The Prevalence of Unexcused School Absenteeism in Germany: An Individual Participant Data Meta-Analysis

Bernd Weiß
Research Institute for Sociology
GESIS-ZA Central Archive for Empirical Social Research
University of Cologne, Germany
bernd.weiss@wiso.uni-koeln.de

May, 18–20, 2009, Oslo
Outline

Introduction and Problem

Data and Methods

Empirical analyses

Conclusion
Outline

Introduction and Problem

Data and Methods

Empirical analyses

Conclusion
Frequent truancy is no peccadillo of youth

- The risk of social disadvantages for life
- Beginning of a criminal career
- Stress for teachers
- Need of medical treatment
- Structural school problems
- Problems within the family
Frequent truancy is no peccadillo of youth

- The risk of social disadvantages for life
- Beginning of a criminal career
- Stress for teachers
- Need of medical treatment
- Structural school problems
- Problems within the family
Frequent truancy is no peccadillo of youth

- The risk of social disadvantages for life
- Beginning of a criminal career
- Stress for teachers
  - Need of medical treatment
  - Structural school problems
  - Problems within the family
Frequent truancy is no peccadillo of youth

- The risk of social disadvantages for life
- Beginning of a criminal career
- Stress for teachers
- Need of medical treatment
- Structural school problems
- Problems within the family
Frequent truancy is no peccadillo of youth

- The risk of social disadvantages for life
- Beginning of a criminal career
- Stress for teachers
- Need of medical treatment
- Structural school problems
- Problems within the family
Frequent truancy is no peccadillo of youth

- The risk of social disadvantages for life
- Beginning of a criminal career
- Stress for teachers
- Need of medical treatment
- Structural school problems
- Problems within the family
Research questions

1. How many students do play truant in Germany?

2. What do we know about (some of) the risk factors? (in meta-analytical terms: Can we explain heterogeneity?)

In contrast to most meta-analysis,

- the dependent variable is not an effect size but a proportion.
- I am using individual participant data (IPD).
Research questions

1. How many students do play truant in Germany?

2. What do we know about (some of) the risk factors? (in meta-analytical terms: Can we explain heterogeneity?)

In contrast to most meta-analysis,

- the dependent variable is not an effect size but a proportion.
- I am using individual participant data (IPD).
Research questions

1. How many students do play truant in Germany?

2. What do we know about (some of) the risk factors? (in meta-analytical terms: Can we explain heterogeneity?)

In contrast to most meta-analysis,

- the dependent variable is not an effect size but a proportion.
- I am using individual participant data (IPD).
Research questions

1. How many students do play truant in Germany?

2. What do we know about (some of) the risk factors? (in meta-analytical terms: Can we explain heterogeneity?)

In contrast to most meta-analysis, . . .

▶ the dependent variable is not an effect size but a proportion.

▶ I am using individual participant data (IPD).
Research questions

1. How many students do play truant in Germany?

2. What do we know about (some of) the risk factors? (in meta-analytical terms: Can we explain heterogeneity?)

In contrast to most meta-analysis, . . .

- the dependent variable is not an effect size but a proportion.
- I am using individual participant data (IPD).
State of research


State of research


Outline

Introduction and Problem

Data and Methods

Empirical analyses

Conclusion
1. Datasets must contain information about unexcused absenteeism (reported from students, teachers or parents).

2. Datasets are available for secondary analysis/meta-analysis.

3. Samples were drawn in Germany but may contain students with Non-German citizenship.
Definition of study/dataset population

1. Datasets must contain information about unexcused absenteeism (reported from students, teachers or parents).

2. Datasets are available for secondary analysis/meta-analysis.

3. Samples were drawn in Germany but may contain students with Non-German citizenship.
Definition of study/dataset population

1. Datasets must contain information about unexcused absenteeism (reported from students, teachers or parents).

2. Datasets are available for secondary analysis/meta-analysis.

3. Samples were drawn in Germany but may contain students with Non-German citizenship.
Data description and quality of data

- 15 datasets ≈ information about 77,000 students (unit of analysis).

- Datasets vary with respect to:
  - the unit of observation (students, teachers, parents)
  - the geographical units (local sample versus national sample)
  - the year of sampling (1971 to 2005)
  - the pupils’ grade (5th to 10th)
  - the pupils’ school type (“Haupt-”, “Real-”, “Gesamtschule”, “Gymnasien”).

