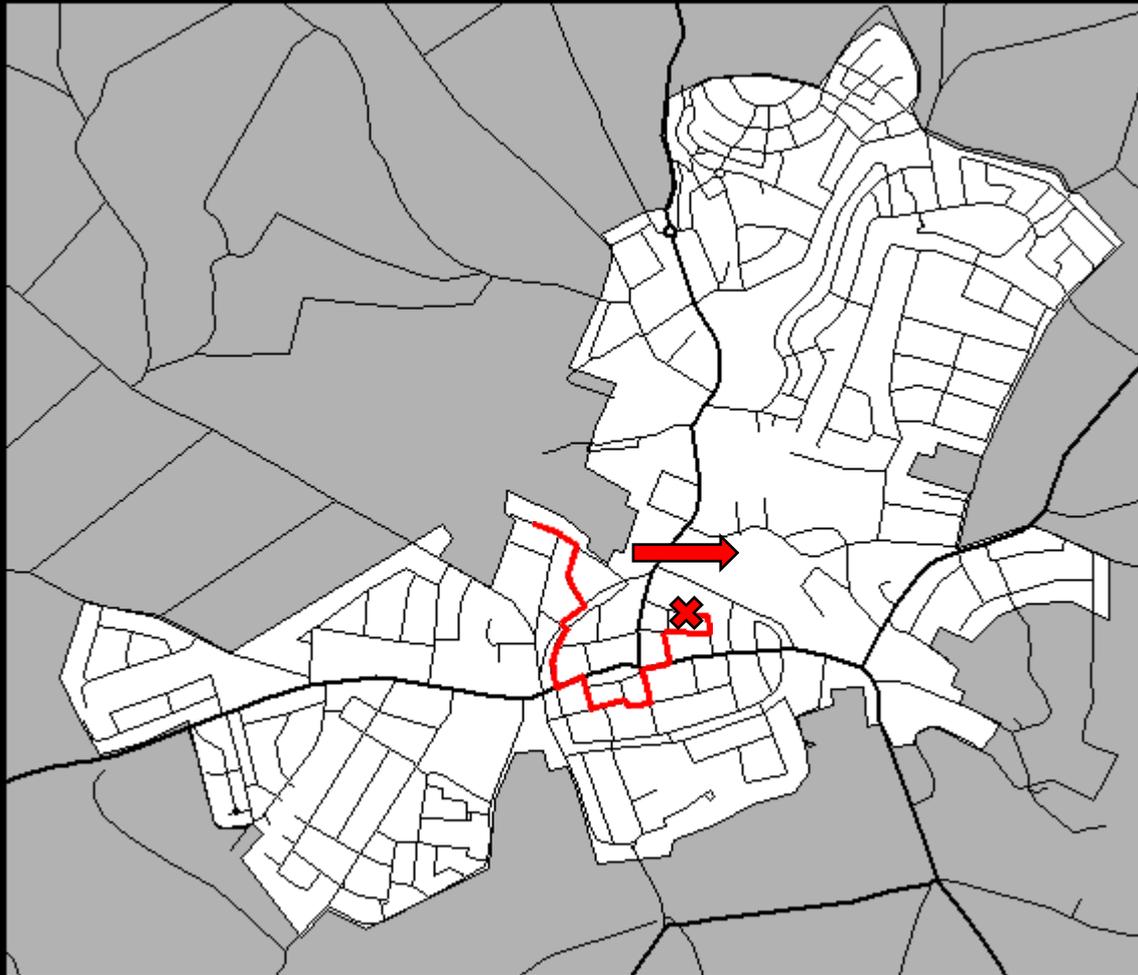


Institut für Soziologie  
Dipl. Soz. Johannes Bauer

# Verzerrungen in Random-Route Stichproben und Lösungsansätze





## Struktur von Random-Route Stichproben:

1. Wahl der Startadresse
2. Wahl der Startrichtung
3. Routenanweisung
4. Interview mit jedem x-ten Haushalt



## Aufbau des Vortrags:

1. Fehler in Auswahlwahrscheinlichkeiten
2. Auswirkung auf Studienergebnisse
3. Lösungsvorschläge



## European Quality of Life Survey 2012

- Erhebung in 34 Ländern darunter alle EU Mitgliedsstaaten

## Gallup World Poll

- Einstellungs und Verhaltensmessung
- “survey data [that] are representative of 95% of the world's adult population” (<http://www.gallup.com/>)

## VA Verbraucheranalyse

- Konsumverhalten sowie die Mediennutzung der Deutschen
- mündliche und schriftliche Befragung von über 30.000 bevölkerungsrepräsentativ ausgewählten Personen

## SOEP Aufwuchssample

## SOEP Innovationsstichprobe ( 2014 )

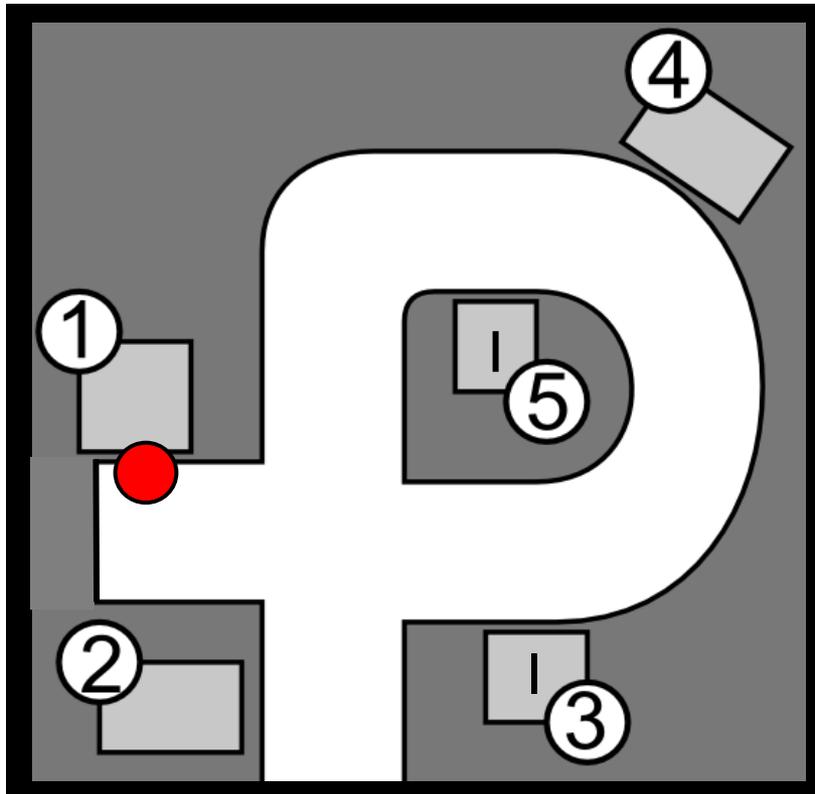


## Anforderungen an Random-Route Stichproben:

1. Gleiche Auswahlwahrscheinlichkeiten
2. Unabhängigkeit

## Bisherige Forschung:

Vergleich von Variablen externer Datenquellen wie dem Mikrozensus oder Einwohnermeldeamtstichproben, mit Random-Route Studien.



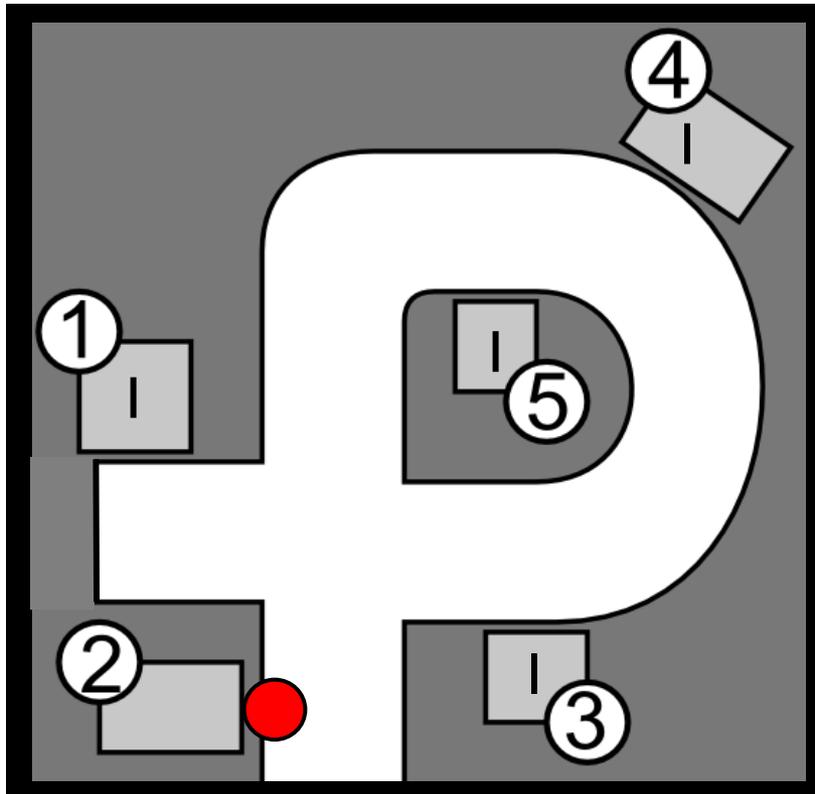
### Random-Route Anweisung:

- Abwechselnd links und rechts abbiegen (beginnend mit rechts)
- Bei rechts (links), laufe auf der rechten (linken) Seite der Straße.

### Sackgassen:

- Auf der gleichen Straßenseite zurück laufen
- Wenn nach rechts (links) betreten wurde, verlasse die Sackgasse rechts (links).

Jeden zweiten Haushalt, insgesamt Zwei Haushalte befragen.



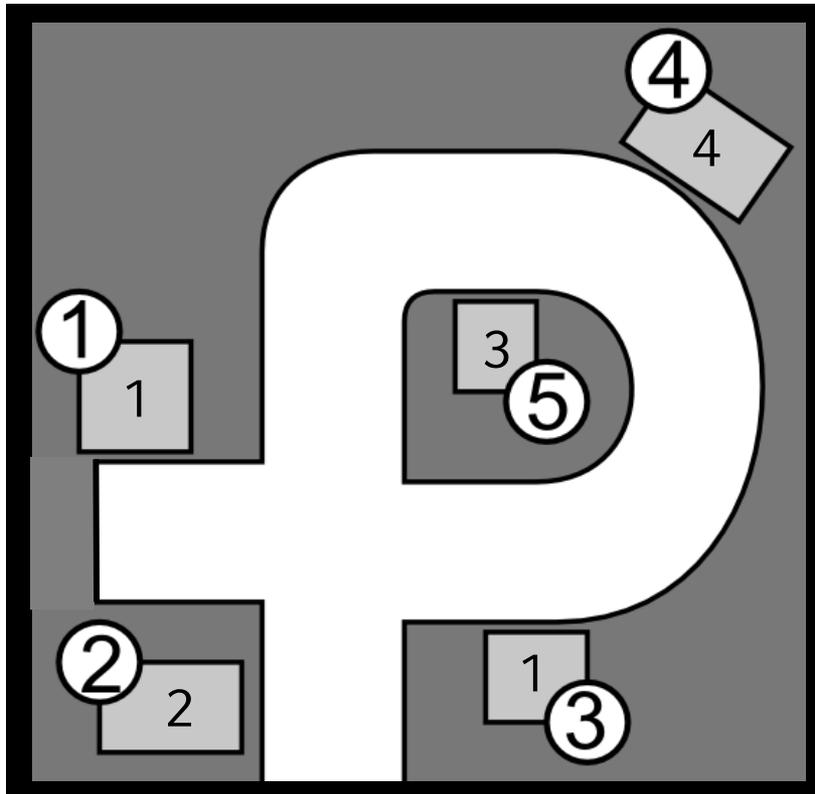
### Random-Route Anweisung:

- Abwechselnd links und rechts abbiegen (beginnend mit rechts)
- Bei rechts (links), laufe auf der rechten (linken) Seite der Straße.

