführung nicht-parametrische Methoden Teil I
22.01.2019	<b>Theoretische Einführung nicht-parametrische Methoden Teil II</b> <b>Forschungsprozess, Prüfungsleistung, Gruppeneinteilung, Themenvergabe</b>
29.01.2019	Praktische Einführung in STATA Teil I (PC-Pool 4 in 46/104)
05.02.2019	Praktische Einführung in STATA Teil II (PC-Pool 4 in 46/104)
Ca. 12.03.2019	<b>!Einzeltermine zu: Zwischenstand, Problemen, Fragen</b> <i>(bitte eine Woche vorher vereinbaren: alexander.kaiser@unibw.de)</i>
29.03.2019	Deadline Exposé

# Unterscheidung nicht-parametrischer Verfahren



Data Envelopment Analysis (DEA):

- A priori Spezifikation der Skalenerträge
- Technologie: Stetige Produktionsfunktion

Varianten:

- CCR-Modell
- BCC-Modell
- DDF-Modell

Free Disposal Hull (FDH):

- **Keine Spezifikation der Skalenerträge**
- Technologie: Diskrete Treppenfunktion

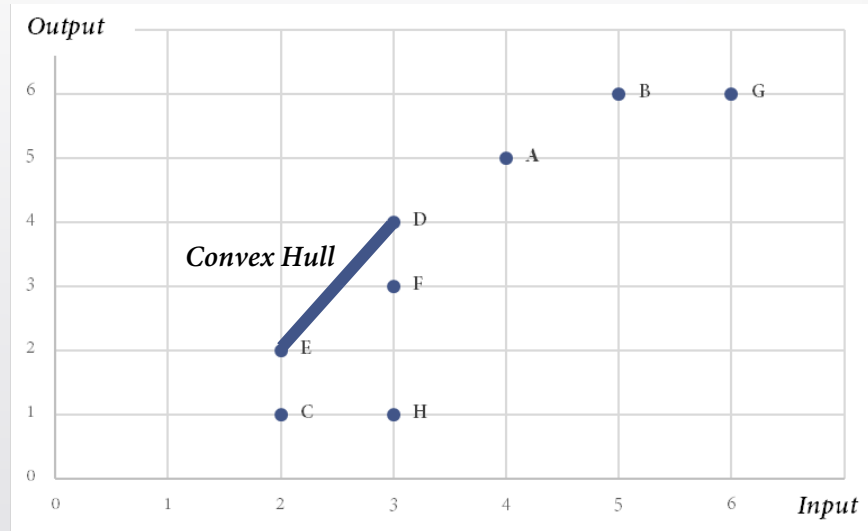
Varianten:

- Order-m Analysis
- **Orderalpha Analysis**

Nicht-parametrische Verfahren und Zeitreihen:

