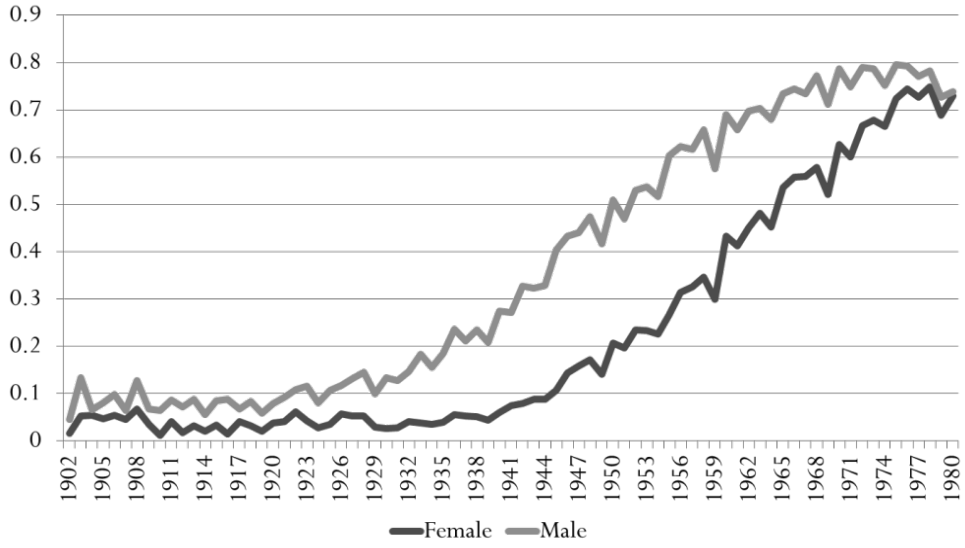
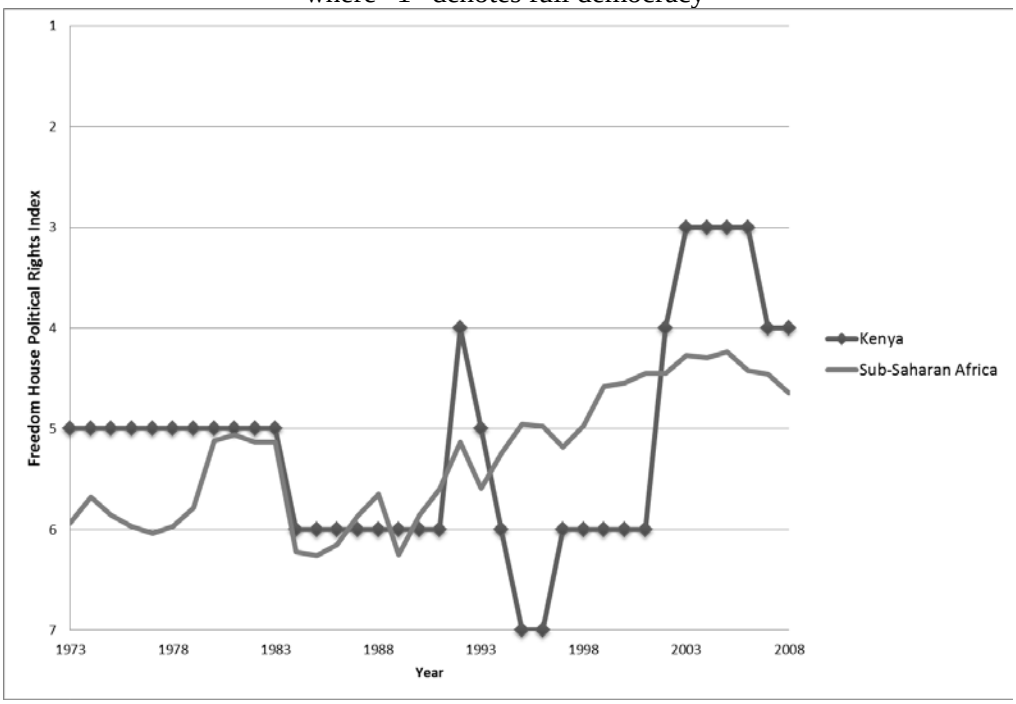
SUPPLEMENTARY APPENDIX MATERIALS [not intended for publication]