- The pooled dataset is biased towards pupils who attend the "Haupt-" and "Gesamtschule" and who are in the 8th, 9th, and 10th grade. It is well known that among these pupils the risk to play truant is higher.

Conclusion: The estimated prevalence is likely to be upwards biased.
Data description and quality of data

- 15 datasets $\approx$ information about 77,000 students (unit of analysis).

- Datasets vary with respect to
  - the unit of observation (students, teacher, parents),
  - the geographical units (local sample versus national sample),
  - the year of sampling (1971 to 2005),
  - the pupils’ grade (5th to 10th),
  - the pupils’ school type ("Haupt-", "Realschule", "Gesamtschulen" and "Gymnasien").

- The pooled dataset is biased towards pupils who attend the "Haupt-" and "Gesamtschule" and who are in the 8th, 9th and 10th grade. It is well known that among these pupils the risk to play truant is higher.

Conclusion: The estimated prevalence is likely to be upwards biased.
Data description and quality of data

- 15 datasets ≈ information about 77,000 students (unit of analysis).

- Datasets vary with respect to
  - the unit of observation (students, teacher, parents)
  - the geographical units (local sample versus national sample),
  - the year of sampling (1971 to 2005),
  - the pupils’ grade (5th to 10th),
  - the pupils’ school type ("Haupt-", "Real-", "Gesamtschulen" and "Gymnasien").

- The pooled dataset is biased towards pupils who attend the "Haupt-" and "Gesamtschule" and who are in the 8th, 9th, and 10th grade. It is well known that among these pupils the risk to play truant is higher.

Conclusion: The estimated prevalence is likely to be upwards biased.
Data description and quality of data

- 15 datasets ≈ information about 77,000 students (unit of analysis).

- Datasets vary with respect to:
  - the unit of observation (students, teacher, parents)
  - the geographical units (local sample versus national sample),
  - the year of sampling (1971 to 2005),
  - the pupils’ grade (5th to 10th),
  - the pupils’ school type ("Haupt-", "Real-", "Gesamtschulen" and "Gymnasien").

- The pooled dataset is biased towards pupils who attend the "Haupt-" and "Gesamtschule" and who are in the 8th, 9th and 10th grade. It is well known that among these pupils the risk to play truant is higher.

Conclusion: The estimated prevalence is likely to be upwards biased.
Data description and quality of data

- 15 datasets ≈ information about 77,000 students (unit of analysis).

- Datasets vary with respect to
  - the unit of observation (students, teacher, parents)
  - the geographical units (local sample versus national sample),
  - the year of sampling (1971 to 2005),
  - the pupils’ grade (5th to 10th),
  - the pupils’ school type ("Haupt-", "Real-", "Gesamtschulen" and "Gymnasien").

- The pooled dataset is biased towards pupils who attend the "Haupt-" and "Gesamtschule" and who are in the 8th, 9th, and 10th grade. It is well known that among these pupils the risk to play truant is higher.

Conclusion: The estimated prevalence is likely to be upwards biased.
Data description and quality of data

▶ 15 datasets ≈ information about 77,000 students (unit of analysis).

▶ Datasets vary with respect to
  ▶ the unit of observation (students, teacher, parents)
  ▶ the geographical units (local sample versus national sample),
  ▶ the year of sampling (1971 to 2005),
  ▶ the pupils’ grade (5th to 10th),
  ▶ the pupils’ school type ("Haupt-", "Real-", "Gesamtschulen" and "Gymnasien").

▶ The pooled dataset is biased towards pupils who attend the "Haupt-" and "Gesamtschule" and who are in the 8th, 9th, and 10th grade. It is well known that among these pupils the risk to play truant is higher.

Conclusion: The estimated prevalence is likely to be upwards biased.
Data description and quality of data

- 15 datasets ≈ information about 77,000 students (unit of analysis).

- Datasets vary with respect to
  - the unit of observation (students, teacher, parents)
  - the geographical units (local sample versus national sample),
  - the year of sampling (1971 to 2005),
  - the pupils’ grade (5th to 10th),
  - the pupils’ school type ("Haupt-", "Real-", "Gesamtschulen" and "Gymnasien").

- The pooled dataset is biased towards pupils who attend the "Haupt-" and "Gesamtschule" and who are in the 8th, 9th, and 10th grade. It is well known that among these pupils the risk to play truant is higher.