- Malmquist Productivity Index
- Malmquist-Luenberger Productivity Index

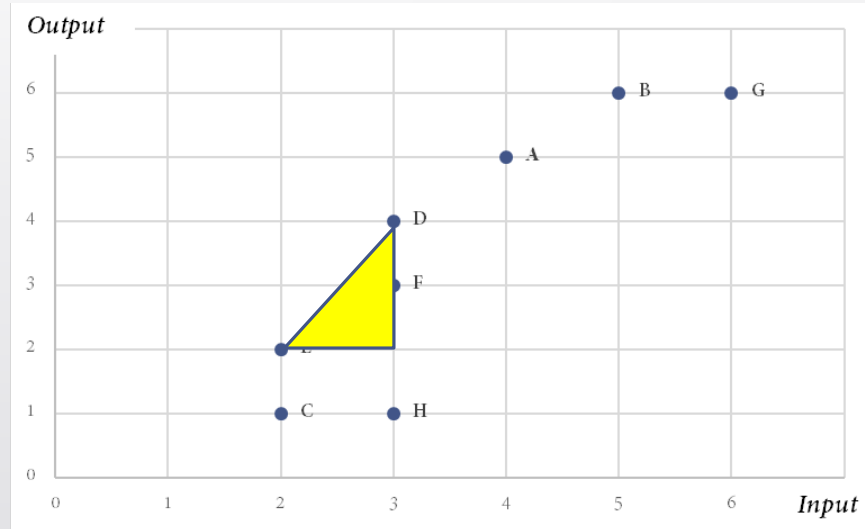
# Grundlagen: Free Disposal Hull



Bisher haben wir angenommen, dass die „Efficiency Frontier“ alle konvexen Möglichkeiten „zwischen“ zwei effizienten Punkten einschließt. Das ist jedoch empirisch schwer begründbar. Es sind Fälle denkbar, in denen zwischen zwei DMUs große Abstände liegen und die „convexity assumption“ eine Effizienzgrenze aufweist, welche für die DMUs nicht erreichbar ist. Mit anderen Worten: Der Bereich zwischen zwei effizienten DMUs kann nicht mit Sicherheit als „machbare“ Input-Output Kombination ausgewiesen werden.

# Grundlagen: Free Disposal Hull

- Die gelb markierte Fläche wird durch die „convexity assumption“ zu machbaren Input-Output-Kombinationen erklärt
- Das ist empirisch nicht beobachtbar
- Ein konvexer Produktionszusammenhang ist begründbar, aber schwer nachweisbar
- Nicht-parametrische Methoden sollen anhand der realen / empirischen Beobachtungen „Efficiency Frontiers“ bilden



**Lösung:** „Free Disposal Hull“, also Bilden einer Treppenfunktion, welche nur aus realen Beobachtungen besteht

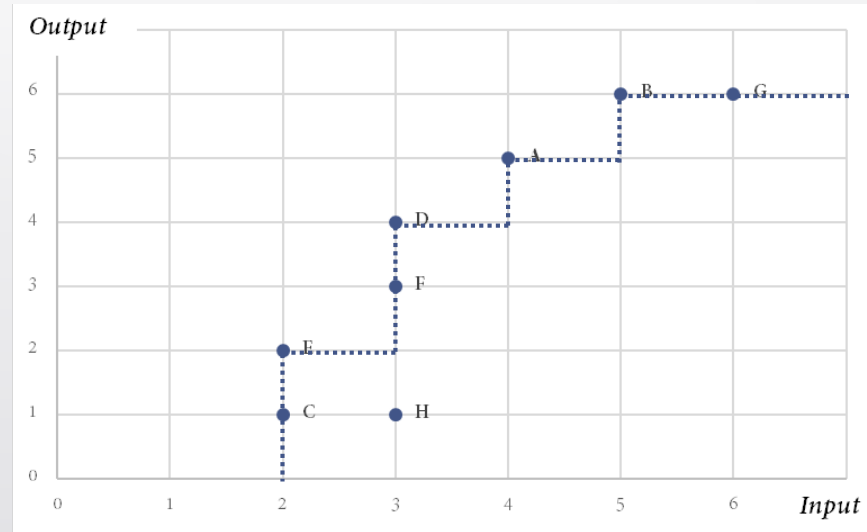
## FDH-Modell (1/4): Deprins, Simar, Tulkens

### FDH berücksichtigt:

- Unteilbarkeit der Inputs und Outputs
- Spezialisierung
- Economies of Scale

### Aber:

- Sofern das Datenset einen konvexen Zusammenhang abbildet, weist der DEA-Schätzer eine höhere Konvergenz zur realen Produktionssituation auf.
- Bei FDH treten verstärkt Slacks auf