Appendix Figure A1: Schooling and democracy trends in Kenya

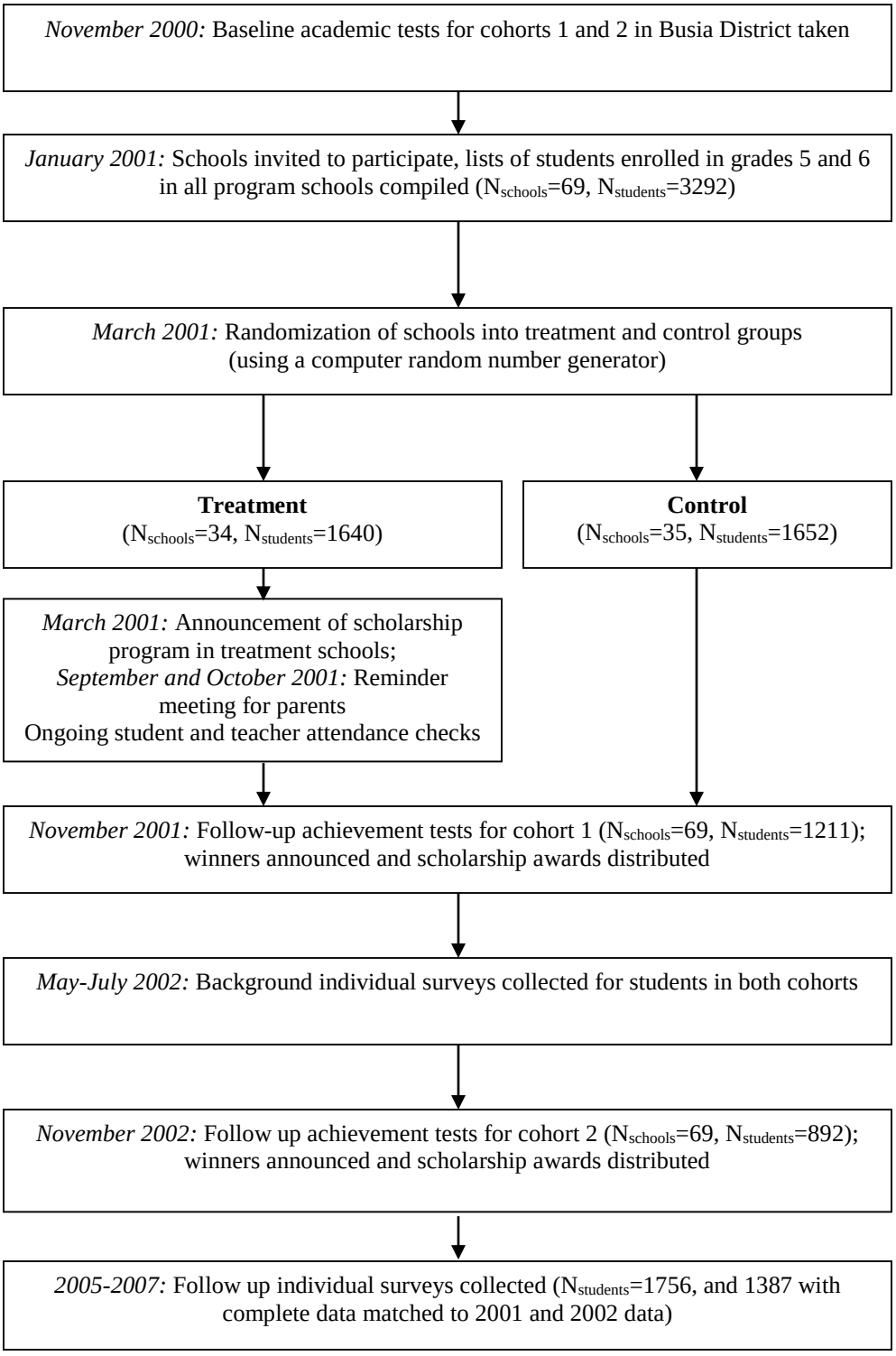
Panel A: Proportion of population completing primary schooling in Kenya, by birth cohort and gender (1902-1980 birth cohorts, source: 1999 Kenya Census)



Panel B: Freedom House Political Rights Index in Kenya and sub-Saharan Africa (1973-2008), where "1" denotes full democracy



Appendix Figure A2: Girls Scholarship Program (GSP) Timeline (2000-2007)



Appendix Table A1: Comparing estimates in the full sample, with and without covariates