### Sackgassen:

- Auf der gleichen Straßenseite zurück laufen
- Wenn nach rechts (links) betreten wurde, verlasse die Sackgasse rechts (links).

Jeden zweiten Haushalt, insgesamt Zwei Haushalte befragen.



### Random-Route Anweisung:

- Abwechselnd links und rechts abbiegen (beginnend mit rechts)
- Bei rechts (links), laufe auf der rechten (linken) Seite der Straße.

### Sackgassen:

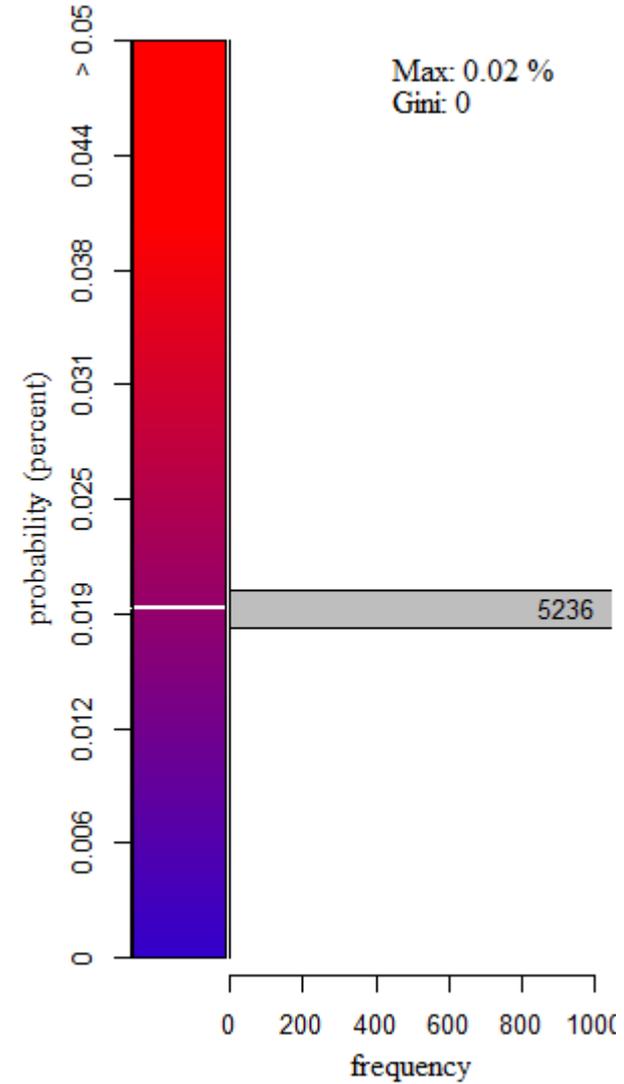
- Auf der gleichen Straßenseite zurück laufen
- Wenn nach rechts (links) betreten wurde, verlasse die Sackgasse rechts (links).

Jeden zweiten Haushalt, insgesamt Zwei Haushalte befragen.



## Benötigte Daten:

- Straßennetzwerk Daten: <http://www.openstreetmap.org/>
- Gebäude
- Personendaten
- Haushaltsdaten
- Telefonadressen
- Inhaltliche Variablen



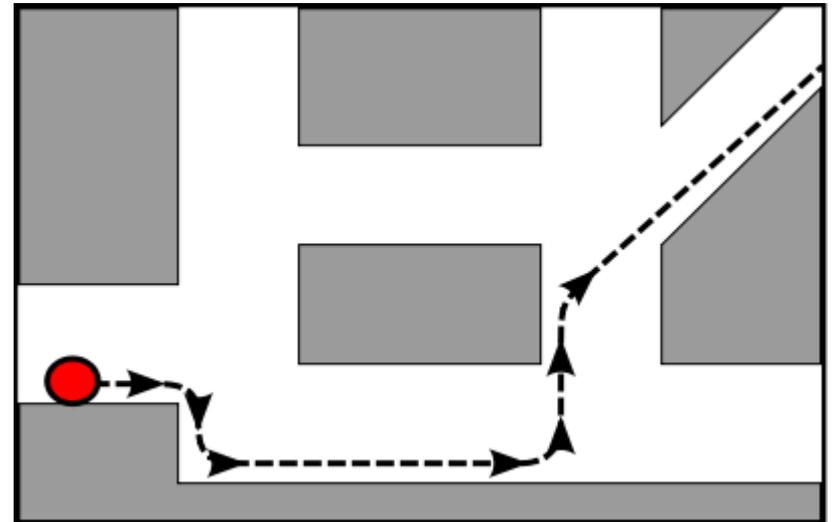
## Noelle-Neumann 1963

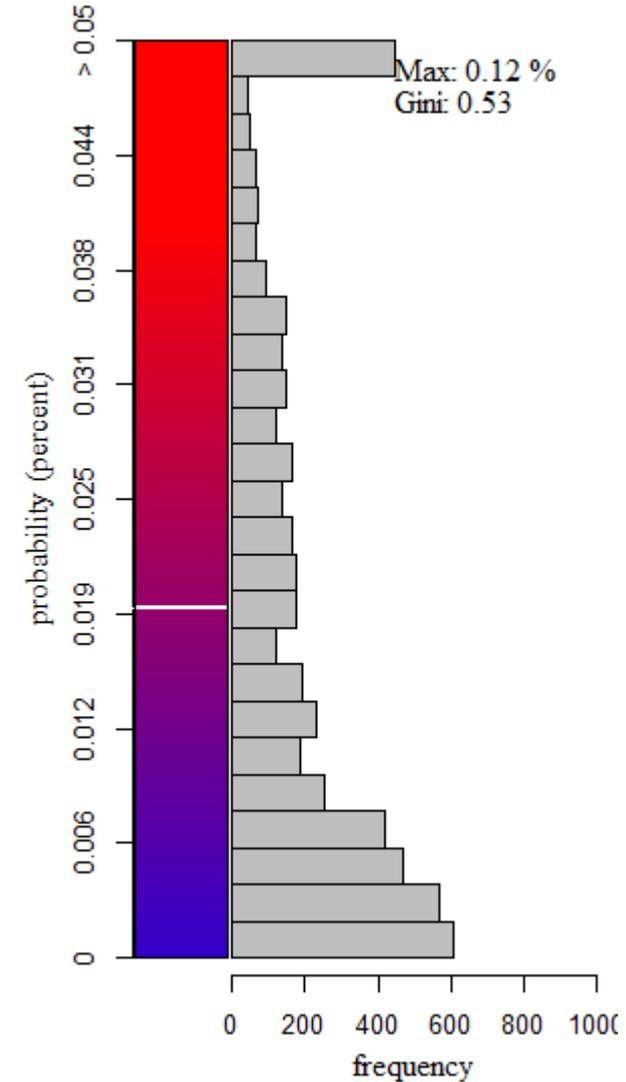
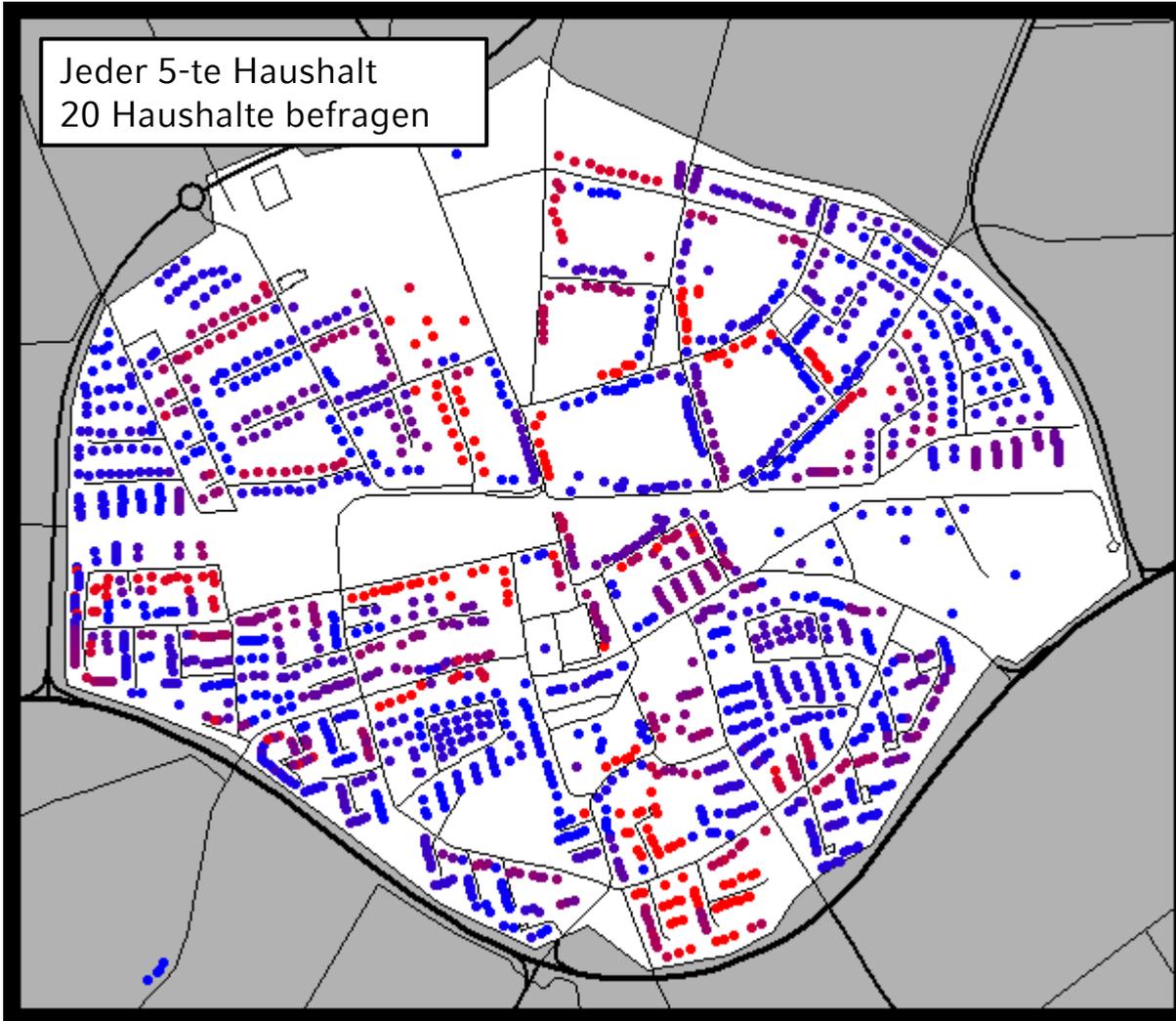
### Anweisung:

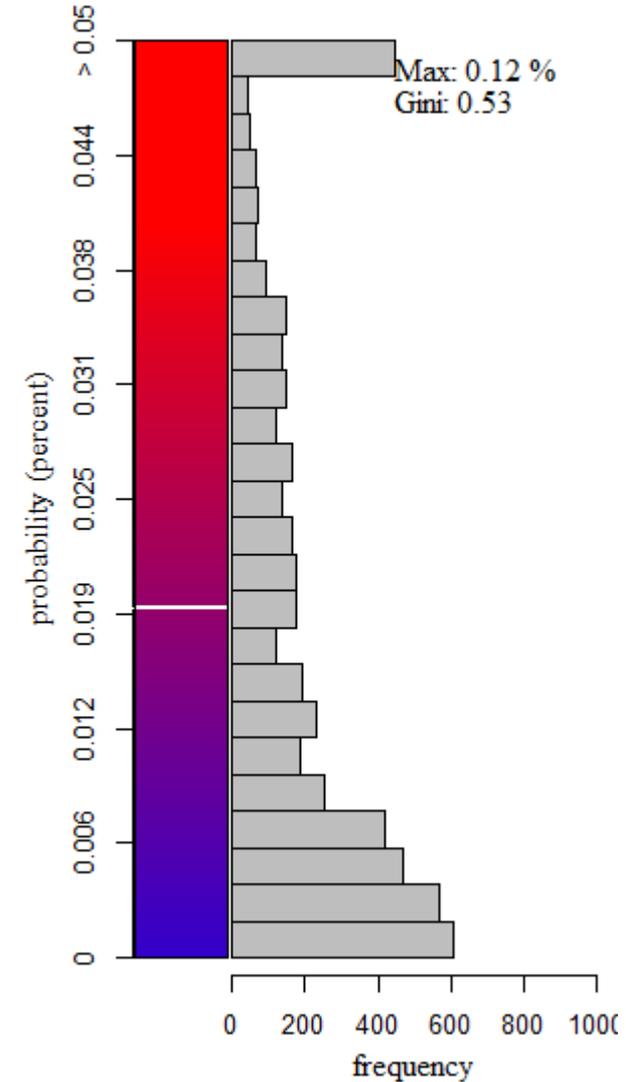
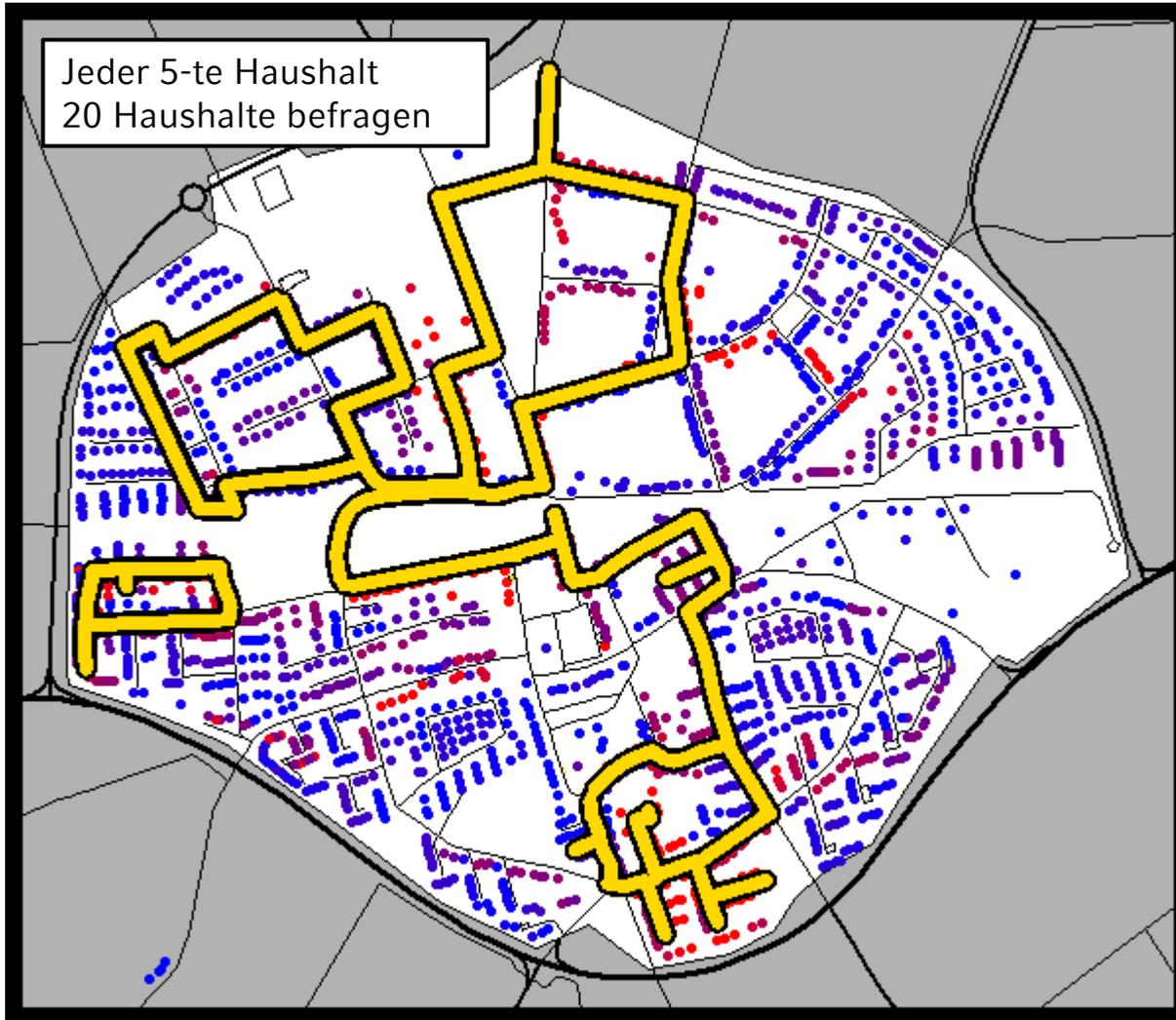
- Abwechselnd links und rechts abbiegen (beginnend mit rechts)
- Bei rechts (links), laufe auf der rechten (linken) Seite der Straße.

### Sackgassen:

- Auf der gleichen Straßenseite zurück laufen
- Wenn nach rechts (links) betreten wurde, verlasse die Sackgasse rechts (links).
- Keine Interviews beim zurück laufen

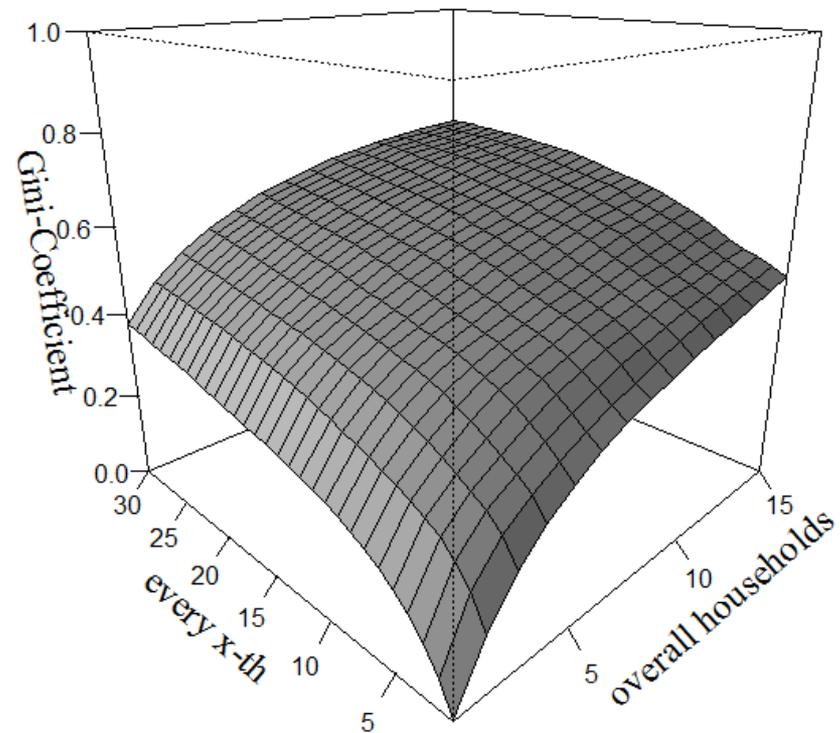








## Gini-Koeffizienten für verschiedene Selektionsintervalle und Kontaktzahlen



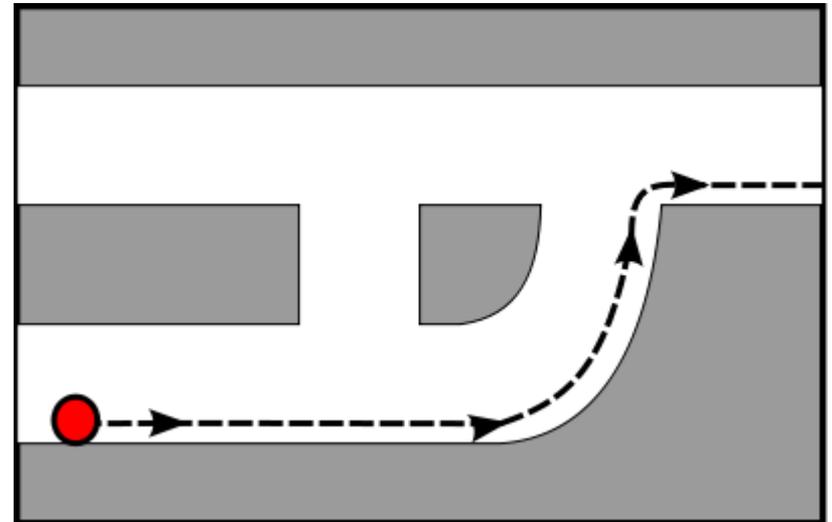
## TNS Infratest

### Anweisung:

- Laufe so lange wie möglich geradeaus.
- Ist geradeaus nicht möglich, biege rechts ab.