Conclusion: The estimated prevalence is likely to be upwards biased.
Data description and quality of data

- 15 datasets ≈ information about 77,000 students (unit of analysis).
- Datasets vary with respect to:
  - the unit of observation (students, teacher, parents)
  - the geographical units (local sample versus national sample),
  - the year of sampling (1971 to 2005),
  - the pupils’ grade (5th to 10th),
  - the pupils’ school type ("Haupt-", "Real-", "Gesamtschulen" and "Gymnasien").

- The pooled dataset is biased towards pupils who attend the "Haupt-" and "Gesamtschule" and who are in the 8th, 9th and 10th grade. It is well known that among these pupils the risk to play truant is higher.

Conclusion: The estimated prevalence is likely to be upwards biased.

Bernd Weiß
University of Cologne
Data description and quality of data

- 15 datasets ≈ information about 77,000 students (unit of analysis).

- Datasets vary with respect to:
  - the unit of observation (students, teacher, parents)
  - the geographical units (local sample versus national sample),
  - the year of sampling (1971 to 2005),
  - the pupils’ grade (5th to 10th),
  - the pupils’ school type ("Haupt-", "Real-", "Gesamtschulen" and "Gymnasien").

- The pooled dataset is biased towards pupils who attend the "Haupt-" and "Gesamtschule" and who are in the 8th, 9th and 10th grade. It is well known that among these pupils the risk to play truant is higher.

Conclusion: The estimated prevalence is likely to be upwards biased.
Data description and quality of data

- 15 datasets ≈ information about 77,000 students (unit of analysis).
- Datasets vary with respect to:
  - the unit of observation (students, teacher, parents)
  - the geographical units (local sample versus national sample),
  - the year of sampling (1971 to 2005),
  - the pupils’ grade (5th to 10th),
  - the pupils’ school type ("Haupt-", "Real-", "Gesamtschulen" and "Gymnasien").
- The pooled dataset is biased towards pupils who attend the "Haupt-" and "Gesamtschule" and who are in the 8th, 9th and 10th grade. It is well known that among these pupils the risk to play truant is higher.

Conclusion: The estimated prevalence is likely to be upwards biased.
Operationalization

Dependent variable: Truancy

- Truancy is defined as unexcused absence from school.
- Two types of truancy:
  - Moderate truancy: 0 = never played truant; 1 = at least once in his/her lifetime
  - Frequent truancy: 0 = never/seldom played truant; 1 = played truant frequently

Independent variables

- Ethnicity (0/1)
- Female (0/1)
- "Hauptschule" (lowest form of schooling; 0/1)
- Grade and grade squared (5–10)
- Unit of observation (0 = {teacher, parents}, 1 = students)
- Observation period (in days)
- Number of categories (truancy variable)
Operationalization

Dependent variable: Truancy

- Truancy is defined as unexcused absence from school.
- Two types of truancy:
  - Moderate truancy: 0 = never played truant; 1 = at least once in his/her lifetime
  - Frequent truancy: 0 = never/seldom played truant; 1 = played truant frequently

Independent variables

- Ethnicity (0/1)
- Female (0/1)
- "Hauptschule" (lowest form of schooling; 0/1)
- Grade and grade squared (5–10)
- Unit of observation (0 = {teacher, parents}, 1 = students)
- Observation period (in days)
- Number of categories (truancy variable)
Operationalization

Dependent variable: Truancy

- Truancy is defined as unexcused absence from school.
- Two types of truancy:
  - Moderate truancy: $0 = \text{never played truant}; \ 1 = \text{at least once in his/her lifetime}$
  - Frequent truancy: $0 = \text{never/seldom played truant}; \ 1 = \text{played truant frequently}$

Independent variables

- Ethnicity (0/1)
- Female (0/1)
- "Hauptschule" (lowest form of schooling; 0/1)
- Grade and grade squared (5–10)
- Unit of observation (0 = {teacher, parents}, 1 = students)
- Observation period (in days)
- Number of categories (truancy variable)
Operationalization

Dependent variable: Truancy

- Truancy is defined as unexcused absence from school.
- Two types of truancy:
  - Moderate truancy: 0 = never played truant; 1 = at least once in his/her lifetime
  - Frequent truancy: 0 = never/seldom played truant; 1 = played truant frequently

Independent variables

- Ethnicity (0/1)
- Female (0/1)
- "Hauptschule" (lowest form of schooling; 0/1)
- Grade and grade squared (5–10)
- Unit of observation (0 = {teacher, parents}, 1 = students)
- Observation period (in days)
- Number of categories (truancy variable)
Operationalization

