Econometrics Assignment 2a

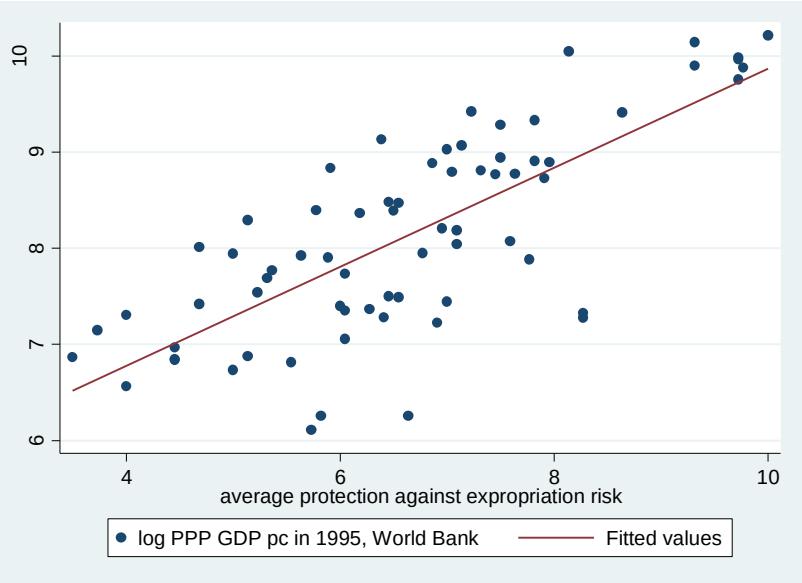
Joost Bouten, SNR: 1265889

Twan Vissers, SNR: 1266283

Fons Strik, SNR: 1257943

III

a.



b.

Source	SS	df	MS	Number of obs	=	70
				F(1, 68)	=	90.69
Model	44.2500667	1	44.2500667	Prob > F	=	0.0000
Residual	33.1773828	68	.487902688	R-squared	=	0.5715
				Adj R-squared	=	0.5652
Total	77.4274494	69	1.12213695	Root MSE	=	.6985

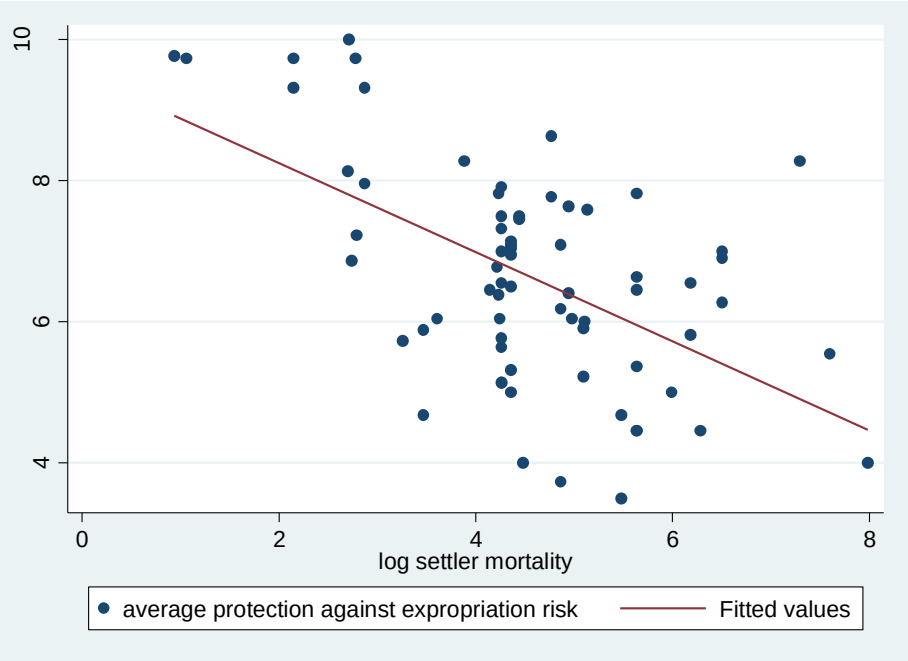
logpgp95	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
avexpr	.5158123	.0541628	9.52	0.000	.4077321	.6238926
_cons	4.713454	.3695264	12.76	0.000	3.976075	5.450832

According to the regression, avexpr positively affects logpgp95 with a coefficient value of 0.5158. This suggest that if the average protection against expropriation risk increases by a unit of 1, PPP GDP per capita in 1995 increases by 51.58%.

c.