**Lösung:** „Free Disposal Hull“, also Bilden einer Treppenfunktion, welche nur aus realen Beobachtungen besteht

## FDH-Modell (2/4): Partielle-“Frontier“-Ansätze

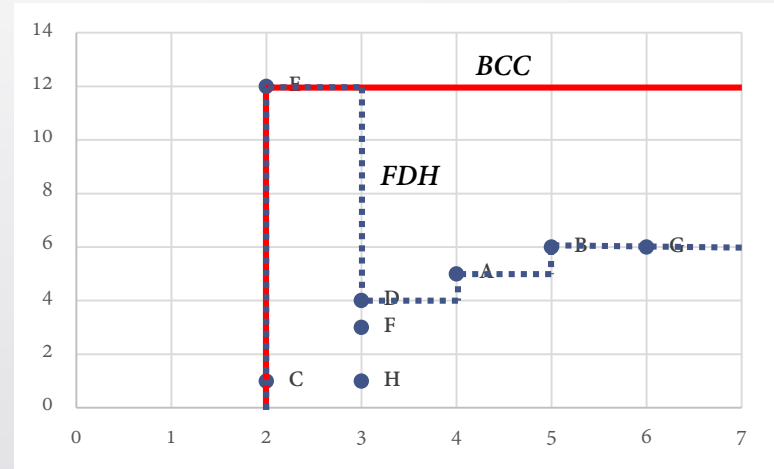
Sowohl DEA- Modelle, als auch das Standard FDH-Modell sind extrem anfällig für Ausreißer

Bei DMU E handelt es sich mit hoher Wahrscheinlichkeit um einen Ausreißer.

Dies kann mehrere Ursachen haben:

- Measurement Bias
- Sampling Bias
- Selection Bias
- Incubation Bias
- Survivorship Bias

Bei einem kleinen Datensatz und einem extremen Fall ist es einfach, Ausreißer zu identifizieren und zu eliminieren



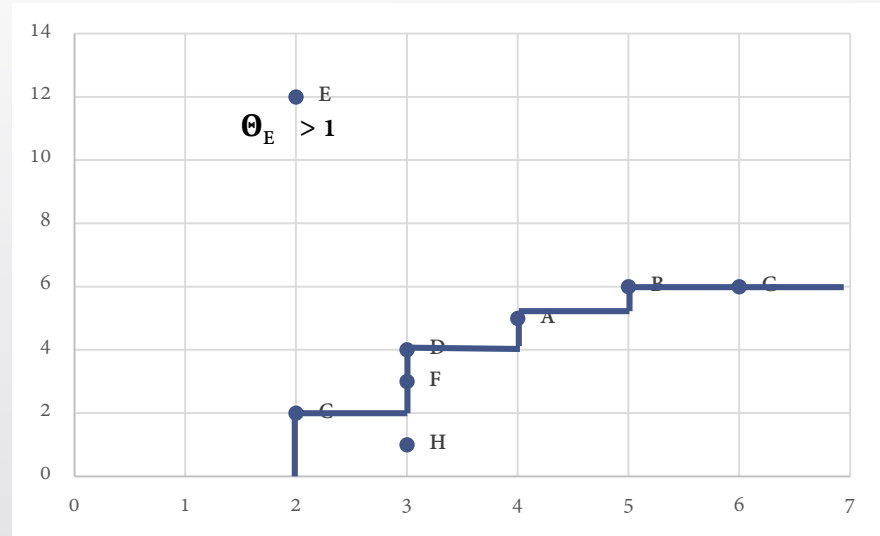
Die Weiterentwicklungen des FDH-Modells, die Order-m und Order-alpha Analyse bilden partielle Effizienzgrenzen, vereinfacht gesagt, sie erlauben Beobachtungen (DMUs), welche über der Effizienzgrenze liegen

## FDH-Modell (3/4): Partielle-“Frontier“-Ansätze

### Order-m Analysis:

- Es werden  $m$  DMUs gezogen
- Für die  $m$  DMUs werden Theta berechnet
- Die beiden Schritte werden  $D$  mal wiederholt
- Order-m Scores werden aus dem Mittel der berechneten Theta der Teilsample gebildet
- Da nicht alle DMUs in jedem Teilsample als Eigenreferenz enthalten sind können Theta  $> 1$  auftreten.

Diese DMUs gelten als supereffizient und liegen damit jenseits der „Efficiency Frontier“.