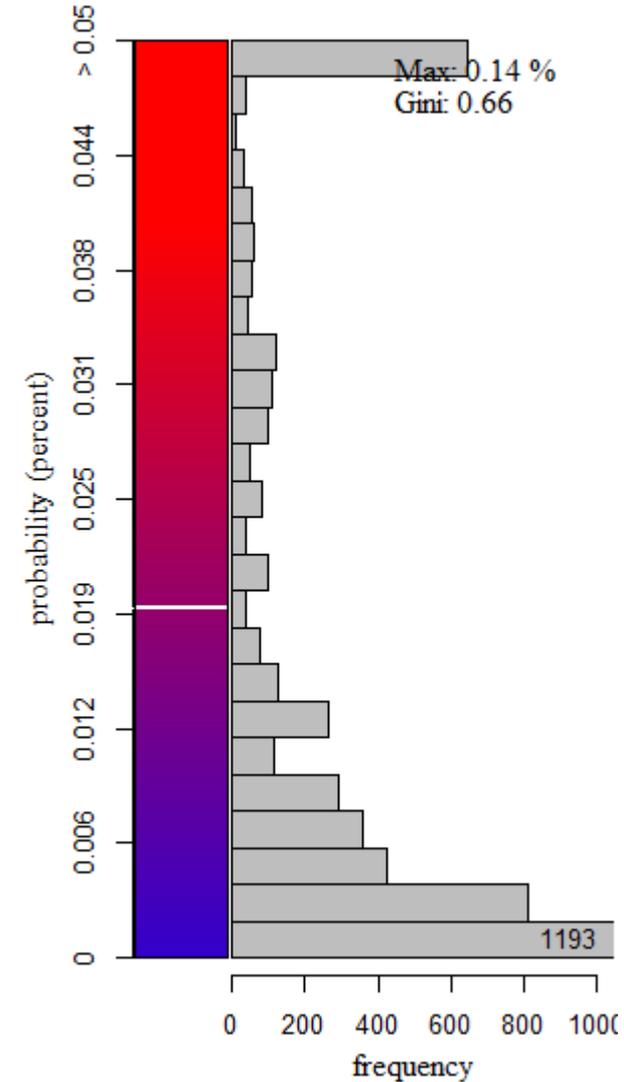
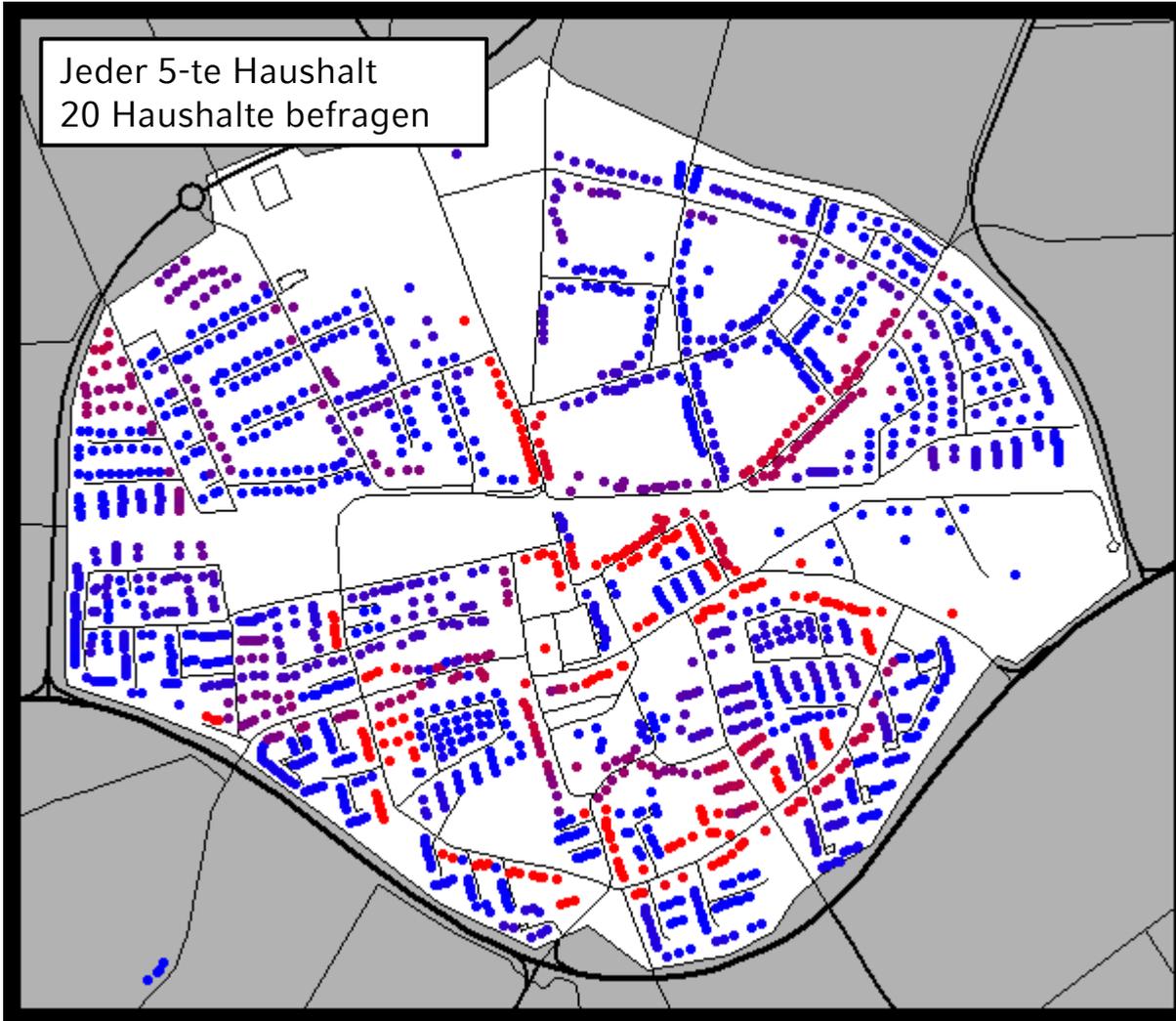
### Sackgasse:

- Laufe auf der anderen Straßenseite zurück.





Jeder 5-te Haushalt  
20 Haushalte befragen

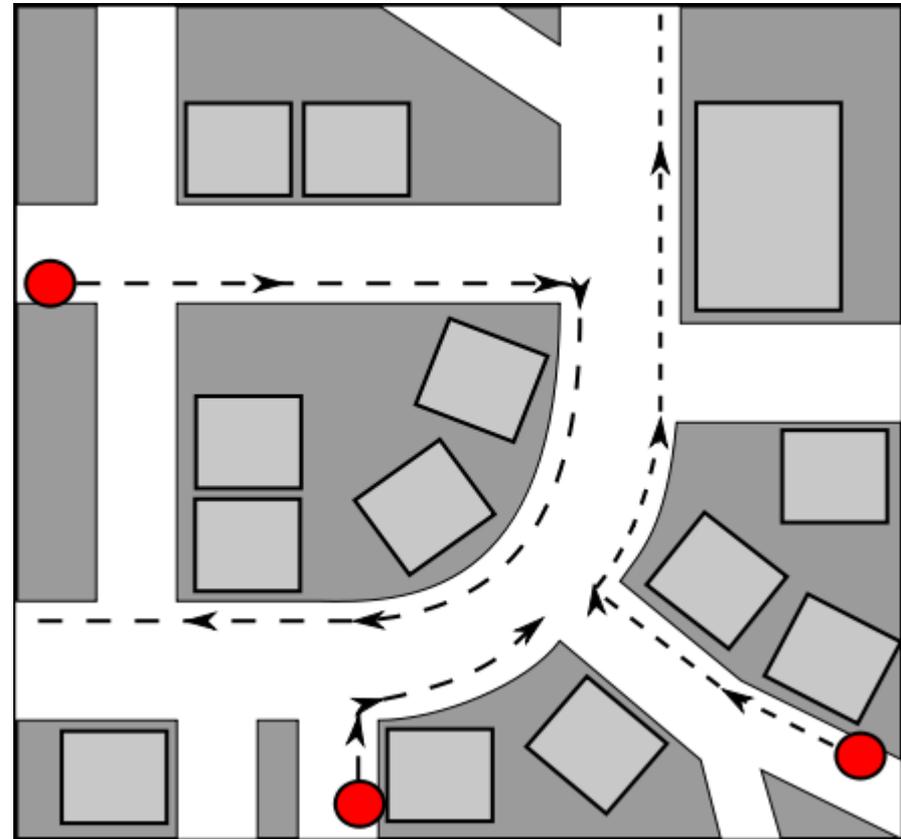


**Anweisung:**

- Laufe so lange wie möglich geradeaus.
- Ist geradeaus nicht möglich, biege rechts ab.

**Sachgasse:**

- Laufe auf der anderen Straßenseite zurück.



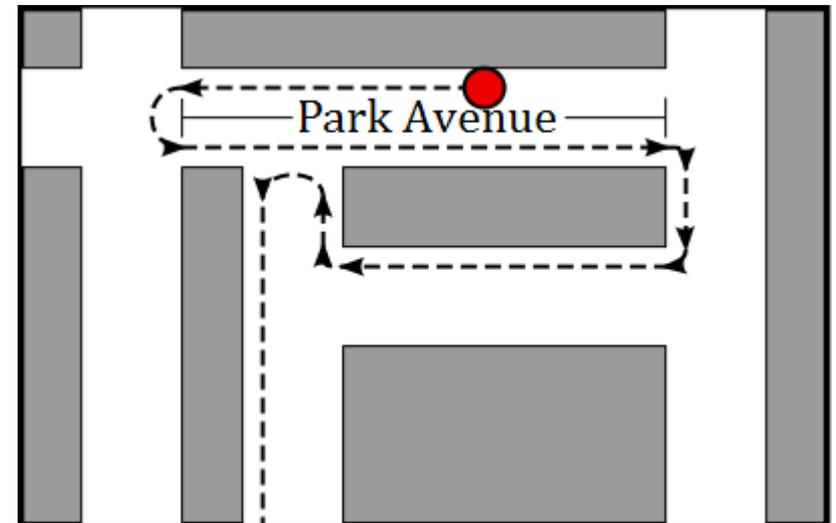
## Gesellschaft für Konsumforschung (GfK)

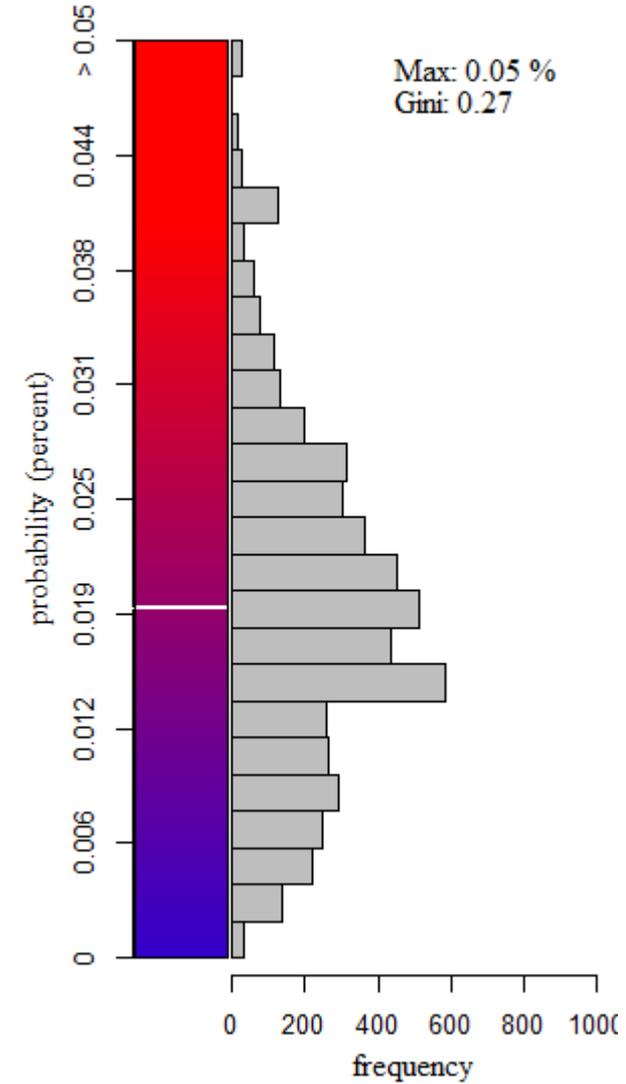
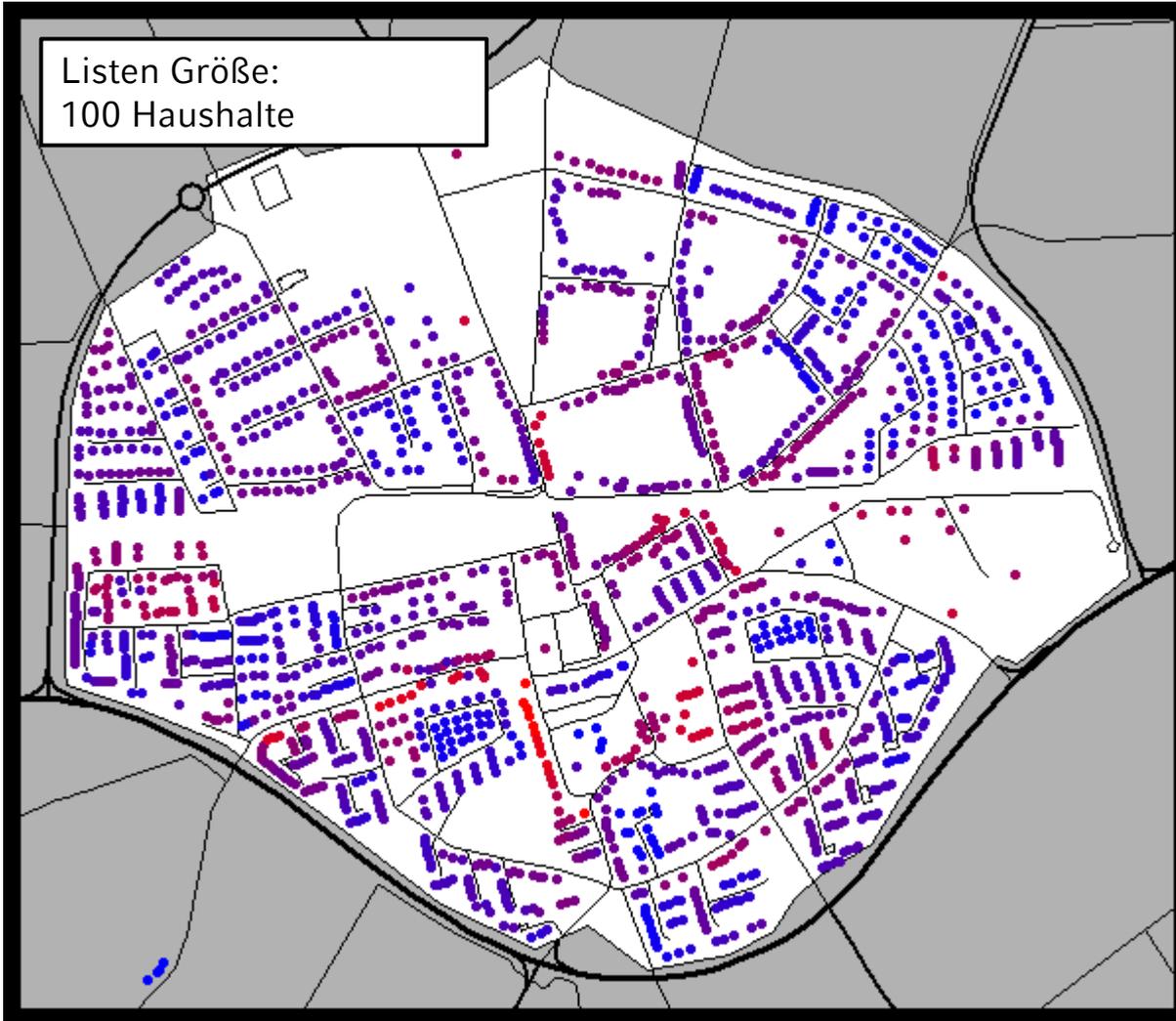
### Anweisung:

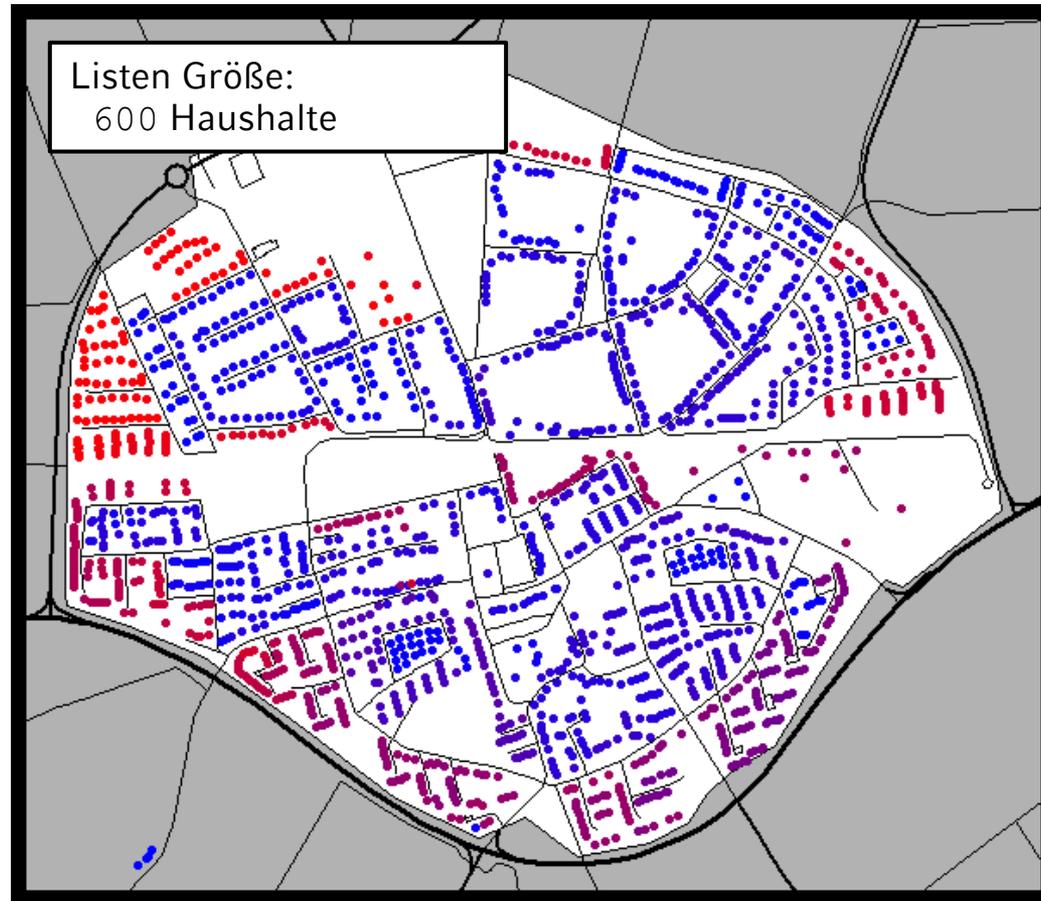
- Folge der Startstraße bis zum ende. Laufe zurück auf der anderen Straßenseite bis zum anderen Ende.
- Danach, biege immer nach rechts ab.
- Laufe niemals zweimal auf der gleichen Straßenseite.

### Sackgasse:

- Laufe auf der anderen Straßenseite zurück.









# Auswirkung auf Studienergebnisse

- Erwartungswert

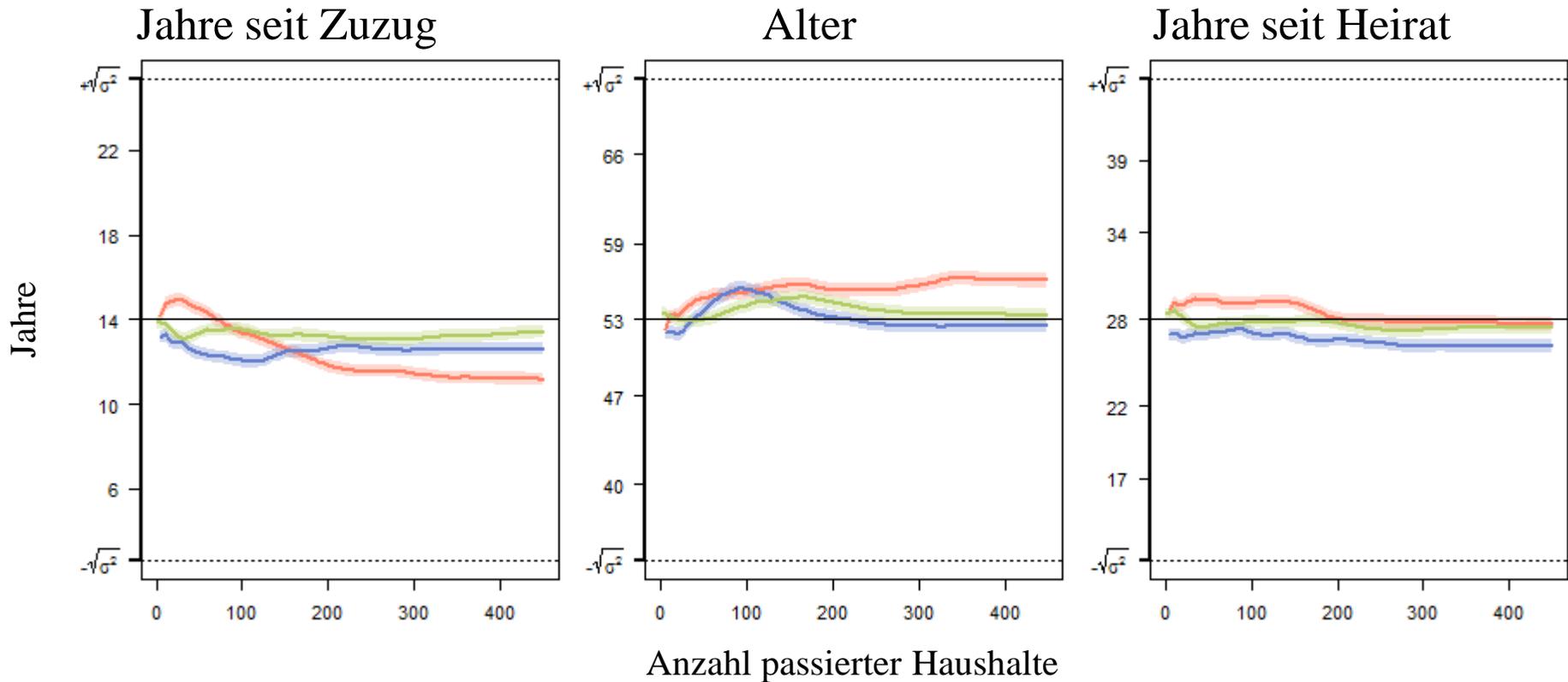
**Zeitlich:** Einzugsdatum, Geburtsdatum, Heiratsdatum

**Demographisch:** Geschlecht, Nationalität, Familienstand

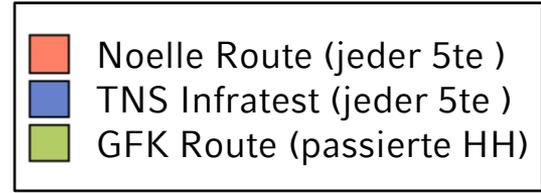
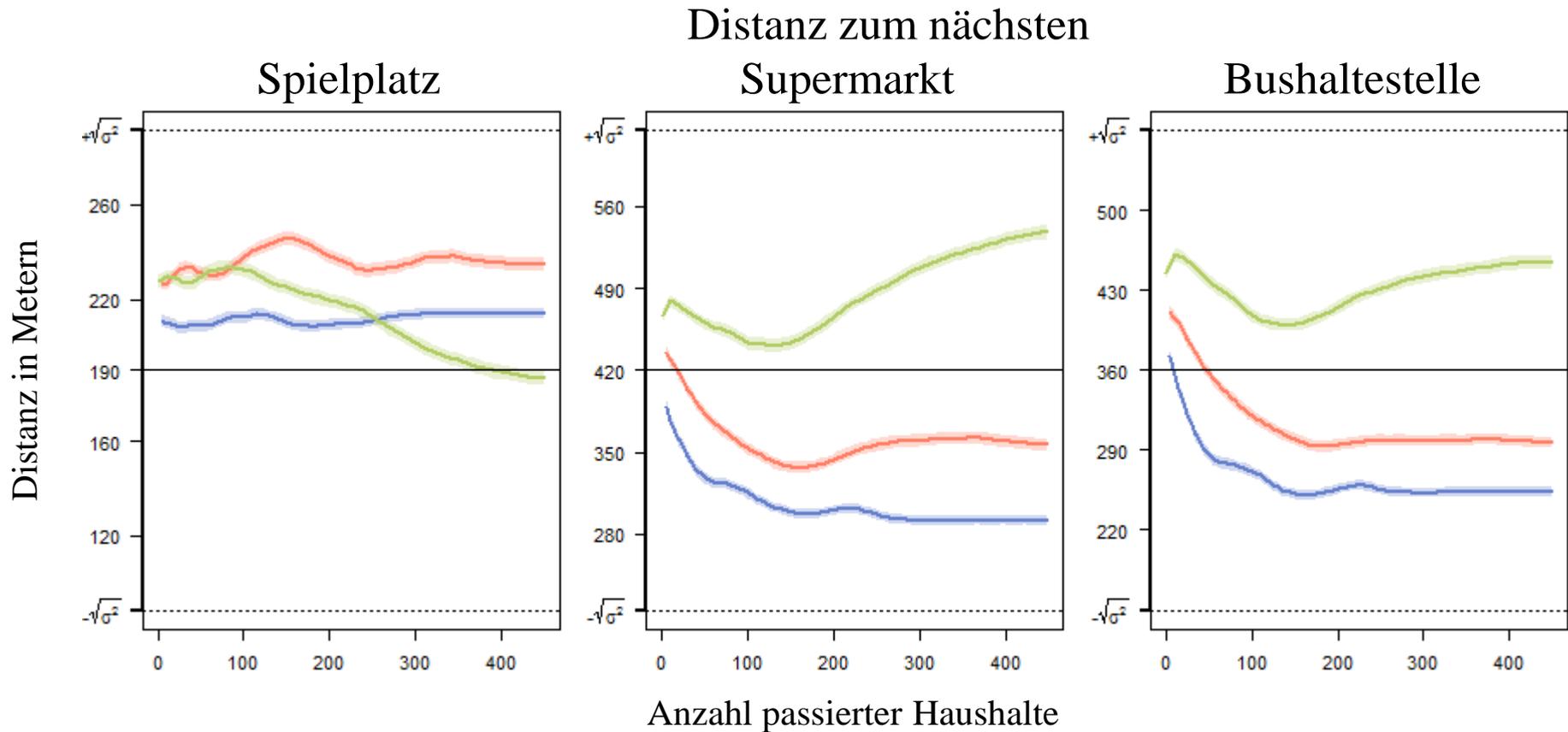
**Räumlich:** Luftdistanz bis zum nächsten Spielplatz, Supermarkt, Bushaltestelle

