Supplement

Poverty and crime in 19th century Germany

This serves as a supplement to the published paper *Poverty and Crime in 19th Century Germany*, by Halvor Mehlum, Edward Miguel and Ragnar Torvik (2006). It outlines minor discrepancies found in data replication which do not, as a whole, affect the reported data analysis published in the paper. See *README.txt* for a detailed update. The tables below come from the output generated with the replication code in *final_do*.

We have replicated the results for the paper, and we show here the replicated tables with any adjustment. There were several typos in the original version, as well as some minor discrepancies between the reproduced and published versions due to changes in STATA over time. Color code: rounding discrepancy, significance discrepancy, typo.

The data replication generated minor discrepancies in **Table 1:** *Descriptive statistics*. The discrepancies generated from germany.do are highlighted in Table 1 below. These discrepancies do not impact the reported data analysis in the paper.

Table 1.A
Descriptive statistics
All Bavaria

VARIABLES	(1) N	(2) mean	(3) sd
Property crime (annual, per 100,000 population)	25.00	298.28	41.84
Ln(Property crime)	25.00	5.69	0.14
Violent crime (annual, per 100,000 population)	25.00	54.88	16.91
Ln(Violent crime)	25.00	3.97	0.26
Murders (annual, per 100,000 population)	25.00	8.16	1.18
Ln(Murders)	25.00	2.09	0.13
Total crime (annual, per 100,000 population)	25.00	417.68	61.35
Ln(Total crime)	25.00	6.02	0.14
Begging, vagrancy arrests (annual, per 100,000 population)	25.00	1,934.56	527.54
Ln(Begging, vagrancy arrests)	25.00	7.54	0.25
Rye prices (Florins per bushel)	25.00	13.59	5.00
Ln(Rye prices)	25.00	2.55	0.37
Real wage (1910 = 100)	25.00	49.72	7.65
Ln(Real wage)	25.00	3.90	0.15
Rainfall, year t	25.00	687.91	70.16
Rainfall, year t-1	25.00	670.44	81.81
Rainfall, year t-2	25.00	671.65	82.02

Note: last three rows divided by a thousand in the published paper

Table 1.B
Descriptive statistics
Bavarian states

VARIABLES	(1) N	(2) mean	(3) sd
Property crime (annual, per 100,000 population)	179.00	290.70	78.82
Ln(Property crime)	179.00	5.64	0.26
Violent crime (annual, per 100,000 population)	179.00	54.44	22.57
Ln(Violent crime)	179.00	3.93	0.36
Murders (annual, per 100,000 population)	179.00	8.04	2.39
Ln(Murders)	179.00	2.04	0.28
Total crime (annual, per 100,000 population)	179.00	407.75	92.83
Ln(Total crime)	179.00	5.99	0.22
Begging, vagrancy arrests (annual, per 100,000 population)	179.00	1,743.75	676.08
Ln(Begging, vagrancy arrests)	179.00	7.37	0.48
Rye prices (Florins per bushel)	179.00	13.57	4.86
Ln(Rye prices)	179.00	2.54	0.36
Real wage (1910 = 100)	179.00	49.86	7.49
Ln(Real wage)	179.00	3.90	0.15
Rainfall, year t	179.00	682.69	76.45
Rainfall, year t-1	179.00	671.11	79.60
Rainfall, year t-2	179.00	671.63	79.68

Note: last three rows divided by a thousand in the published paper

In the published version of Table 1:
- Column 1, Row 9: published Mean is 1934, while reproduced value is 1935
- Column 4, Row 1: published Mean is 290, while reproduced value is 291

The data replication generated minor discrepancies in **Table 2:** *First-stage and reduced-form results,* highlighted below. There are also several minor typos in the published version. These discrepancies and typos do not impact the reported data analysis in the paper.

Table 2.A
First-stage and reduced-form results

VARIABLES	(1) ln(Rye price) OLS	(2) ln(Rye price) OLS	(3) ln(Rye price) OLS
Rainfall, year t	1.04	0.61	0.83**
	(0.83)	(0.51)	(0.40)
Rainfall, year t-1	1.95**	1.71***	1.66***
	(0.82)	(0.47)	(0.34)
Rainfall, year t-2	1.20	1.08**	1.02**
	(0.82)	(0.49)	(0.42)
Time trend		-1.90	-1.99
		(50.07)	(43.33)
(Time trend)^2		5.61	5.63*
		(3.51)	(3.04)
Post-1847		-4,195.15**	-4,145.96***
		(1,607.31)	(1,389.52)
Time trend*Post-1847		455.32**	452.41***
		(166.76)	(144.18)
(Time trend)^2*Post-1847		-16.01***	-16.03***
		(5.09)	(4.41)
Observations	25	25	179
R-squared	0.28	0.84	0.83
Root MSE	0.333	0.180	0.150
F stat	2.361	7.885	14.61