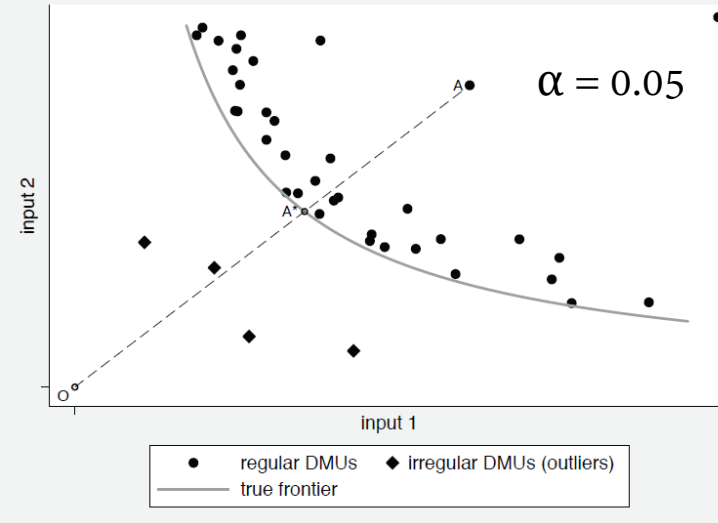


Order-m Analysis kann DMUs als supereffizient klassifizieren und damit Outlier entdecken. Je öfter ein Sample gezogen wird ( $D \rightarrow \infty$ ), umso akkurater werden die Scores (zu Last hoher Berechnungszeit). Je kleiner  $m$  gewählt wird, umso mehr DMUs werden supereffizient. Für  $m \rightarrow \infty$  gilt: order-m = FDH

## FDH-Modell (4/4): Partielle-“Frontier“-Ansätze

### Order-alpha Analysis:

- Anstatt ein sample zu ziehen wird bei der Order-alpha Analysis das  $(100 - \alpha)$  Perzentil als Referenz-Set (Input-Minimierung in der Abbildung) verwendet.
- Das  $\alpha$  Perzentil wird a priori als supereffizient eingestuft
- Allerdings ist die Wahl des  $\alpha$  Wertes ebenfalls arbiträr und kaum empirisch rechtfertigbar.
- In der Literatur werden häufig die Signifikanzwerte angewendet (0.01, 0.05, 0.1)



Order-alpha Analysis erlaubt eine a priori Spezifikation des als supereffizient zu klassifizierenden Perzentils. Obwohl die Wahl des  $\alpha$  Wertes arbiträr ist, ist es bei mehrdimensionalen Produktionszusammenhängen und gleichzeitig großen Datensätzen sinnvoll per se Ausreißer zuzulassen. Für  $\alpha = 0$  gilt FDH.





## Prüfungsleistung: Anwendung einer nicht-parametrischen Methode und Exposé

1. Gruppeneinteilung
2. Wahl eines geeigneten Themas
3. Festlegen einer Forschungsfrage
4. Beschaffen eines Datensatzes
5. Durchführen der DEA / FDH in Stata in euren Gruppen
6. Aufbereitung des Forschungsprozesses in einem Exposé / Extended Abstract (individuell)

Abgabe: 29.03.2018

Umfang: ca. 8-10 Seiten (, davon min. 4-5 Seiten zu Methodik und Ergebnisse der angewandten Methode)

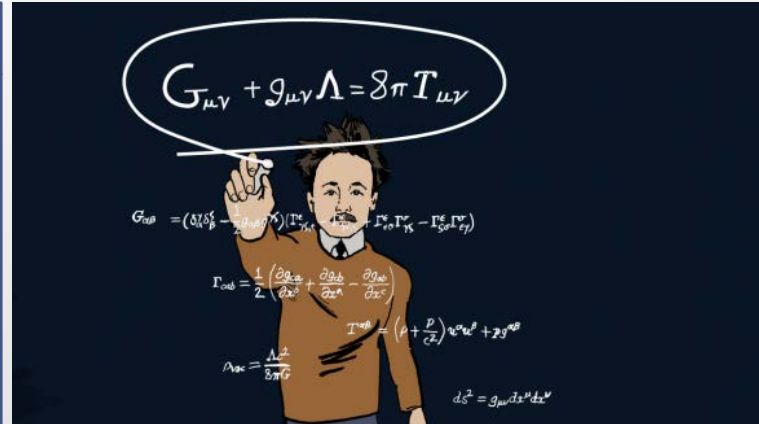
# Die empirische Methode

Der wissenschaftliche Forschungsprozess wird wesentlich von der gewählten Forschungsphilosophie bestimmt

