Acknowledgement of Country

I acknowledge the Traditional Owners and their custodianship of the lands on which we meet today.

On behalf of us all, I pay our respects to their Ancestors and their descendants, who continue cultural and spiritual connections to Country.

We recognise their valuable contributions to Australian and global society.

Image: Digital reproduction of *A guidance through time* by Casey Coolwell and Kyra Mancktelow

Lecture 10: Quantile regression

Julie Moschion
University of Queensland

Reading for Lecture 10

- Mostly Harmless Econometrics (Angrist and Pischke): Chapter 7

Introduction

- In a vast majority of papers, applied econometricians focus on averages, i.e. on evaluating the impact of a treatment on the mean of the outcome.
- For example, if a training program raises average earnings enough to offset the costs, we consider it to be successful.
- And to some extent this makes sense: obtaining a good estimate of the average causal effect is hard enough! And if the impacts are positive overall, it can justify investing public money.
- But when we just compare the means of an outcome, we don't know what is happening in the upper and lower quartiles or quintiles or deciles.
- Does the training program increase the earnings of workers in the bottom quartile of the earnings distribution as well as those at the top?

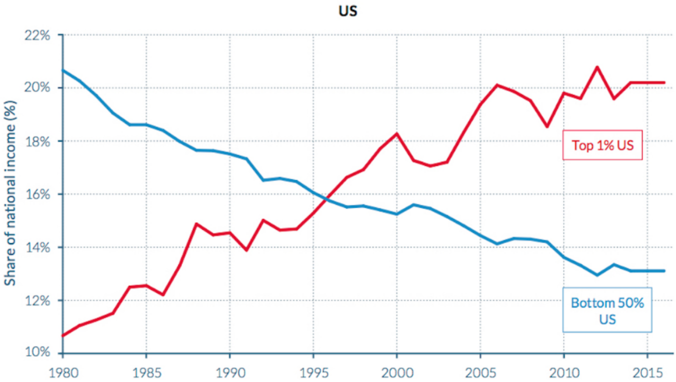
Introduction

- Applied economists and policy makers increasingly want to know what is happening to the entire distribution, i.e. for those at the bottom and those at the top, as well as to the averages.
- There could be winners and losers, i.e. the impact of the program could be positive for some and negative for others.
- This can even happen with the program impact being insignificant overall, i.e. the outcome means being similar before and after the program.
- So a simple examination of averages might not always be enough.
- And many outcomes have continuous distributions and allow this type of analysis: earnings, test scores and even dummy outcomes (e.g. winning a scholarship with the dummy representing the probability of winning).

Introduction

- Policy-makers and labour economists have been especially concerned with changes in the wage distribution since the 1980s, with important implications for inequality.
- And we know that average wages are only a small part of the story.
- Earnings in the upper quantiles have been increasing (a lot), while lower quantiles have not grown as fast.
- In other words, the rich are getting richer and the poor are getting poorer.
- Wage inequality has grown as a result (if you would like to read more about this, have a look at Thomas Piketty's books).

Introduction



The richest 1 per cent exceeded the poorest 50 per cent in the last decades

© World inequality report 2018

Methodology

- Quantile regression is a powerful tool that makes the task of modelling distributions easy, even when the underlying story is complex.
- We can use this tool to see whether participation in a training program or membership in a labor union affects earnings inequality as well as average earnings.
- A quantile regression models the relationship between a set of explanatory variables (X) and specific percentiles (or quantiles) of the dependent variable (Y).
- **The estimated parameter is interpreted as the change in a specified quantile of the dependent variable produced by a one unit change in the explanatory variable.**
- For example, participation in the training program increases the 0.25 quantile (or quartile) of earnings by β dollars.