Dependent variable: Truancy

- Truancy is defined as unexcused absence from school.
- Two types of truancy:
  - Moderate truancy: 0 = never played truant; 1 = at least once in his/her lifetime
  - Frequent truancy: 0 = never/seldom played truant; 1 = played truant frequently

Independent variables

- Ethnicity (0/1)
- Female (0/1)
- "Hauptschule" (lowest form of schooling; 0/1)
- Grade and grade squared (5–10)
- Unit of observation (0 = {teacher, parents}, 1 = students)
- Observation period (in days)
- Number of categories (truancy variable)
### Operationalization

#### Dependent variable: Truancy
- Truancy is defined as unexcused absence from school.
- Two types of truancy:
  - Moderate truancy: 0 = never played truant; 1 = at least once in his/her lifetime
  - Frequent truancy: 0 = never/seldom played truant; 1 = played truant frequently

#### Independent variables
- Ethnicity (0/1)
- Female (0/1)
- "Hauptschule" (lowest form of schooling; 0/1)
- Grade and grade squared (5–10)
- Unit of observation (0 = {teacher, parents}, 1 = students)
- Observation period (in days)
- Number of categories (truancy variable)
Operationalization

Dependent variable: Truancy

- Truancy is defined as unexcused absence from school.
- Two types of truancy:
  - Moderate truancy: 0 = never played truant; 1 = at least once in his/her lifetime
  - Frequent truancy: 0 = never/seldom played truant; 1 = played truant frequently

Independent variables

- Ethnicity (0/1)
- Female (0/1)
- "Hauptschule" (lowest form of schooling; 0/1)
- Grade and grade squared (5–10)
- Unit of observation (0 = {teacher, parents}, 1 = students)
- Observation period (in days)
- Number of categories (truancy variable)
## Operationalization

### Dependent variable: Truancy
- Truancy is defined as unexcused absence from school.
- Two types of truancy:
  - Moderate truancy: 0 = never played truant; 1 = at least once in his/her lifetime
  - Frequent truancy: 0 = never/seldom played truant; 1 = played truant frequently

### Independent variables
- Ethnicity (0/1)
- Female (0/1)
- "Hauptschule" (lowest form of schooling; 0/1)
- Grade and grade squared (5–10)
- Unit of observation (0 = {teacher, parents}, 1 = students)
- Observation period (in days)
- Number of categories (truancy variable)
Operationalization

**Dependent variable: Truancy**

- Truancy is defined as unexcused absence from school.
- Two types of truancy:
  - Moderate truancy: 0 = never played truant; 1 = at least once in his/her lifetime
  - Frequent truancy: 0 = never/seldom played truant; 1 = played truant frequently

**Independent variables**

- Ethnicity (0/1)
- Female (0/1)
- "'Hauptschule'" (lowest form of schooling; 0/1)
- Grade and grade squared (5–10)
  - Unit of observation (0 = {teacher, parents}, 1 = students)
  - Observation period (in days)
  - Number of categories (truancy variable)
### Operationalization

**Dependent variable: Truancy**

- Truancy is defined as unexcused absence from school.
- Two types of truancy:
  - Moderate truancy: 0 = never played truant; 1 = at least once in his/her lifetime
  - Frequent truancy: 0 = never/seldom played truant; 1 = played truant frequently

**Independent variables**

- Ethnicity (0/1)
- Female (0/1)
- "‘Hauptschule’" (lowest form of schooling; 0/1)
- Grade and grade squared (5–10)
- Unit of observation (0 = {teacher, parents}, 1 = students)
- Observation period (in days)
- Number of categories (truancy variable)
Operationalization

Dependent variable: Truancy

- Truancy is defined as unexcused absence from school.
- Two types of truancy:
  - Moderate truancy: 0 = never played truant; 1 = at least once in his/her lifetime
  - Frequent truancy: 0 = never/seldom played truant; 1 = played truant frequently

Independent variables

- Ethnicity (0/1)
- Female (0/1)
- "’Hauptschule’" (lowest form of schooling; 0/1)
- Grade and grade squared (5–10)
- Unit of observation (0 = {teacher, parents}, 1 = students)
- Observation period (in days)
- Number of categories (truancy variable)
Operationalization

Dependent variable: Truancy

- Truancy is defined as unexcused absence from school.
- Two types of truancy:
  - Moderate truancy: 0 = never played truant; 1 = at least once in his/her lifetime
  - Frequent truancy: 0 = never/seldom played truant; 1 = played truant frequently

Independent variables

- Ethnicity (0/1)
- Female (0/1)
- "‘Hauptschule’" (lowest form of schooling; 0/1)
- Grade and grade squared (5–10)
- Unit of observation (0 = {teacher, parents}, 1 = students)
- Observation period (in days)
- Number of categories (truancy variable)
Example from PISA 2000

Q29 → How many times in the previous two school weeks did you:

(Please <tick> only one box on each row.)