## Empirismus

### Merkmale:

- Ein subjekt-unabhängiger Realitäts-Begriff – Welt existiert unabhängig von unserem Wissen, sie ist objektiv begreifbar, der Forscher ist unabhängig
- Erfahrung ist Grundlage der Erkenntnis
- Gültiges Wissen sind beobachtbare Phänomene, Daten, Fakten, Zahlen, Fokus auf Ursache-Wirkungszusammenhänge
- Methoden: Datenerhebungen und -analysen, quantitativ wie qualitativ



Verschaffen uns die mathematischen Ausdrücke zur Raumzeit „Wissen“ über Raum und Zeit?

Der Empirismus hat sich in den vergangenen Jahrhunderten für die Naturwissenschaften (und in der Folge für den technologischen Fortschritt) als sehr ergiebig erwiesen.

*Aber: Die Wirtschaftswissenschaften sind keine Naturwissenschaft!*



## Exkurs: Die empirische Methode – Grenzen

Es ist erstaunlich, dass ein empirisches Weltbild unhinterfragt ‚common sense‘ geworden ist.

*Annahme vieler Forscher der Empirismus stehe auf einem sicheren Erkenntnisfundament. Aber was genau soll dieses Fundament denn sein?*

Einige Beispiele für simple, grundlegende Fragen, auf welche die Empirie bei genauerer Betrachtung keine Antwort hat.

1. *Alles ‚geistige‘ ist psychologisch zu erklären! – ABER: Unmöglichkeit einer Privatsprache?*
2. *Alles hat Ursache und Wirkung – Man erlernt eine Sache, dann beherrscht man die Sache (klare Grenzen)  
– ABER: Regelfolgenproblem (Man kann einer Regel weder privat, noch einmalig folgen)*
3. *Wissen ist fundamentaler als begründen – ABER: Grundbegriffe wie „Sein“, „Wahrheit“, „Verstehen“ zu definieren  
scheitert an der Zirkularität innerhalb der Sprache; man verlässt sich auf sie, aber kann man sich auf sie verlassen?*
4. *Zeit ist messbar – ABER: Was ist Zeit?*
5. *Alles ist letztlich körperlich – ABER: „Das Leib-Seele Problem“*

# Exkurs: Die empirische Methode – Grenzen

Monismus	Dualismus
Nur Körper	Körper und Seele

Versuch eine der beiden ‚Philosophien‘ zu beweisen:

*Peter Frederic Strawson: Individuals. An Essay in Descriptive Metaphysics*

M-Prädikate	P-Prädikate
Wiegt 5 kg	Spielt Schach
Steht vor dem Fernseher	Hat Schmerzen
Hat Klamotten an	Freut sich
Ist brennbar	Denkt nach

# Exkurs: Die empirische Methode – Grenzen

## Monismus

M-Prädikate	P-Prädikate
Wiegt 5 kg	Spielt Schach

*Hypothese: wenn der Monismus gilt, dann haben P-Prädikate keine Berechtigung. Alle mit P-Prädikaten ausgedrückte Sachverhalte müssen auch mit M-Prädikaten auszudrücken sein.*

**Bsp:** „All my experiences are had by body B“. „my“ muss noch substituiert werden, denn es ist ein P-Prädikat. Der Sinn des Satzes mit M-Prädikaten muss kontingent sein, also die Fakten des P-Satzes repräsentieren (In der Aussagen- / Prädikatenlogik sind Formeln kontingent, wenn sie nicht ‚automatisch‘ wahr, oder falsch sind):

1. „All experiences of Person P are had by body B“. – Person P ebenfalls P-Prädikat (zirkulär)
2. „All experiences are had by body B“ – offensichtlich falsch, inhaltlich nicht richtig repräsentiert
3. „All experiences of body B are had by body B“ – tautologisch (nicht kontingent), nicht informativ

*Der Ich-Charakter der Selbsterfahrung ist irreduzibel. Damit ist der Monismus widerlegt.*