Dependent variable:	Coefficient estimate (s.e.) on program indicator OLS (1)	Coefficient estimate (s.e.) on program indicator, without covariates (2)
Lack of autonomy mean effect	-0.181**	-0.199**
	(0.077)	(0.077)
Ethnic identity is not "very important" (0-1)	-0.033*	-0.035*
	(0.020)	(0.020)
Religious identity is not "very important" (0-1)	0.005	0.004
	(0.006)	(0.006)
Democratic attitudes mean effect	0.023	0.041
	(0.098)	(0.096)
Days read a newspaper in last week (0 to 7)	0.134*	0.176**
	(0.074)	(0.081)
Political knowledge mean effect	0.203**	0.228***
	(0.085)	(0.084)
Satisfaction with authority mean effect	-0.239***	-0.251***
	(0.061)	(0.064)
Perceived political efficacy mean effect	0.055	0.043
	(0.066)	(0.067)
Participation in politics and civic affairs mean effect	-0.038	-0.033
	(0.073)	(0.073)
Agree with "It is sometimes necessary to use violence in support of a just cause." (0-1)	0.059**	0.054*
(vs. "The use of violence is never justified in politics.")	(0.029)	(0.029)

Notes: Each cell contains results from a separate regression. Significant at 90% (*), 95% (**), 99% (***) confidence. Details on the mean effect analysis are in the text. Standard errors are clustered by school. The sample is N=1,387 for all dependent variables.
In column 1, the outcome variable is regressed on the GSP program (treatment) indicator, an indicator for student cohort, student age at time of the survey, educational attainment of each parent, and timing of the follow-up survey (coefficient estimates not shown). In column 2, the specification drops the student age at time of the survey, and educational attainment of each parent covariates.

Appendix Table A2: Comparing estimates in the full sample vs. among those with "low" baseline test scores

Dependent variable:	Coefficient estimate (s.e.) on program indicator OLS (1)	Coefficient estimate (s.e.) on program indicator, baseline test score < +2 s.d. (2)
Lack of autonomy mean effect	-0.204**	-0.176*
	(0.088)	(0.091)
Ethnic identity is not "very important" (0-1)	-0.024	-0.024
	(0.024)	(0.025)
Religious identity is not "very important" (0-1)	-0.007	-0.005
	(0.007)	(0.007)
Democratic attitudes mean effect	-0.037	-0.077
	(0.098)	(0.096)
Days read a newspaper in last week (0 to 7)	0.109	0.116
	(0.097)	(0.105)
Political knowledge mean effect	0.109	0.107
	(0.094)	(0.097)
Satisfaction with authority mean effect	-0.181**	-0.147
	(0.089)	(0.088)
Perceived political efficacy mean effect	0.083	0.086
	(0.077)	(0.081)
Participation in politics and civic affairs mean effect	-0.144	-0.138
	(0.096)	(0.098)
Agree with "It is sometimes necessary to use violence in support of a just cause." (0-1) (vs. "The use of violence is never justified in politics.")	0.032	0.029
	(0.036)	(0.037)

Notes: Each cell contains results from a separate regression. Significant at 90% (*), 95% (**), 99% (***) confidence. Details on the mean effect analysis are in the text. Standard errors are clustered by school. The sample in column 1 consists of all individuals with baseline test score data (N=919). In column 2, the sample is restricted to individuals with baseline normalized test scores less than +2 s.d. (N=881).
In columns 1 and 2, the outcome variable is regressed on the GSP program (treatment) indicator, an indicator for student cohort, student age at time of the survey, educational attainment of each parent, and timing of the follow-up survey (coefficient estimates not shown).

Appendix Table A3: Effects in schools with different numbers of predicted GSP winners

	Coefficient estimate (s.e.) on program indicator OLS	Coefficient estimate (s.e.) on interaction between the program indicator and predicted number of winners
Dependent variable:	(1)	(2)
Lack of autonomy mean effect	-0.193**	0.005
	(0.096)	(0.009)
Ethnic identity is not "very important" (0-1)	-0.039**	0.005**
	(0.018)	(0.002)
Religious identity is not "very important" (0-1)	-0.003	-0.0004
	(0.007)	(0.0006)
Democratic attitudes mean effect	-0.040	0.009
	(0.100)	(0.013)
Days read a newspaper in last week (0 to 7)	0.113	-0.004
	(0.077)	(0.007)
Political knowledge mean effect	0.205**	0.006
	(0.095)	(0.009)
Satisfaction with authority mean effect	-0.204***	-0.001
	(0.063)	(0.005)
Perceived political efficacy mean effect	0.070	0.002
	(0.068)	(0.008)
Participation in politics and civic affairs mean effect	-0.022	-0.008
	(0.073)	(0.009)
Agree with "It is sometimes necessary to use violence in support of a just cause." (0-1) (vs. "The use of violence is never justified in politics.")	0.045	0.007**
	(0.027)	(0.003)

Notes: Each row contains results from a separate regression. Significant at 90% (*), 95% (**), 99% (***) confidence. Details on the mean effect analysis are in the text. Standard errors are clustered by school. The sample is N=1,369 for all dependent variables.