$$E_g(X) = \sum_{i=1}^N x_i \cdot w_i$$

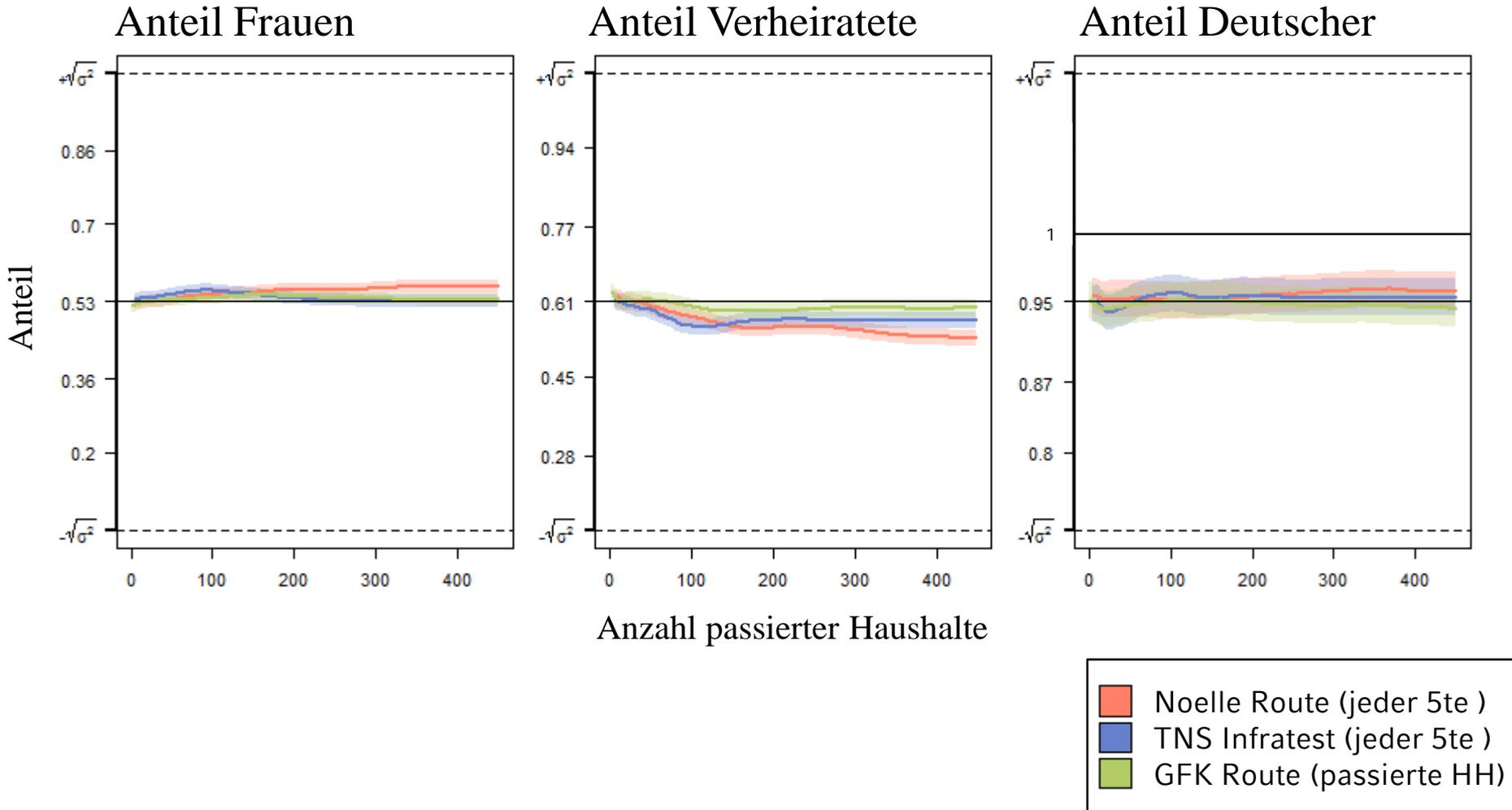
- Auswahlwahrscheinlichkeit  $w_i$  :
  - Gewichtung nach Haushaltsgröße
  - Telefonbuchadressen als Startadressen



- Noelle Route (jeder 5te )
- TNS Infratest (jeder 5te )
- GFK Route (passierte HH)



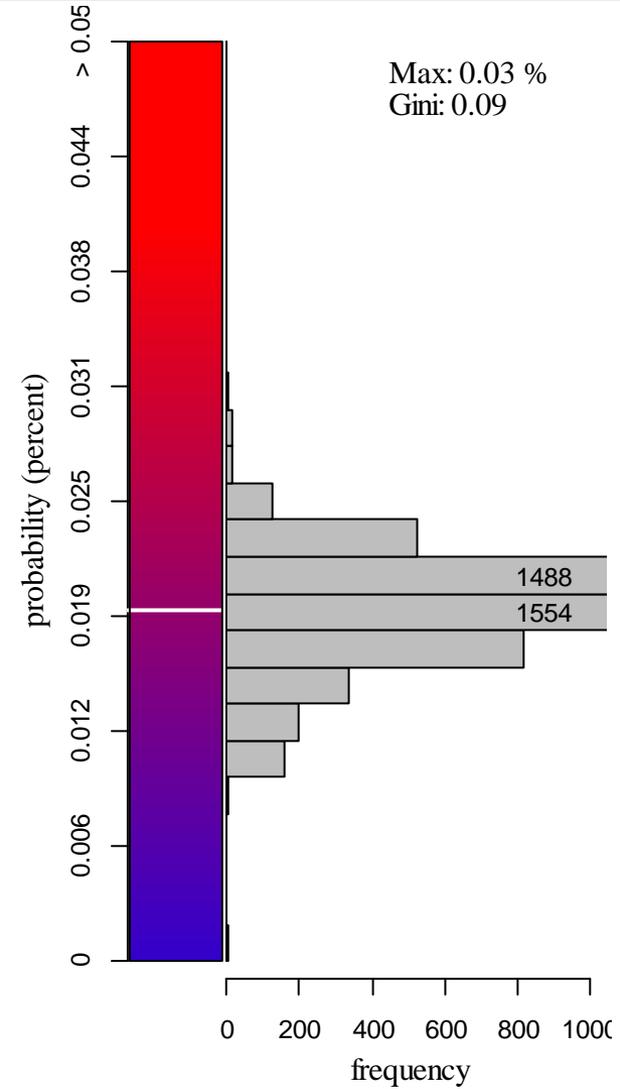
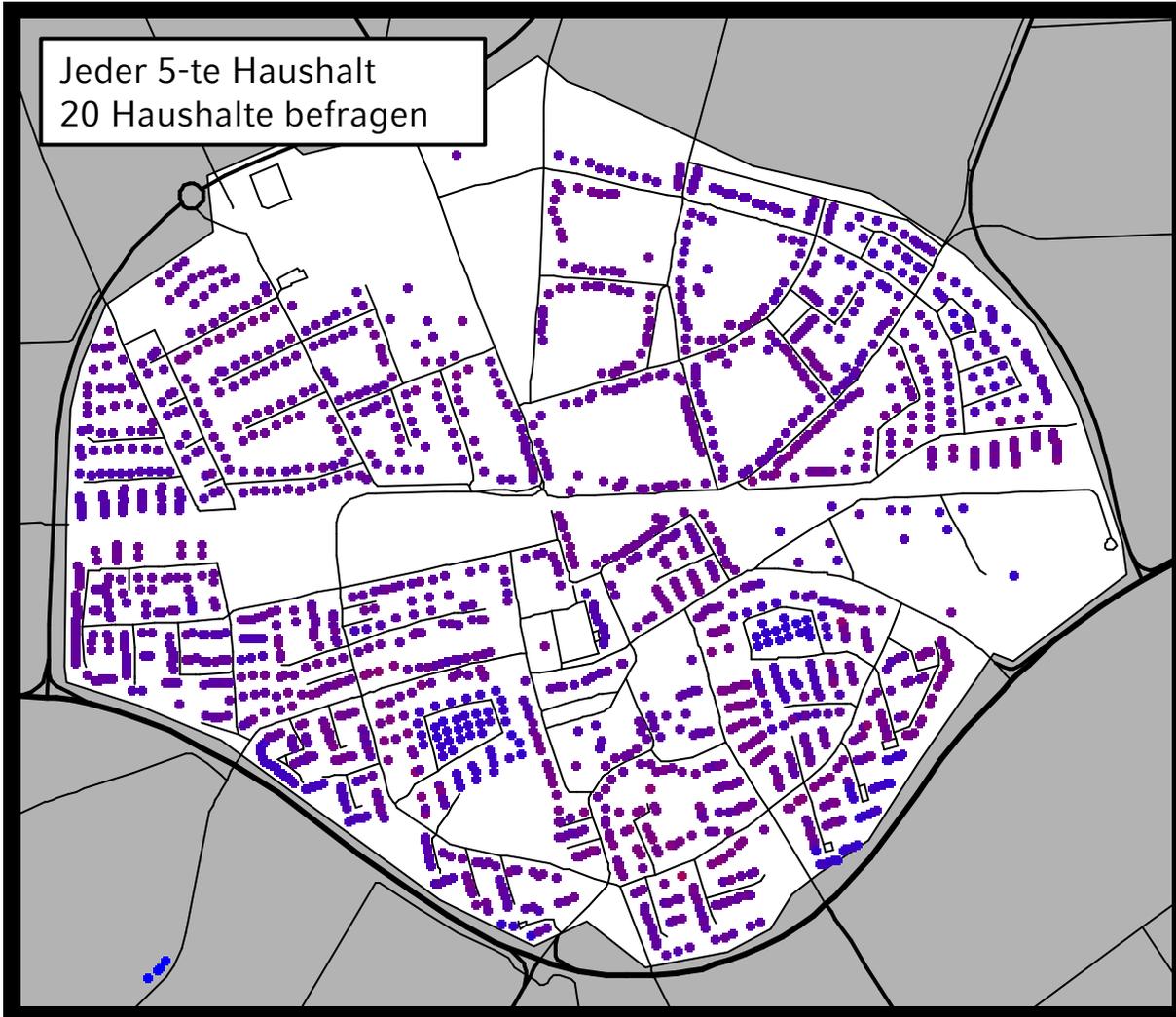