- → → → 5
  None → 1 or 2 → 3 or 4 → or more
  a) miss school? 

Moderate truancy: 0 = None, 1 = {1 or 2, 3 or 4, 5 or more}
Frequent truancy: 0 = {None, 1 or 2, 3 or 4}, 1 = 5 or more
Observation period = 14 days
Number of categories: 4
Example from PISA 2000

Q.29 How many times in the previous two school weeks did you:

(Please <tick> only one box on each row.)

- Moderate truancy: 0 = None, 1 = {1 or 2, 3 or 4, 5 or more}
- Frequent truancy: 0 = {None, 1 or 2, 3 or 4}, 1 = 5 or more
- Observation period = 14 days
- Number of categories: 4
Example from PISA 2000

Q.29: How many times in the previous two school weeks did you:

(Please tick only one box on each row.)

Moderate truancy: 0 = None, 1 = {1 or 2, 3 or 4, 5 or more}
Frequent truancy: 0 = {None, 1 or 2, 3 or 4}, 1 = 5 or more
Observation period = 14 days
Number of categories: 4
Example from PISA 2000

Q.29 → How many times in the previous two school weeks did you:

(Please <tick> only one box on each row.)

None → 1 or 2 → 3 or 4 → or more

a) Miss school?

- Moderate truancy: 0 = None, 1 = {1 or 2, 3 or 4, 5 or more}
- Frequent truancy: 0 = {None, 1 or 2, 3 or 4}, 1 = 5 or more
- Observation period = 14 days
- Number of categories: 4
Methodology: IPD and APD meta-analysis

Meta-analysis using aggregate participant data (APD)

▶ Only for illustrative purposes: forest plots

Meta-analysis using individual participant data (IPD)

Multilevel models (= random effects models)

▶ to estimate an overall estimator of the prevalence of truancy and
▶ to conduct heterogeneity analyses.
Methodology: IPD and APD meta-analysis

Meta-analysis using aggregate participant data (APD)

- Only for illustrative purposes: forest plots

Meta-analysis using individual participant data (IPD)

Multilevel models (= random effects models)

- to estimate an overall estimator of the prevalence of truancy and
- to conduct heterogeneity analyses.
Methodology: IPD and APD meta-analysis

Meta-analysis using aggregate participant data (APD)

- Only for illustrative purposes: forest plots

Meta-analysis using individual participant data (IPD)

Multilevel models (= random effects models)
- to estimate an overall estimator of the prevalence of truancy and
- to conduct heterogeneity analyses.
Methodology: IPD and APD meta-analysis

Meta-analysis using aggregate participant data (APD)

- Only for illustrative purposes: forest plots

Meta-analysis using individual participant data (IPD)

Multilevel models (= random effects models)

- to estimate an overall estimator of the prevalence of truancy and
- to conduct heterogeneity analyses.
Outline

Introduction and Problem

Data and Methods

Empirical analyses

Conclusion
Diagnostics

- No indications for "publication bias" (Egger-test, Rang-correlation test).
- Log-odds of moderate truancy do not differ from a normal distribution (Lilliefors-, Pearson-, Shapiro-Francia-Test).
- Log-odds of frequent truancy are slightly rightward biased.
Diagnostics

- No indications for "publication bias" (Egger-test, Rang-correlation test).
- Log-odds of moderate truancy do not differ from a normal distribution (Lilliefors-, Pearson-, Shapiro-Francia-Test).
- Log-odds of frequent truancy are slightly rightward biased.
Diagnostics

- No indications for "publication bias" (Egger-test, Rang-correlation test).
- Log-odds of moderate truancy do not differ from a normal distribution (Lilliefors-, Pearson-, Shapiro-Francia-Test).
- Log-odds of frequent truancy are slightly rightward biased.
Empirical analyses
Overall estimate
Explaining heterogeneity
APD/IPD meta-analysis for moderate truancy