The outcome variable is regressed on the GSP program (treatment) indicator, the GSP indicator interacted with the predicted number of GSP winners, an indicator for student cohort, student age at time of the survey, educational attainment of each parent, and timing of the follow-up survey (coefficient estimates not shown). The predicted number of GSP winners is estimated in treatment group schools by regressing the actual number of scholarship winners in the school on various quantiles of the baseline test score distribution for students in that school; the predicted number of winners is then assigned to both treatment and control schools.

Appendix B: Model appendix

This appendix contains some of the algebra and analysis underlying the discussion of the model in section 6. Equation 7 implies that, in the absence of a merit scholarship program (as in our control group), the variance of education will be equal to:

(eqn. A1) $$\sigma_H^2 = \frac{\sigma_B^2 + \beta_1^2 \sigma_R^2}{4}.$$

This implies that the willingness to accept authority for person *i* at time 1 will be:

(eqn. A2) $$R_{1,i} = R_{0,i} + \gamma\left(\frac{B_{0,i}+\beta_1 R_{0,i}+\beta_2 M_i}{2}\right).$$

In the absence of a merit scholarship program, the variance of $R_{1,i}$ (suppressing subscripts) is:

(eqn. A3)
$$\begin{aligned} Var(R_1) \equiv \sigma_{R_1}^2 &= Var(R_0 + \gamma H^*) \\ &= Var(R_0) + \gamma^2 Var(H^*) + 2Cov(R_0, \gamma H^*) \\ &= \sigma_R^2 + \gamma^2\sigma_H^2 + 2Cov(R_0, \gamma\tfrac{B_0+\beta_1 R_0}{2}) \\ &= \sigma_R^2 + \gamma^2\sigma_H^2 + 2[0 + Cov(R_0, \gamma\tfrac{\beta_1 R_0}{2}) \\ &= \sigma_R^2(1+\gamma\beta_1) + \gamma^2\sigma_H^2 \end{aligned}$$

In the absence of a merit scholarship program, the covariance between H^* and R_1 will be:

(eqn. A4)
$$\begin{aligned} Cov(R_1, H^*) &= Cov(R_0 + \gamma\left[\tfrac{B_0+\beta_1 R_0}{2}\right], \tfrac{B_0+\beta_1 R_0}{2}) \\ &= Cov(R_0, \tfrac{B_0}{2}) + Cov(R_0, \tfrac{\beta_1 R_0}{2}) + \gamma Var(\tfrac{B_0+\beta_1 R_0}{2}) \\ &= \tfrac{2\beta_1\sigma_R^2 + \gamma\sigma_B^2 + \gamma\beta_1^2\sigma_R^2}{4} \end{aligned}$$

Taking the difference between equations 8 and 9 in section 6 allows us to estimate the bias, and solve for β_1 in terms of known parameters:

(eqn. A5) $$\gamma_{OLS} - \gamma_{IV} = \frac{2\beta_1\sigma_R^2}{\sigma_B^2 + \beta_1^2\sigma_R^2}$$

This implies β_1 is

(eqn. A6) $$\beta_1 = \frac{2\left(\gamma_{OLS} - \gamma_{IV}\right)\text{var}(H^*)}{\sigma_R^2}$$

It is possible to solve for β_1 based on observed parameters. To do so, we first rewrite equation A3 and then substitute the variance of R_0 into equation A6:

(eqn. A7) $$\sigma_R^2 = \frac{\sigma_{R_1}^2 - \gamma^2\sigma_H^2}{\left(1+\gamma\beta_1\right)}$$

(eqn. A8) $$\beta_1 = \frac{2\left(\gamma_{OLS} - \gamma_{IV}\right)\sigma_H^2}{\frac{\sigma_{R_1}^2 - \gamma^2\sigma_H^2}{\left(1+\gamma\beta_1\right)}}$$

To simplify let $z = \left[\frac{2\left(\gamma_{OLS} - \gamma_{IV}\right)\sigma_H^2}{\sigma_{R_1}^2 - \gamma^2\sigma_H^2}\right]$,

(eqn. A9) $$\beta_1 = z(1+\gamma\beta_1)$$

Solving for β_1 yields:

(eqn. A10) $$\beta_1 = \frac{1}{\left(\frac{1}{z} - \gamma\right)}$$

Re-writing yields equation 10 in the text.