Methodology

- Assume that we want to estimate the effect of mothers' characteristics (X) on infant birth weight (Y).
- QR allows us to compare how specific quantiles of infant birth weight are affected by mothers characteristics compared to other quantiles.
- This is reflected by the change in the size of the regression coefficient at different quantiles of the outcomes.
- For example, smoking during pregnancy may decrease the 0.25 quantile (or quartile) of birth weight by 200 grams and the 0.75 quantile of birth weight by 50 grams (these are fictitious estimates!).

Methodology

- Estimates: smoking during pregnancy may decrease the 0.25 quantile (or quartile) of birth weight by 200 grams and the 0.75 quantile of birth weight by 50 grams.
- The 0.25 quantile is such that 25% of the sample is below and 75% is above.
- So in the above example, if 25% of non-smoking mothers give birth to babies weighing less than 3kgs $\Rightarrow$ the estimated effect suggests that 25% of smoking mothers give birth to babies weighing less than 2.8kgs.
- How would you interpret the effect on the 0.75 quantile of birth weight (assuming a 0.75 quantile for non-smoking mothers of 4kgs)?
- In comparison, an OLS regression would estimate that smoking during pregnancy may decrease the mean of birth weight by X grams.

Angrist et al. (2006): education & wage distribution

- As we have discussed in a few classes already, applied economists are interested in the returns to education, i.e the effect of one extra year of education on wages.
- This is an important policy question for instance to determine the legal school leaving age or set up incentives to pursue education for specific subgroups (e.g. disengaged youth) if it turns out that this would improve their potential wages.
- As suggested above, 1 more year of education may not mean the same thing for everyone, i.e. it may not translate to the same increase in wages depending on where one is on the wage distribution.

Angrist et al. (2006): education & wage distribution

- Angrist et al. (2006) derive important properties for QR and propose an application estimating the returns to education on the quantiles of wages using US Cenus data in 1980, 1990, and 2000.
- The outcome variable is *wages* and the variable of interest is *years of schooling*.
- They estimate returns to schooling by OLS and by QR for the following quantiles: 0.1, 0.25, 0.5 (median), 0.75, 0.9.

Angrist et al. (2006): education & wage distribution

Table 7.1.1: Quantile regression coefficients for schooling in the 1970, 1980, and 2000 Censuses

		Desc. Stats.		Quantile Regression Estimates					OLS Estimates	
Census	Obs.	Mean	SD	0.1	0.25	0.5	0.75	0.9	Coeff.	Root MSE
				.074	.074	.068	.070	.079	.072	
1980	65023	6.4	0.67	(.002)	(.001)	(.001)	(.001)	(.001)	(.001)	0.63
				.112	.110	.106	.111	.137	.114	
1990	86785	6.46	0.06	(.003)	(.001)	(.001)	(.001)	(.003)	(.001)	0.64
				.092	.105	.111	.120	.157	.114	
2000	97397	6.5	0.75	(.002)	(.001)	(.001)	(.001)	(.004)	(.001)	0.69

Notes: Adapted from Angrist, Chernozhukov, and Fernandez-Val (2006). The tables reports quantile regression estimates of the returns to schooling, with OLS estimates shown at the right for comparison. The sample includes US-born white and black men aged 40-49. Standard errors are reported in parentheses. All models control for race and potential experience. Sampling weights were used for the 2000 Census estimates.

Angrist et al. (2006): education & wage distribution

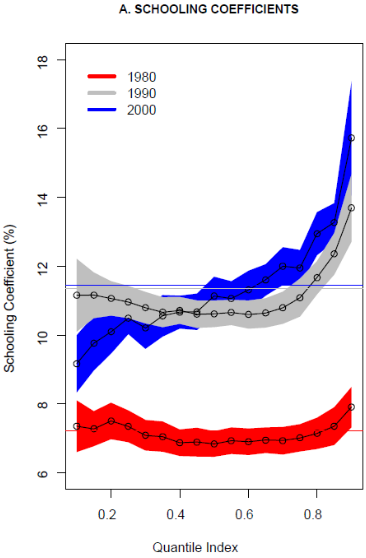
- In 1980, 1 extra year of schooling increases the 0.1 quantile (or decile) of wages by 7.4%.
- The quantile coefficients range from 6.8% to 7.9% with no particular pattern.
- They are not very different from the OLS effect of 7.2% increase in mean wages for an extra year of schooling.
- Overall, 1 extra year of education buys the same increase in wages.