Study
LSB2005 (4%)
PAK–KID1994/P (6%)
KHB2003 (7%)
DWD2002 (10%)
PISA2000 (11%)
ESPAD2003 (11%)
PAK–KID1994/Y (12%)
NRW–KIDS2001 (15%)
KLS1993 (21%)
KJS1980 (22%)
S1990/A (24%)
S1990/B (27%)
MPI1999 (38%)
KFN2000 (52%)
AVS1971 (54%)

Overall (FEM) (22%)
Overall (REM) (17%)

APD (k = 15):
▶ Only REM appropriate
▶ Overall prevalence (REM): 17.11%

IPD (Null model; N = 77469, k = 15):
▶ Overall prevalence (MLM): 16.94%
▶ \( \tau^2 = 0.93 \)
APD/IPD meta-analysis for moderate truancy

Study
LSB2005 (4%)
PAK−KID1994/P (6%)
KHB2003 (7%)
DWD2002 (10%)
PISA2000 (11%)
ESPAD2003 (11%)
PAK−KID1994/Y (12%)
NRW−KIDS2001 (15%)
KLS1993 (21%)
KJS1980 (22%)
S1990/A (24%)
S1990/B (27%)
MPI1999 (38%)
KFN2000 (52%)
AVS1971 (54%)

Overall (FEM) (22%)
Overall (REM) (17%)

APD (k = 15):
- Only REM appropriate
- Overall prevalence (REM): 17.11%
- $\tau^2 = 1.26$

IPD (Null model; N = 77469, k = 15):
- Overall prevalence (MLM): 16.94%
- $\tau^2 = 0.93$

Bernd Weiß
University of Cologne
APD/IPD meta-analysis for moderate truancy

Study
LSB2005 (4%)
PAK–KID1994/P (6%)
KHB2003 (7%)
DWD2002 (10%)
PISA2000 (11%)
ESPAD2003 (11%)
PAK–KID1994/Y (12%)
NRW–KIDS2001 (15%)
KLS1993 (21%)
KJS1980 (22%)
S1990/A (24%)
S1990/B (27%)
MPI1999 (38%)
KFN2000 (52%)
AVS1971 (54%)

Overall (FEM) (22%)
Overall (REM) (17%)

APD (k = 15):
► Only REM appropriate
► Overall prevalence (REM): 17.11%
► $\tau^2 = 1.26$

IPD (Null model; N = 77469, k = 15):
► Overall prevalence (MLM): 16.94%
► $\tau^2 = 0.93$
APD/IPD meta-analysis for frequent truancy

**Study**
- LSB2005 (0%)
- ESPAD2003 (1%)
- PAK–KID1994/P (1%)
- NRW–KIDS2001 (1%)
- S1990/A (1%)
- KJS1980 (1%)
- AVS1971 (1%)
- KHB2003 (1%)
- PISA2000 (1%)
- S1990/B (2%)
- PAK–KID1994/Y (2%)
- DWD2002 (3%)
- KFN2000 (6%)
- KLS1993 (21%)
- MPI1999 (38%)

**Overall (FEM) (7%)**
**Overall (REM) (2%)**

**APD (k = 15):**
- Only REM appropriate
- Overall prevalence (REM): 1.89%
- $\tau^2 = 3.21$

**IPD (Null model; N = 77469, k = 15):**
- Overall prevalence (MLM): 1.86%
- $\tau^2 = 2.03$

Bernd Weiß
University of Cologne
APD/IPD meta-analysis for frequent truancy

Study
- LSB2005 (0%)
- ESPAD2003 (1%)
- PAK–KID1994/P (1%)
- NRW–KIDS2001 (1%)
- S1990/A (1%)
- KJS1980 (1%)
- AVS1971 (1%)
- KHB2003 (1%)
- PISA2000 (1%)
- S1990/B (2%)
- PAK–KID1994/Y (2%)
- DWD2002 (3%)
- KFN2000 (6%)
- KLS1993 (21%)
- MPI1999 (38%)

Overall (FEM) (7%)
Overall (REM) (2%)

APD (k = 15):
- Only REM appropriate
- Overall prevalence (REM): 1.89%
- $\tau^2 = 3.21$

IPD (Null model; N = 77469, k = 15):
- Overall prevalence (MLM): 1.86%
- $\tau^2 = 2.03$