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Controls (Time trend, (Time trend)^2, Post-1847, Time trend*Post-1847, (Time trend)^2*Post-1847) divided by a thousand in the published paper. See Table 1 for "Mean of dependent variables" – there is a typo for the first three columns (where ln(Rye price) is the dependent variable). Mean should be 2.55 instead of 2.54

Table 2.B
First-stage and reduced-form results

VARIABLES	(1) ln(Real wage) OLS	(2) ln(Property crime) OLS	(3) ln(Violent crime) OLS
Rainfall, year t	-0.05	-0.16	-0.36
	(0.29)	(0.33)	(0.27)
Rainfall, year t-1	-0.41*	0.38*	-0.66**
	(0.22)	(0.20)	(0.28)
Rainfall, year t-2	-0.33	0.42	-0.50
	(0.25)	(0.36)	(0.30)
Time trend	-21.79	-59.00*	38.63
	(29.70)	(33.49)	(28.35)
(Time trend)^2	0.68	5.91**	-3.06
	(2.26)	(2.43)	(2.03)
Post-1847	3,208.22***	-340.68	1,937.86
	(1,014.24)	(734.82)	(1,261.93)
Time trend*Post-1847	-308.87***	85.09	-257.07*
	(106.95)	(83.68)	(131.18)
(Time trend)^2*Post-1847	7.31**	-6.20*	9.77**
	(3.37)	(3.09)	(3.72)
Observations	25	25	25
R-squared	0.71	0.51	0.87
Root MSE	0.101	0.119	0.114
F stat	1.749	1.308	4.327

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Controls (Time trend, (Time trend)^2, Post-1847, Time trend*Post-1847, (Time trend)^2*Post-1847) divided by a thousand in the published paper. See Table 1 for "Mean of dependent variables"

In the published version of Table 2:
- Column 2: published r-squared value is 0.83, while reproduced value is 0.84
- Column 3, Row 6: published RSE is 0.40, while reproduced RSE is 0.42
- Column 3, Row 9: published coefficient is not significant, while reproduced value is significant at 90% confidence
- Column 3: published F-stat is 14.1, while reproduced value is 14.6
- Column 4, Row 13: published coefficient is significant at the 95% level, while reproduced value is significant at 99% confidence
- Column 4, Row 15: published coefficient is not significant, while reproduced value is significant at 95% confidence
- TYPO -- Column 4, Row 9: coefficient should be 0.0007 rather than -0.0007
- TYPO -- Columns 1-3: mean of dependent variable is 2.55 rather than 2.54 (see reproduced values in Table 1)

- TYPO -- label at the bottom of the Table should say "* Significantly different than zero at 90% confidence" (currently reads: * Significantly different than zero at 99% confidence)

The data replication generated one minor discrepancy in **Table 3:** *Poverty and property crimes,* highlighted below, which does not impact the reported data analysis in the paper.

Table 3.A
Poverty and property crimes

VARIABLES	(1) ln(Property crime) OLS	(2) ln(Property crime) OLS	(3) ln(Property crime) IV-2SLS	(4) ln(Property crime) IV-2SLS
Ln(Rye prices)	0.29***	0.41***	0.20**	0.22**
	(0.05)	(0.08)	(0.09)	(0.09)
Observations	25	25	25	179
R-squared	0.60	0.78	0.69	0.68
Time controls	No	Yes	Yes	Yes
District fixed effects	No	No	No	Yes
Root MSE	0.0893	0.0752	0.0887	0.150

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: See Table 1 for “Mean of dependent variables”

Table 3.B
Poverty and property crimes

VARIABLES	(1) ln(Property crime) OLS	(2) ln(Property crime) OLS	(3) ln(Property crime) IV-2SLS	(4) ln(Property crime) IV-2SLS
Ln(Real wage)	-0.46***	-0.88***	-1.00**	-0.93**
	(0.15)	(0.14)	(0.47)	(0.36)
Observations	25	25	25	179
R-squared	0.26	0.76	0.76	0.72
Time controls	No	Yes	Yes	Yes
District fixed effects	No	No	No	Yes
Root MSE	0.121	0.0776	0.0784	0.142