Countries with a relatively high GPD per capita self-select into the treatment of generally having good institutions. To illustrate: if a country has favorable geographical characteristics it is generally richer and has better institutions. A problem is that we first need to find out whether the direction of causality is indeed from the treatment to the outcome.

d.



e.

Source	SS	df	MS	Number of obs	=	70
				F(1, 68)	=	66.27
Model	38.2141393	1	38.2141393	Prob > F	=	0.0000
Residual	39.2133101	68	.576666325	R-squared	=	0.4935
				Adj R-squared	=	0.4861
Total	77.4274494	69	1.12213695	Root MSE	=	.75939

logpgp95	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
logem4	-.5483295	.0673584	-8.14	0.000	-.6827411	-.4139179
_cons	10.63256	.3191752	33.31	0.000	9.99566	11.26947

From these results we can see that the coefficient of the reduced form is negative with a value of -0.5483. This coefficient is sufficiently strong as it is significant at the 1% level. As we can see a strong effect in the reduced form, it 'looks good'.

f.

Source	SS	df	MS	Number of obs	=	70
				F(1, 68)	=	29.80
Model	50.6746952	1	50.6746952	Prob > F	=	0.0000
Residual	115.639953	68	1.70058755	R-squared	=	0.3047
				Adj R-squared	=	0.2945
Total	166.314648	69	2.41035722	Root MSE	=	1.3041

avexpr	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
logem4	-.6314299	.1156722	-5.46	0.000	-.8622502	-.4006096
_cons	9.514584	.5481083	17.36	0.000	8.420851	10.60832

We find that the coefficient of the instrument ("logem4") on the independent variable ("avexpr") is negative with a value of -0.6314. The t statistic has a value of -5.46, suggesting that the instrument is sufficiently strong at the 1% level ($P > |t| = 0.000$). This means that the log of settler mortality is a strong predictor of the average protection against expropriation risk.

g.

Source	SS	df	MS	Number of obs	=	70
				F(1, 68)	=	66.27
Model	38.2141398	1	38.2141398	Prob > F	=	0.0000
Residual	39.2133096	68	.576666318	R-squared	=	0.4935
				Adj R-squared	=	0.4861
Total	77.4274494	69	1.12213695	Root MSE	=	.75939

logpgp95	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
avexpr_hat	.8683933	.106676	8.14	0.000	.6555247	1.081262
_cons	2.370164	.714766	3.32	0.001	.9438703	3.796457

We now find that the effect of the average protection against expropriation risk, predicted using the instrumental variable, on the dependent variable (logpgp95) is estimated to be positive at a value of 0.8684. This value again seems to be significant at the 1% level. This coefficient seems to be greater than the coefficient estimated under question (b) (0.8684>0.5158), this indicates that the effect is underestimated when using a simple OLS regression compared to an IV regression.

h.

Source	SS	df	MS	Number of obs	=	70
				F(2, 67)	=	19.53
Model	61.2431764	2	30.6215882	Prob > F	=	0.0000
Residual	105.071472	67	1.56823093	R-squared	=	0.3682
				Adj R-squared	=	0.3494
Total	166.314648	69	2.41035722	Root MSE	=	1.2523

avexpr	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
logem4	-.4537058	.1304823	-3.48	0.001	-.7141496	-.1932619
lat_abst	3.125466	1.203964	2.60	0.012	.7223438	5.528588
_cons	8.094943	.7590112	10.67	0.000	6.57995	9.609936

Source	SS	df	MS	Number of obs	=	70
				F(2, 67)	=	36.61
Model	40.4326897	2	20.2163448	Prob > F	=	0.0000
Residual	36.9947598	67	.552160593	R-squared	=	0.5222
				Adj R-squared	=	0.5079
Total	77.4274494	69	1.12213695	Root MSE	=	.74308

logpgp95	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
avexpr_hat	.739435	.122618	6.03	0.000	.4946884	.9841816
lat_abst	1.432	.7143996	2.00	0.049	.0060527	2.857948
_cons	2.94671	.7562474	3.90	0.000	1.437234	4.456186

The new estimation results show that the effect of the treatment on the outcome variable reduces when including latitude in the regression. This decrease in estimated effect of the treatment variable indicates an overestimation when not including latitude in the regression.

The point of including this covariate is to reduce the omitted variable bias, latitude (through climate and soil conditions) can affect both the outcome variable as well as the treatment variable.

i.