Bernd Weiß
University of Cologne
APD/IPD meta-analysis for frequent truancy

**Study**
- LSB2005 (0%)
- ESPAD2003 (1%)
- PAK–KID1994/P (1%)
- NRW–KIDS2001 (1%)
- S1990/A (1%)
- KJS1980 (1%)
- AVS1971 (1%)
- KHB2003 (1%)
- PISA2000 (1%)
- S1990/B (2%)
- PAK–KID1994/Y (2%)
- DWD2002 (3%)
- KFN2000 (6%)
- KLS1993 (21%)
- MPI1999 (38%)

**Overall (FEM) (7%)**
**Overall (REM) (2%)**

**APD (k = 15):**
- Only REM appropriate
- Overall prevalence (REM): 1.89%
  \[ \tau^2 = 3.21 \]

**IPD (Null model; N = 77469, k = 15):**
- Overall prevalence (MLM): 1.86%
  \[ \tau^2 = 2.03 \]
Outline

Empirical analyses
  Overall estimate
  Explaining heterogeneity
### Results of heterogeneity analyses

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Moderate truancy</th>
<th>Frequent truancy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>APD</td>
<td>IPD</td>
</tr>
<tr>
<td><strong>Predictor at the participant level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hauptschule</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Grade</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Grade(^2)</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Female</td>
<td>0</td>
<td>−</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ethnicity × Hauptschule</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td><strong>Predictors at the dataset level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit of observation (pupils = 1)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Number of categories (truancy)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Observation period (&gt; 6 month / 1 year)</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>Average grade</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>Prop. pupils attending &quot;Hauptschule&quot;</td>
<td>−</td>
<td>0</td>
</tr>
<tr>
<td>Year of sampling</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

−: negative effect \((p \leq 0.10)\)  
0: no effect \((p > 0.10)\)  
+: positive effect \((p \leq 0.10)\)
Outline

Introduction and Problem

Data and Methods

Empirical analyses

Conclusion
Conclusion

1. How many students do play truant in Germany?
   - About 17% of the students play truant moderately, about 2% play truant frequently.
   - Due to heterogeneity issues, the overall prevalence must be seen with caution!

2. What do we know about (some of) the risk factors?
   - A considerable amount of between-study heterogeneity is caused by compositional effects.
   - At the participant level: Female (+), Hauptschule (+), grade (+) and ethnicity (+).

3. Other sources of heterogeneity are study design and operationalization of truancy:
   - Unit of observation: Students (+)
   - Observation period (+)
   - Average grade (+)
Conclusion

1. How many students do play truant in Germany?
   ▶ About 17% of the students play truant moderately, about 2% play truant frequently.
   ▶ Due to heterogeneity issues, the overall prevalence must be seen with caution!

2. What do we know about (some of) the risk factors?
   ▶ A considerable amount of between-study heterogeneity is caused by compositional effects.
   ▶ At the participant level: Female (+), Hauptschule (+), grade (+) and ethnicity (+).

3. Other sources of heterogeneity are study design and operationalization of truancy:
   ▶ Unit of observation: Students (+)
   ▶ Observation period (+)
   ▶ Average grade (+)
Conclusion

1. How many students do play truant in Germany?
   ▶ About 17% of the students play truant moderately, about 2% play truant frequently.
   ▶ Due to heterogeneity issues, the overall prevalence must be seen with caution!

2. What do we know about (some of) the risk factors?
   ▶ A considerable amount of study between-study heterogeneity is caused by compositional effects.
   ▶ At the participant level: Female (+), Hauptschule (+), grade (+) and ethnicity (+).

3. Other sources of heterogeneity are study design and operationalization of truancy:
   ▶ Unit of observation: Students (+)
   ▶ Observation period (+)
   ▶ Average grade (+)
Conclusion

1. How many students do play truant in Germany?
   - About 17% of the students play truant moderately, about 2% play truant frequently.
   - Due to heterogeneity issues, the overall prevalence must be seen with caution!

2. What do we know about (some of) the risk factors?
   - A considerable amount of between-study heterogeneity is caused by compositional effects (IPD).
   - At the participant level: Female (−), Hauptschule (+), grade (+) and ethnicity (+).

3. Other sources of heterogeneity are study design and operationalization of truancy:
   - Unit of observation: Students (+)
   - Observation period (+)
   - Average grade (+)
Conclusion

1. How many students do play truant in Germany?
   ▶ About 17% of the students play truant moderately, about 2% play truant frequently.
   ▶ Due to heterogeneity issues, the overall prevalence must be seen with caution!