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: See Table 1 for "Mean of dependent variables"

In published version of Table 3:

- Column 6, Row 3 (which in reproduced table is Column 2, Row 1 of Table 3.B): coefficient is significantly different than zero at 99%, rather than 95% confidence

There were no discrepancies found in published version of **Table 4**, *Poverty and major crime categories.*

Table 4
Poverty and major crime categories

VARIABLES	(1) Ln(Begging vagrancy) IV-2SLS	(2) ln(Property crime) IV-2SLS	(3) ln(Violent crime) IV-2SLS	(4) ln(Murder) IV-2SLS	(5) ln(Total crime) IV-2SLS
Ln(Rye prices)	0.60***	0.20**	-0.42***	-0.31**	0.08
	(0.11)	(0.09)	(0.08)	(0.14)	(0.08)
Observations	25	25	25	25	25
R-squared	0.94	0.69	0.95	0.63	0.80
Time controls	Yes	Yes	Yes	Yes	Yes
Root MSE	0.0723	0.0887	0.0661	0.0939	0.0751

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: See Table 1 for "Mean of dependent variables"

The data replication generated one minor discrepancy in **Table 5:** *Poverty and violent crimes and murder,* highlighted below, which does not impact the reported data analysis in the paper.

Table 5.A
Poverty and violent crimes and murder

VARIABLES	(1) ln(Violent crime) OLS	(2) ln(Violent crime) IV-2SLS	(3) ln(Violent crime) IV-2SLS
Ln(Rye prices)	-0.48***	-0.42***	-0.37***
	(0.05)	(0.08)	(0.08)
Observations	25	25	179
R-squared	0.95	0.95	0.75
Time controls	Yes	Yes	Yes
District fixed effects	No	No	Yes
Root MSE	0.0650	0.0661	0.186

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: See Table 1 for "Mean of dependent variables"

Table 5.B
Poverty and violent crimes and murder

VARIABLES	(1) ln(Murder) OLS	(2) ln(Murder) IV-2SLS	(3) ln(Murder) IV-2SLS
Ln(Rye prices)	-0.25**	-0.31**	-0.25**
	(0.10)	(0.14)	(0.12)
Observations	25	25	179
R-squared	0.64	0.63	0.52
Time controls	Yes	Yes	Yes
District fixed effects	No	No	Yes
Root MSE	0.0930	0.0939	0.203

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: See Table 1 for "Mean of dependent variables"

In the published version of Table 5:
- Column 4, Row 5: published r-squared value is 0.63, while reproduced value is 0.64

The data replication generated minor discrepancies in **Table A.1:** *Poverty and property crimes—using only lagged rainfall as instruments,* highlighted below. There are also one minor typo in the published version. These discrepancies and typo do not impact the reported data analysis in the paper.

Appendix Table A.1.A
Poverty and property crimes—using only lagged rainfall as instruments

VARIABLES	(1) ln(Property crime) OLS	(2) ln(Property crime) IV-2SLS	(3) ln(Property crime) IV-2SLS
Ln(Rye prices)	0.41***	0.26**	0.24**
	(0.08)	(0.10)	(0.11)
Observations	25	25	179
R-squared	0.78	0.73	0.69
Time controls	Yes	Yes	Yes
District fixed effects	No	No	Yes
Root MSE	0.0752	0.0822	0.149

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: See Table 1 for "Mean of dependent variables"

Appendix Table A.1.B
Poverty and property crimes—using only lagged rainfall as instruments

VARIABLES	(1) ln(Property crime) OLS	(2) ln(Property crime) IV-2SLS	(3) ln(Property crime) IV-2SLS
Ln(Real wage)	-0.88***	-1.05*	-1.06*
	(0.14)	(0.50)	(0.58)
Observations	25	25	179
R-squared	0.76	0.75	0.72
Time controls	Yes	Yes	Yes
District fixed effects	No	No	Yes
Root MSE	0.0776	0.0794	0.142

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: See Table 1 for "Mean of dependent variables"

In the published version of the Appendix Table A.1
- Column 2, Row 2: published RSE is 0.09, while reproduced RSE is 0.10
- Column 5, Row 8: published root MSE is 0.12, while reproduced root MSE is 0.08
- TYPO -- label at the bottom of the Table should say "* Significantly different than zero at 90% confidence" (currently reads: * Significantly different than zero at 99% confidence)