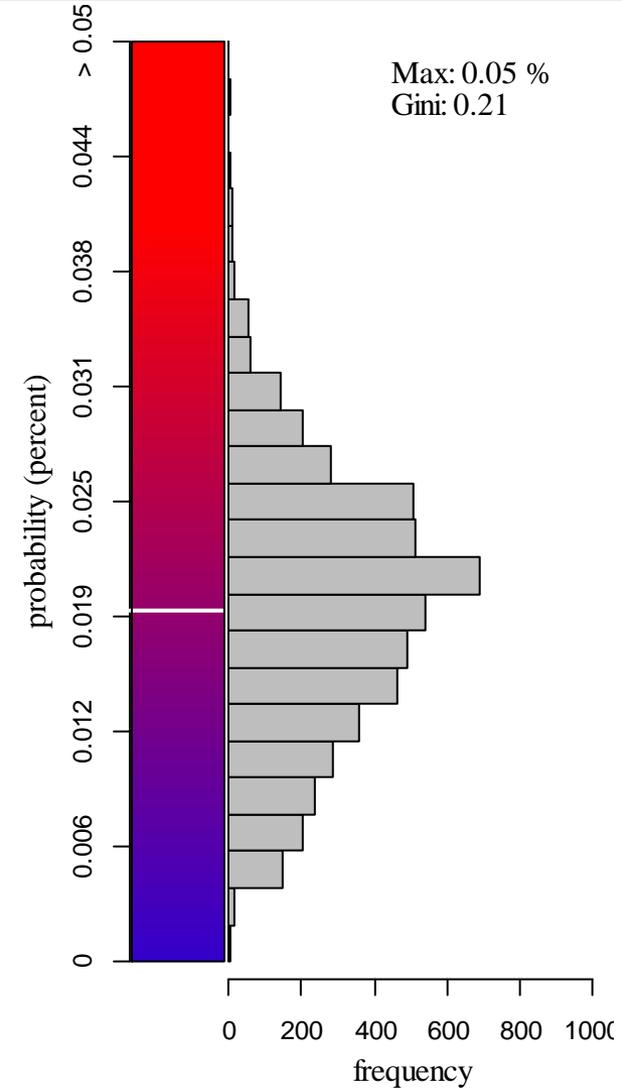
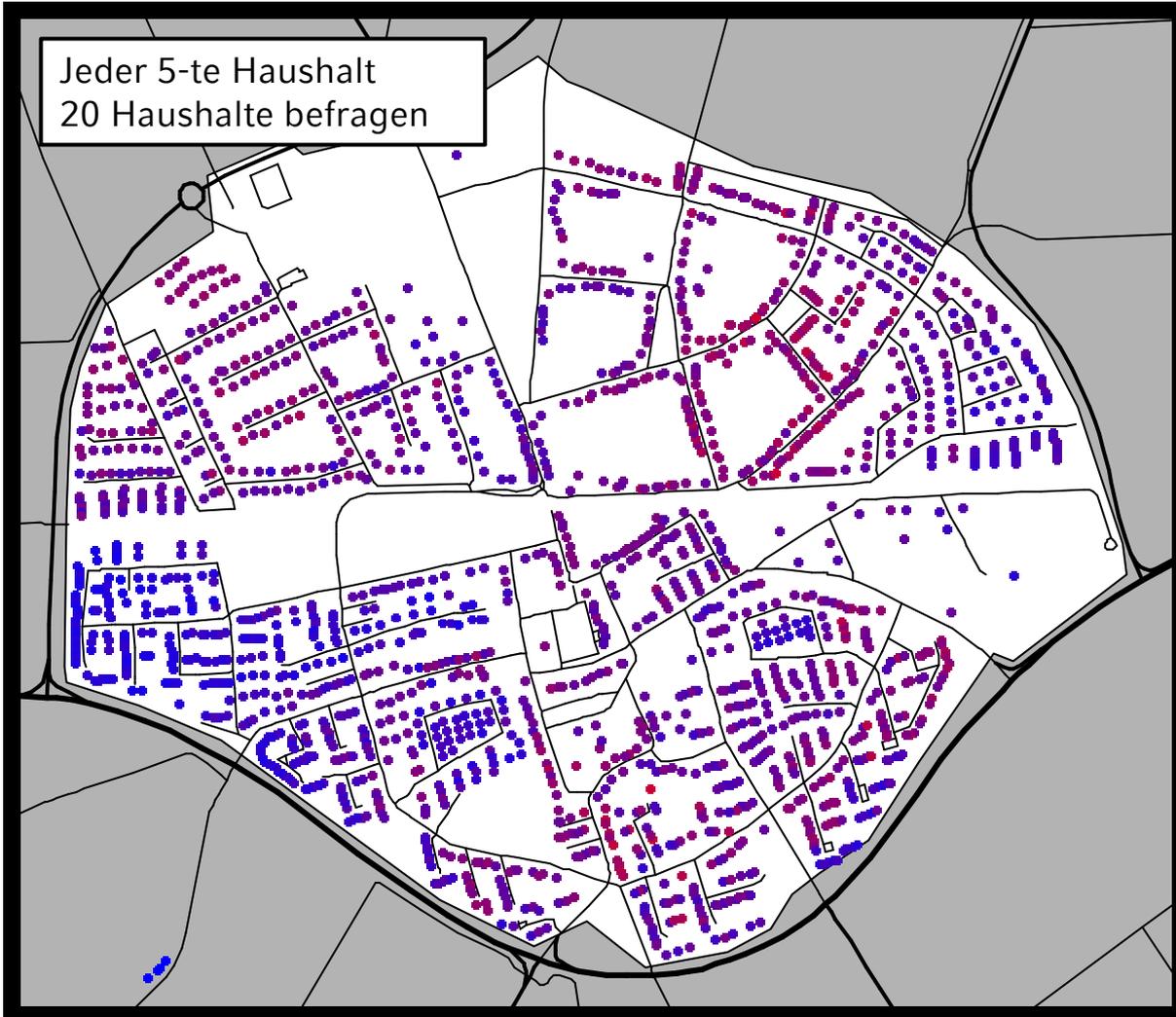