2. What do we know about (some of) the risk factors?
   ▶ A considerable amount of between-study heterogeneity is caused by compositional effects (IPD).
   ▶ At the participant level: Female (−), Hauptschule (+), grade (+) and ethnicity (+).

3. Other sources of heterogeneity are study design and operationalization of truancy:
   ▶ Unit of observation: Students (+)
   ▶ Observation period (+)
   ▶ Average grade (+)
Conclusion

1. How many students do play truant in Germany?
   ▶ About 17% of the students play truant moderately, about 2% play truant frequently.
   ▶ Due to heterogeneity issues, the overall prevalence must be seen with caution!

2. What do we know about (some of) the risk factors?
   ▶ A considerable amount of between-study heterogeneity is caused by compositional effects (IPD).
   ▶ At the participant level: Female (−), Hauptschule (+), grade (+) and ethnicity (+).

3. Other sources of heterogeneity are study design and operationalization of truancy:
   ▶ Unit of observation: Students (+)
   ▶ Observation period (+)
   ▶ Average grade (+)
Conclusion

1. How many students do play truant in Germany?
   - About 17% of the students play truant moderately, about 2% play truant frequently.
   - Due to heterogeneity issues, the overall prevalence must be seen with caution!

2. What do we know about (some of) the risk factors?
   - A considerable amount of between-study heterogeneity is caused by compositional effects (IPD).
   - At the participant level: Female (−), Hauptschule (+), grade (+) and ethnicity (+).

3. Other sources of heterogeneity are study design and operationalization of truancy:
   - Unit of observation: Students (+)
   - Observation period (+)
   - Average grade (+)
Conclusion

1. How many students do play truant in Germany?
   ▶ About 17% of the students play truant moderately, about 2% play truant frequently.
   ▶ Due to heterogeneity issues, the overall prevalence must be seen with caution!

2. What do we know about (some of) the risk factors?
   ▶ A considerable amount of between-study heterogeneity is caused by compositional effects (IPD).
   ▶ At the participant level: Female (−), Hauptschule (+), grade (+) and ethnicity (+).

3. Other sources of heterogeneity are study design and operationalization of truancy:
   ▶ Unit of observation: Students (+)
     ▶ Observation period (+)
     ▶ Average grade (+)
Conclusion

1. How many students do play truant in Germany?
   - About 17% of the students play truant moderately, about 2% play truant frequently.
   - Due to heterogeneity issues, the overall prevalence must be seen with caution!

2. What do we know about (some of) the risk factors?
   - A considerable amount of between-study heterogeneity is caused by compositional effects (IPD).
   - At the participant level: Female (−), Hauptschule (+), grade (+) and ethnicity (+).

3. Other sources of heterogeneity are study design and operationalization of truancy:
   - Unit of observation: Students (+)
   - Observation period (+)
   - Average grade (+)
Conclusion

1. How many students do play truant in Germany?
   ▶ About 17% of the students play truant moderately, about 2% play truant frequently.
   ▶ Due to heterogeneity issues, the overall prevalence must be seen with caution!

2. What do we know about (some of) the risk factors?
   ▶ A considerable amount of between-study heterogeneity is caused by compositional effects (IPD).
   ▶ At the participant level: Female (−), Hauptschule (+), grade (+) and ethnicity (+).

3. Other sources of heterogeneity are study design and operationalization of truancy:
   ▶ Unit of observation: Students (+)
   ▶ Observation period (+)
   ▶ Average grade (+)
Recommendations for further research

- A common definition/operationalization of truancy.
- More longitudinal data.
- ...
Recommendations for further research

- A common definition/operationalization of truancy.
- More longitudinal data.
- ...
Recommendations for further research

- A common definition/operationalization of truancy.
- More longitudinal data.
- ...
Thank you very much for your attention!
Data retrieval process

- Identified potentially relevant publications or datasets, obtained more detailed information (n=37)

- Publications or datasets available (n=25)
  - Publications or datasets out of date or not available (n=12)
  - Delivery of data refused (n=2)
  - Excluded publications or datasets for not meeting inclusion criteria (n=10)

- Original datasets from research projects available (n=13)

- Split up two datasets; basis for a meta-analysis of the prevalence of truancy (n=15)

- Basis for a meta-analysis on the relation between migration background and truancy (n=12)

- No measures of migration background included (n=3)