First-stage regressions

Source	SS	df	MS	Number of obs	=	70
				F(2, 67)	=	19.53
Model	61.2431764	2	30.6215882	Prob > F	=	0.0000
Residual	105.071472	67	1.56823093	R-squared	=	0.3682
				Adj R-squared	=	0.3494
Total	166.314648	69	2.41035722	Root MSE	=	1.2523

avexpr	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lat_abst	3.125466	1.203964	2.60	0.012	.7223438	5.528588
logem4	-.4537058	.1304823	-3.48	0.001	-.7141496	-.1932619
_cons	8.094943	.7590112	10.67	0.000	6.57995	9.609936

Instrumental variables (2SLS) regression

Source	SS	df	MS	Number of obs	=	70
				F(2, 67)	=	18.75
Model	5.18804532	2	2.59402266	Prob > F	=	0.0000
Residual	72.2394041	67	1.07820006	R-squared	=	0.0670
				Adj R-squared	=	0.0392
Total	77.4274494	69	1.12213695	Root MSE	=	1.0384

logpgp95	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
avexpr	1.029084	.2384636	4.32	0.000	.5531086	1.505059
lat_abst	-1.784366	1.527363	-1.17	0.247	-4.832995	1.264263
_cons	1.65175	1.35228	1.22	0.226	-1.047413	4.350913

Instrumented: avexpr
Instruments: lat_abst logem4

j.

The condition of the exclusion restriction that the instrument may only affect the outcome variable through the treatment variable and not via any other mechanism. We think that it is highly likely that the exclusion restriction will hold, as we believe that is highly unlikely that settler mortality still affects current GDP per capita in any other way than through institutions.

Project: Econometrics Assignment 2a

```
      name:  <unnamed>
       log:  C:\Users\u1265889\Desktop\2a.smcl
  log type:  smcl
 opened on:   7 Sep 2017, 17:16:45

 1 . do "C:\Users\u1265889\AppData\Local\Temp\STD01000000.tmp"

 2 . * Computer Assignment 2 Econometrics, Sep 2017
 3 . *I
 4 . use C:\Users\u1265889\Downloads\AJR_140915, clear

 5 . *IIa
 6 . summ logpgp95 avexpr logem4 lat_abst

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
    logpgp95 |         70    8.141596     1.05931   6.109248   10.21574
      avexpr |         70    6.646104    1.552533        3.5         10
      logem4 |         70    4.542832    1.357206   .9360933   7.986165
    lat_abst |         70    .1958971    .1470903          0   .6666667

 7 . *IIb
 8 . histogram logpgp95
   (bin=8, start=6.1092477, width=.51331156)

 9 . histogram avexpr
   (bin=8, start=3.5, width=.8125)

10 . *IIIa
11 . graph twoway scatter logpgp95 avexpr

12 . graph twoway lfit logpgp95 avexpr

13 . graph twoway (scatter logpgp95 avexpr) (lfit logpgp95 avexpr)

14 . *IIIb
15 . reg logpgp95 avexpr

      Source |       SS           df       MS      Number of obs   =        70
-------------+----------------------------------   F(1, 68)        =     90.69
       Model |  44.2500667         1  44.2500667   Prob > F        =    0.0000
    Residual |  33.1773828        68  .487902688   R-squared       =    0.5715
-------------+----------------------------------   Adj R-squared   =    0.5652
       Total |  77.4274494        69  1.12213695   Root MSE        =    .6985

------------------------------------------------------------------------------
    logpgp95 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      avexpr |   .5158123   .0541628     9.52   0.000     .4077321    .6238926
       _cons |   4.713454   .3695264    12.76   0.000     3.976075    5.450832
------------------------------------------------------------------------------

16 . *EXPLANATION
17 . *IIIc
18 . *EXPLANATION
19 . *IIId
20 . graph twoway (scatter avexpr logem4) (lfit avexpr logem4)

21 . *IIIe
```

```
22 . reg logpgp95 logem4
```

Source	SS	df	MS	Number of obs	=	70
				F(1, 68)	=	66.27
Model	38.2141393	1	38.2141393	Prob > F	=	0.0000
Residual	39.2133101	68	.576666325	R-squared	=	0.4935
				Adj R-squared	=	0.4861
Total	77.4274494	69	1.12213695	Root MSE	=	.75939

logpgp95	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
logem4	-.5483295	.0673584	-8.14	0.000	-.6827411	-.4139179
_cons	10.63256	.3191752	33.31	0.000	9.99566	11.26947

```
23 . *EXPLANATION
24 . *IIIf
25 . reg avexpr logem4
```