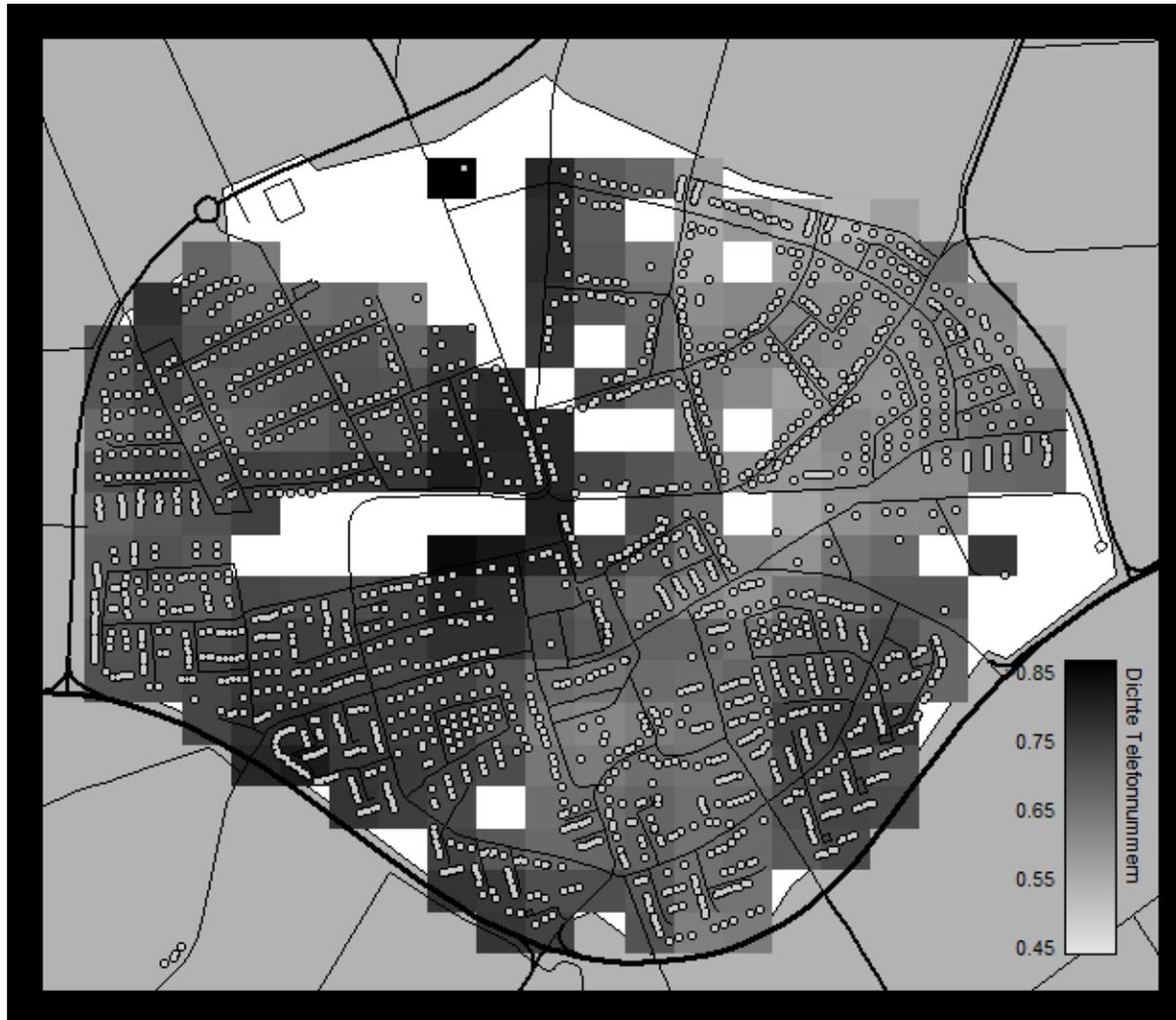
# Lösungsansätze

1. Gewichtung
2. Reine Zufallsroute
3. Straßenabschnitt Stichprobe



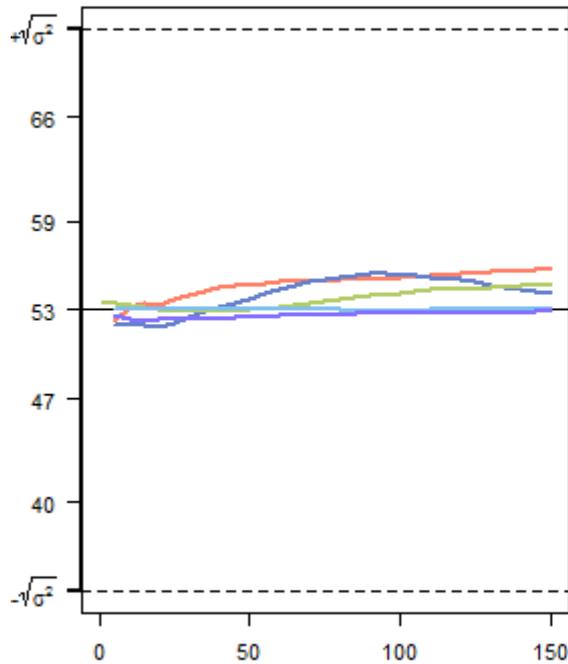




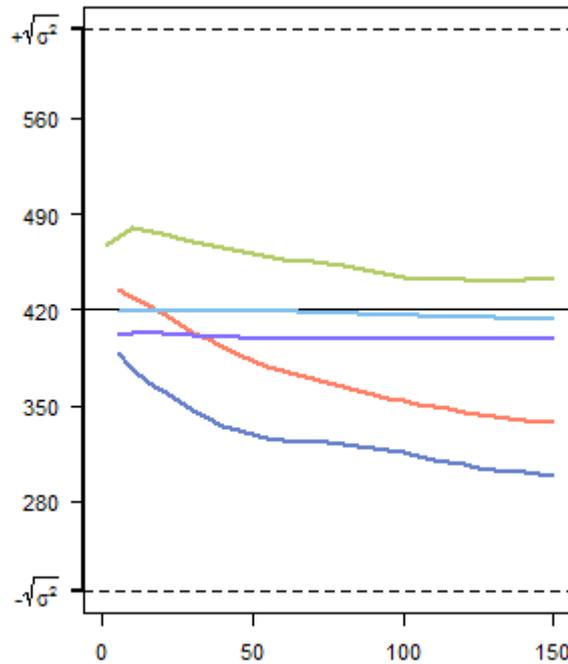




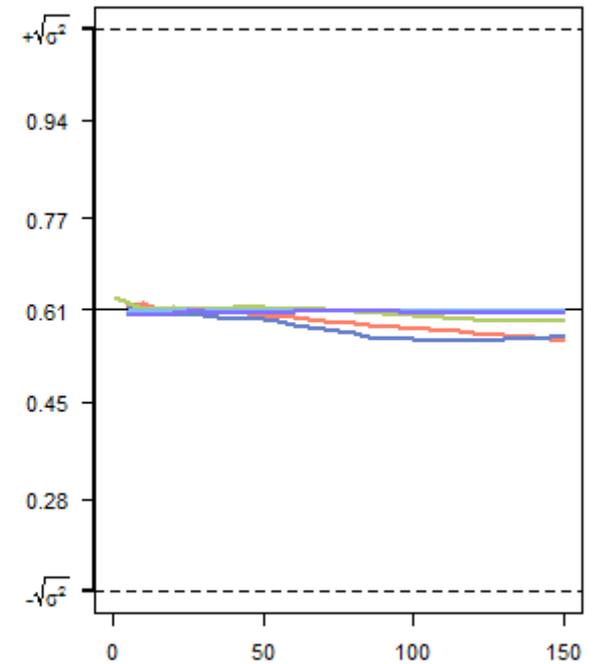
Alter (in Jahren)



Distanz zum nächsten  
Supermarkt (in Meter)



Anteil Verheiratete



Anzahl passierter Haushalte

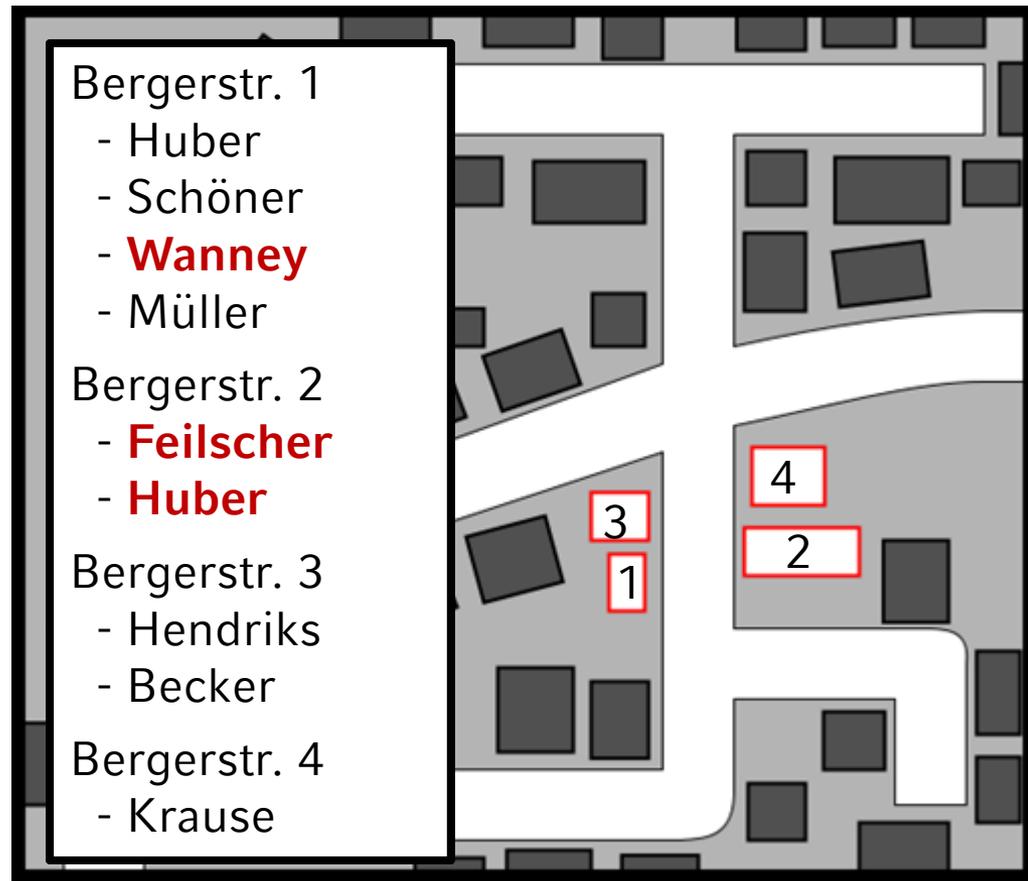
- Noelle Route (jeder 5te )
- TNS Infratest (jeder 5te )
- GFK Route (passierte HH)

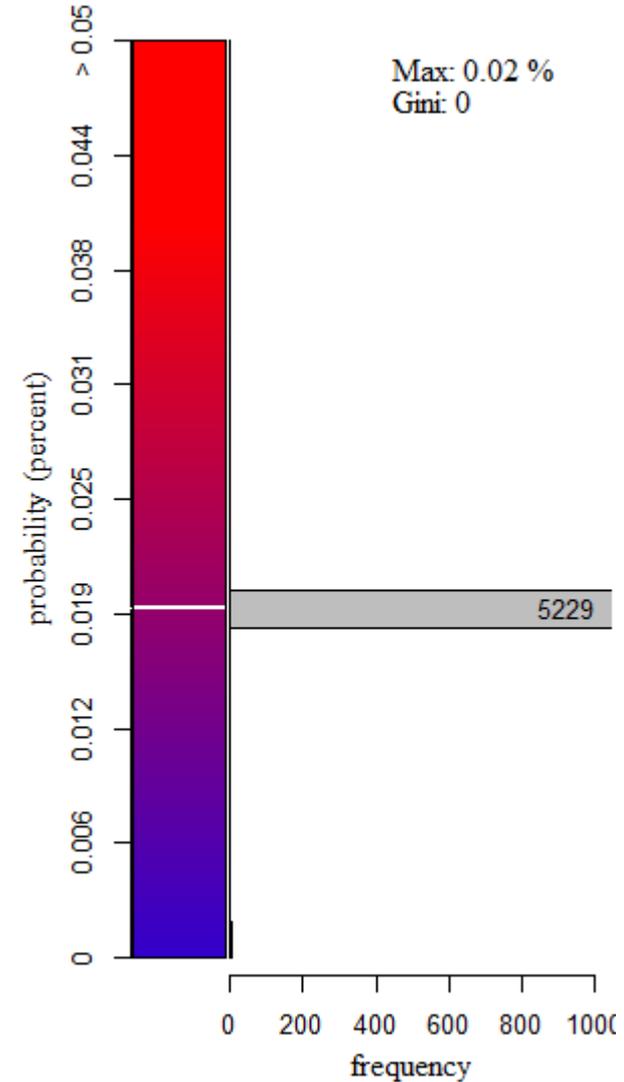
- Alle Haushalte als Start
- Telefonb.haush. als Start

## Straßenstück Stichprobe

1. Einteilen aller Straßen in Abschnitte zwischen zwei Kreuzungen
2. Zufällige Ziehung von Straßenstücken
3. Notieren aller Haushalte
4. X-Prozent Stichprobe aus allen notierten Haushalten

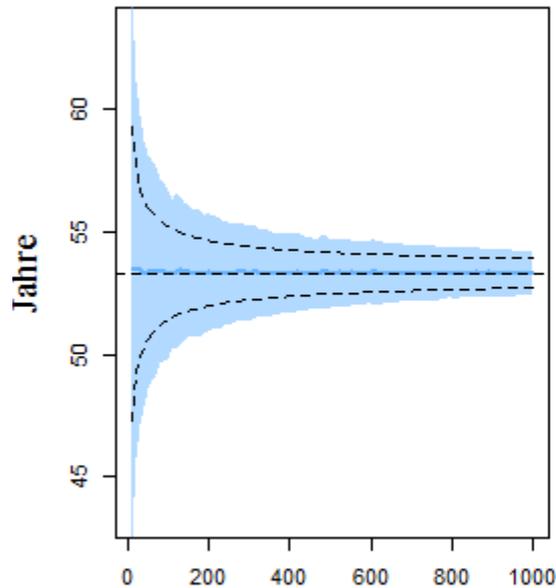
$$P(HH) = \frac{1}{n_{\text{Straßenstücke}}} \cdot \frac{X}{100}$$



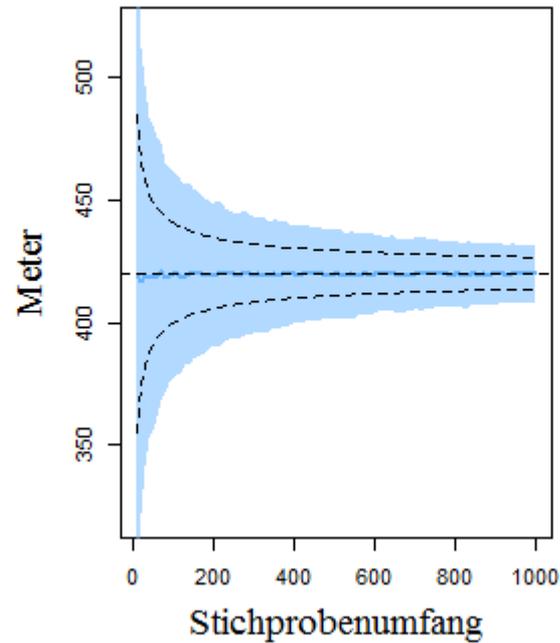




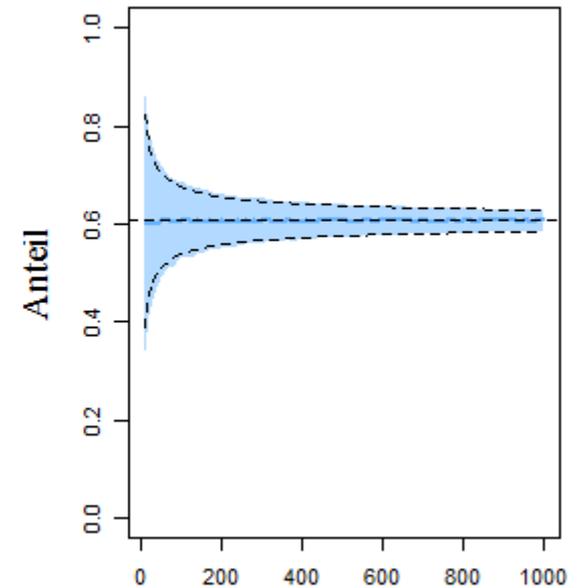
## Alter



## Distanz zum nächsten Supermarkt



## Anteil Verheiratete





**Danke für die Aufmerksamkeit**