# Exkurs: Die empirische Methode – Grenzen

## Dualismus

M-Prädikate	P-Prädikate
Wiegt 5 kg	Spielt Schach

*Hypothese:*

*Descartes:*

*M-Prädikate = res extensa*

*P-Prädikate = res cogitans*

*I) Eine Person kann sich selbst P-Prädikate nur dann stabil zuschreiben, wenn sie bereit und fähig ist, sie auch anderen Personen zuzuschreiben (Unmöglichkeit der Privatsprache)*

*II) Zwar kann ich mir selbst P-Prädikate ohne Rekurs auf M-Prädikate zuschreiben, doch für andere Personen ist das unmöglich. (Fremdzuschreibung)*

*Eine isolierte ‚res cogitans‘ ist also auch ausgeschlossen. M und P-Prädikate kommen nicht ohne einander aus.*

**Strawson: The concept of a person is not to be analysed as that of an animated body or an embodied anima.  
The concept of a person is logically prior to that of an individual consciousness.**

## Exkurs: Die empirische Methode – Grenzen

*Annahme vieler, die Welt sei objektiv begreifbar, oder der Empirismus stehe auf einem sicheren Erkenntnisfundament ist beweisbar falsch. Die Sachverhalte sind weder alle messbar und objektiv, noch völlig esoterisch. Welche der beiden Extreme man wählt, spielt dabei keine Rolle, beides ist gleichermaßen falsch.*



Die Metapher des Fundaments ist vor allem irreführend wegen seines statischen Charakters. Man scheint zur Wahl gezwungen zwischen dem Bild des Elefanten, der auf einer Schildkröte ruht und dem Bild einer großen Hegel'schen Schlange des Wissens, die sich in den Schwanz beißt. Keines von beiden ist zutreffend. Denn das empirische Wissen wie seine differenzierte Erweiterung, die Wissenschaft, ist rational, nicht weil es ein Fundament hat, sondern weil es sich um ein selbst-korrigierendes Unternehmen handelt, das jede Behauptung in Frage stellen kann, wenn auch nicht alle auf einmal.

# Methodologie – Qualitativ vs. Quantitativ

Man unterscheidet in der empirischen Forschung: (1) qualitative und (2) quantitative Forschungsmethoden

## Qualitativ

- Explorativ und hypothesengenerierend
- Schrittweise Theoriebildung
- Ziel: Wirklichkeit anhand der subjektiven Sicht relevanter Gesprächspartner abzubilden.
- Ursachen für Verhalten und das Verhalten verstehen

## Quantitativ

- Prüfen vorher festgelegter Hypothesen
- Beobachtbares Verhalten wird beschrieben und vorhersagbar gemacht
- Durch Modelle, Zusammenhänge und zahlenbasierte Ausprägungen
- Vollstandardisiert und strukturiert um Vergleichbarkeit bei erhobenen Messwerten zu erreichen

In der VWL sind beide Ansätze von Nöten, z.B. Makroökonomische Modelle lassen sich quantitativ stützen (Modellierung); In der Mikroökonomie lässt sich das Verhalten der Konsumenten qualitativ erklären (Ernst Fehr – Behavioral Economics)





# Methodologie – Hans-Werner Sinn und Ernst Fehr

Häufig geht in der Forschung Spezialisierung zu Lasten von Interdisziplinarität, welche nötig wäre um Zusammenhänge umfassend zu erkennen und zu erklären.

Ein Beispiel für interdisziplinäre Zusammenhänge:  
<https://www.youtube.com/watch?v=OnCNUoMgvC4>

Qualitativ

Mixed-Methods

Quantitativ

In der VWL sind beide Ansätze von Nöten, z.B. Makroökonomische Modelle lassen sich quantitativ stützen (Modellierung); In der Mikroökonomie lässt sich das Verhalten der Konsumenten qualitativ erklären (Ernst Fehr – Behavioral Economics)