Source	SS	df	MS	Number of obs	=	70
				F(1, 68)	=	29.80
Model	50.6746952	1	50.6746952	Prob > F	=	0.0000
Residual	115.639953	68	1.70058755	R-squared	=	0.3047
				Adj R-squared	=	0.2945
Total	166.314648	69	2.41035722	Root MSE	=	1.3041

avexpr	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
logem4	-.6314299	.1156722	-5.46	0.000	-.8622502	-.4006096
_cons	9.514584	.5481083	17.36	0.000	8.420851	10.60832

```
26 . *EXPLANATION
27 . *IIIg
28 . predict avexpr_hat
   (option xb assumed; fitted values)

29 . reg logpgp95 avexpr_hat
```

Source	SS	df	MS	Number of obs	=	70
				F(1, 68)	=	66.27
Model	38.2141398	1	38.2141398	Prob > F	=	0.0000
Residual	39.2133096	68	.576666318	R-squared	=	0.4935
				Adj R-squared	=	0.4861
Total	77.4274494	69	1.12213695	Root MSE	=	.75939

logpgp95	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
avexpr_hat	.8683933	.106676	8.14	0.000	.6555247	1.081262
_cons	2.370164	.714766	3.32	0.001	.9438703	3.796457

```
30 . *EXPLANATION
31 . *IIIh
32 . reg avexpr logem4 lat_abst
```

Source	SS	df	MS	Number of obs	=	70
				F(2, 67)	=	19.53
Model	61.2431764	2	30.6215882	Prob > F	=	0.0000
Residual	105.071472	67	1.56823093	R-squared	=	0.3682
				Adj R-squared	=	0.3494
Total	166.314648	69	2.41035722	Root MSE	=	1.2523

avexpr	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
logem4	-.4537058	.1304823	-3.48	0.001	-.7141496	-.1932619
lat_abst	3.125466	1.203964	2.60	0.012	.7223438	5.528588
_cons	8.094943	.7590112	10.67	0.000	6.57995	9.609936

```
33 . reg logpgp95 avexpr_hat lat_abst
```

Source	SS	df	MS		
Model	40.4326897	2	20.2163448	Number of obs	= 70
Residual	36.9947598	67	.552160593	F(2, 67)	= 36.61
				Prob > F	= 0.0000
Total	77.4274494	69	1.12213695	R-squared	= 0.5222
				Adj R-squared	= 0.5079
				Root MSE	= .74308

logpgp95	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
avexpr_hat	.739435	.122618	6.03	0.000	.4946884	.9841816
lat_abst	1.432	.7143996	2.00	0.049	.0060527	2.857948
_cons	2.94671	.7562474	3.90	0.000	1.437234	4.456186

```
34 . *EXPLANATION
35 . *IIIi
36 . ivreg logpgp95 lat_abst (avexpr=logem4), first
```

First-stage regressions

Source	SS	df	MS		
Model	61.2431764	2	30.6215882	Number of obs	= 70
Residual	105.071472	67	1.56823093	F(2, 67)	= 19.53
				Prob > F	= 0.0000
Total	166.314648	69	2.41035722	R-squared	= 0.3682
				Adj R-squared	= 0.3494
				Root MSE	= 1.2523

avexpr	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lat_abst	3.125466	1.203964	2.60	0.012	.7223438	5.528588
logem4	-.4537058	.1304823	-3.48	0.001	-.7141496	-.1932619
_cons	8.094943	.7590112	10.67	0.000	6.57995	9.609936

Instrumental variables (2SLS) regression

Source	SS	df	MS		
Model	5.18804532	2	2.59402266	Number of obs	= 70
Residual	72.2394041	67	1.07820006	F(2, 67)	= 18.75
				Prob > F	= 0.0000
Total	77.4274494	69	1.12213695	R-squared	= 0.0670
				Adj R-squared	= 0.0392
				Root MSE	= 1.0384

logpgp95	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
avexpr	1.029084	.2384636	4.32	0.000	.5531086	1.505059
lat_abst	-1.784366	1.527363	-1.17	0.247	-4.832995	1.264263
_cons	1.65175	1.35228	1.22	0.226	-1.047413	4.350913

Instrumented: avexpr
Instruments: lat_abst logem4

```
37 . *IIIj
38 . *EXPLANATION
39 .
   end of do-file

40 . log close
          name:  <unnamed>
           log:  C:\Users\u1265889\Desktop\2a.smcl
     log type:  smcl
    closed on:   7 Sep 2017, 17:17:20
```