# Der Forschungsprozess - Literaturrecherche


(Primärliteratur)	Sekundärliteratur	(Tertiärliteratur)
<ul style="list-style-type: none"><li>• <i>Berichte</i></li><li>• <i>Abschlussarbeiten</i></li><li>• <i>Tagungsberichte</i></li><li>• <i>Geschäftsberichte</i></li><li>• <i>Unveröffentlichte Manuskripte</i></li><li>• <i>Ministerialveröffentlichungen</i></li></ul>	<ul style="list-style-type: none"><li>• <b>Journalartikel</b></li><li>• <i>Monographien</i></li><li>• <i>Herausgeberwerke</i></li><li>• <i>Sammelbände</i></li></ul>	<ul style="list-style-type: none"><li>• <i>Datenbanken</i></li><li>• <i>Enzyklopädien</i></li><li>• <i>Wörterbücher</i></li><li>• <i>Bibliographien</i></li><li>• <i>Kataloge</i></li></ul>

Grundsätzlich gilt: Sekundärliteratur ist zitationsfähige, wissenschaftliche Literatur. Graue Literatur (nicht zu einem wissenschaftlichen Zweck verlegt) ist zulässig, sofern sie für die Arbeit unentbehrlich oder deren Gegenstand ist (z.B. Metaanalyse von dt. Tageszeitungen → dann ist die FAZ als Quelle Gegenstand)

## Die Suche nach Sekundärliteratur:

• **Opac UniBW:** Mono, Hrsg., Sammelb. <https://opac.unibw.de/InfoGuideClient.ubysis/start.do?Login=iguby>

• **OpacPlus der StaBi:** Journal-Artikel <https://opacplus.bsb-muenchen.de/metaopac/start.do>



# Von der Idee zur Forschungsfrage in 5 Schritten

1. Keine Ahnung	<ul style="list-style-type: none"><li>• Nachhaltigkeit hat mich schon immer irgendwie interessiert...</li><li>• Es ist ja auch politisch / gesellschaftlich gerade relevant...</li><li>• Ich hab mich schon in meiner Bachelorarbeit / Projektstudium...</li></ul>
2. Thema eingegrenzt	<ul style="list-style-type: none"><li>• Besonders relevant in der Debatte um „nachhaltige Entwicklung“ erscheinen mir die Herausforderungen in der modernen Landwirtschaft</li><li>• Vor allem die intensiven Anbaupraktiken und die Überdüngung mit Stickstoff scheint die umgebenden Ökosysteme negativ zu beeinflussen</li></ul>
3. Der Versuch einer Forschungsfrage	<ul style="list-style-type: none"><li>• Wie beeinflussen die intensiven Anbaupraktiken und die Übernutzung von Stickstoffdünger die umgebenden Ökosysteme?</li></ul>
4. Forschungslücke entdeckt	<ul style="list-style-type: none"><li>• Ein Weg zu einer nachhaltigen Landwirtschaft scheint steigende Effizienz zwischen den Produktionsmitteleinsätzen und den erzielten Outputs (Hektarerträgen)</li><li>• In der Literatur werden Schäden an den Ökosystemen häufig als negative Outputs beschrieben</li><li>• Die Effizienzanalyse zwischen Inputs und Outputs könnte mit nicht-parametrischen Methoden durchgeführt werden</li></ul>
5. (Arbeits-) Forschungsfrage	Effizienzanalyse in der Landwirtschaft unter Miteinbeziehung unerwünschter ökologischer Outputs – Eine nicht-parametrische Analyse für die Region X von A bis O

# Von der Idee zur Forschungsfrage in 5 Schritten

<p>1. Keine Ahnung</p>	<ul style="list-style-type: none"> <li>Nachhaltigkeit hat mich schon immer irgendwie interessiert...</li> <li>Es ist ja auch politisch / gesellschaftlich gerade relevant...</li> <li>Ich hab mich schon in meiner Bachelorarbeit / Projektstudium...</li> </ul>	<ol style="list-style-type: none"> <li>„Opacplus“ der Bayrischen Staatsbibliothek (Ausweis beschaffen)</li> <li>Nach „Nachhaltigkeit“ suchen</li> <li>Zu verschiedenen Themen „Abstract“ und „Introduction“ lesen</li> </ol>
<p>2. Thema eingegrenzt</p>	<ul style="list-style-type: none"> <li>Besonders relevant in der Debatte um „nachhaltige Entwicklung“ erscheinen mir die Herausforderungen in der modernen Landwirtschaft</li> <li>Vor allem die intensiven Anbaupraktiken und die Überdüngung mit Stickstoff scheint die umgebenden Ökosysteme negativ zu beeinflussen</li> </ul>	<ol style="list-style-type: none"> <li>Gezielt Schlagwörter (englisch) zu int. Thema im Opacplus suchen</li> <li>Nach Papern (Aufsätzen) mit klarem Themenbezug filtern</li> <li>Abstracts lesen und nicht zusammenhängende aussortieren</li> <li>3-5 interessante Paper komplett und strukturiert durcharbeiten</li> </ol>
<p>3. Der Versuch einer Forschungsfrage</p>	<ul style="list-style-type: none"> <li>Wie beeinflussen die intensiven Anbaupraktiken und die Übernutzung von Stickstoffdünger die umgebenden Ökosysteme?</li> </ul>	<p>Inhalte der Paper zusammenfassen und Probleme / Herausforderungen verstehen</p>
<p>4. Forschungslücke entdeckt</p>	<ul style="list-style-type: none"> <li>Ein Weg zu einer nachhaltigen Landwirtschaft scheint steigende Effizienz zwischen den Produktionsmitteleinsätzen und den erzielten Outputs (Hektarerträgen)</li> <li>In der Literatur werden Schäden an den Ökosystemen häufig als negative Outputs beschrieben</li> <li>Die Effizienzanalyse zwischen Inputs und Outputs könnte mit nicht-parametrischen Methoden durchgeführt werden</li> </ul>	<p>Hier beginnt der eigentliche Forschungsprozess mit Literaturrecherche</p>
<p>5. (Arbeits-) Forschungsfrage</p>	<p>Effizienzanalyse in der Landwirtschaft unter Miteinbeziehung unerwünschter ökologischer Outputs – Eine nicht-parametrische Analyse für die Region X von A bis O</p>	

# Das Exposé (oder Extended Abstract)

4. Forschungs-  
lücke entdeckt

5. (Arbeits-)  
Forschungsfrage

## Die Prüfungsleistung:

- Ein Exposé bietet eine detaillierte Übersicht über das Forschungsvorhaben (soz. eine wissenschaftliche Projektskizze)
- Ein Extended Abstract beinhaltet zusätzlich bereits erste Forschungsergebnisse und ist in der Regel kürzer
- Als Prüfungsleistung erwarten wir im wesentlichen ein Exposé, welches die Ergebnisse aus der praktischen Durchführung eurer nicht-parametrischen Analyse beinhaltet.
- Der Fokus liegt dabei auf den Teilen „Methodologie“ und „Ergebnisse“, welche ca. 5-6 der 8-10 Seiten (exkl. Referenzen) einnehmen sollen

## Aufbau eines Exposés / Extended Abstracts

1. Titel
2. Motivation (Herausarbeiten der Forschungslücke aus der Literatur)
3. Forschungsfragen- und ziele
- 4. Methodologie und Methodik**
- 5. Ergebnisse**
6. (*Zeitplan*)
7. (*Ressourcen*)
8. Referenzen

## 4. Methodologie und Methodik

Beschreibung des Forschungsprozesses:

- Forschungsdesign (Erläuterung: Kontext Motivation → Forschungsfrage → Beantwortung via empirischer Forschungsphilosophie)
- Techniken zur Datenerhebung, bzw. -beschaffung
- Struktur der Daten, Umfang, deskriptive Statistiken
- Methode zur Datenanalyse (DEA, FDH, Orderalpha)
- Detaillierte Ausführung zu der Anwendung der Methode (Begründung der Variablen, Optionen, Spezifikationen für die gewählte Forschungsfrage)



## Gruppeneinteilung und Themenvergabe

### **Bis zum nächsten Mal:**

- *Sofern ihr vorhabt außerhalb des Campus / mit Privat-Laptop zu arbeiten, installiert euch Stata (via StudiSoft-Portal)  
– zur Not hilft ein Gang zum RZ)*
- *Sucht in den gängigen Datenbanken (EuroStat, Stat. Bundesamt) nach Datensätzen zu dem gewählten Thema und ladet sie in eure Drive*