Angrist et al. (2006): education & wage distribution

- In 1990, the returns to schooling increase sharply: from 0.068 to 0.106 at the median and from 0.072 to 0.114 at the mean.
- The largest increase is at the upper decile (from 0.079 to 0.137, almost 6pp), while the increase at all other deciles is around 4pp $\Rightarrow$ 1 extra year of schooling is particularly meaningful at the top of the wage distribution.
- In 2000, while the returns to education continue increasing above the median, they start decreasing below the median.
- An additional year of schooling increases the lower decile of wages by 9.2%, the median by 11.1% and the upper decile by 15.7%.

Angrist et al. (2006): education & wage distribution

This figure shows the coeff of schooling in the QR of log wages on years of schooling, race, and a quadratic function of experience; and robust simultaneous 95% confidence bands. Horizontal lines correspond to OLS estimates.



Abadie & al.(2002): effects of training on earning quantiles

- This paper estimates the impacts of the Job Training Partnership Act (JTPA) program on earning quantiles introducing a combination of QR and IV to estimate quantile treatment effects (QTE).
- This method relies on essentially the same assumptions as for average causal effects (First stage and exogeneity).
- The JTPA was a large federal program that provided subsidized training to American workers facing barriers for employment in the 1980s (e.g. dislocated workers, disadvantaged young adults).
- The program was delivered in 649 areas, also called Service Delivery Areas (SDAs), located throughout the country.
- A non-random subgroup of 16 SDAs were selected to participate in the evaluation. Like in many RCTs external validity may be an issue...

Abadie & al.(2002): effects of training on earning quantiles

- Within participating SDAs, some individuals were randomly assigned to the treatment group and offered training while those assigned to the control group were excluded from the training for 18 months.
- But like in many RCTs some participants declined to participate in the program: only 60% of those assigned to treatment actually got the training.
- However, exclusion from treatment was enforced and those assigned to the control were in the control group.
- This one-sided non compliance results in a selection problem affecting the actual treatment status and making the OLS potentially biased because of OVB.
- There are indeed good reasons to think that factors affecting the decision to refuse training may be correlated with earnings.

Abadie & al.(2002): effects of training on earning quantiles

- As seen previously, in this case the initial allocation to treatment (which is random) can serve as an instrument for the actual training status.
- The two are obviously correlated and the allocation to treatment is independent of potential outcomes.
- The authors evaluate the effect of the training on earnings in the 30 months period after random assignment for 15,981 people.
- The table below shows OLS (panel A) & IV estimates (panel B) for men at the mean and on the following quantiles: 0.15, 0.25, 0.5, 0.75, 0.85.

Abadie & al.(2002): effects of training on earning quantiles

A. OLS and Quantile Regression Estimates

	OLS	Quantile				
		0.15	0.25	0.50	0.75	0.85
Training	3,754	1,187	2,510	4,420	4,678	4,806
	(536)	(205)	(356)	(651)	(937)	(1,055)
% Impact of Training	21.20	135.56	75.20	34.50	17.24	13.43
High school or GED	4,015	339	1,280	3,665	6,045	6,224
	(571)	(186)	(305)	(618)	(1,029)	(1,170)
Black	-2,354	-134	-500	-2,084	-3,576	-3,609
	(626)	(194)	(324)	(684)	(1087)	(1,331)
Hispanic	251	91	278	925	-877	-85
	(883)	(315)	(512)	(1,066)	(1,769)	(2,047)
Married	6,546	587	1,964	7,113	10,073	11,062
	(629)	(222)	(427)	(839)	(1,046)	(1,093)
Worked less than 13	-6,582	-1,090	-3,097	-7,610	-9,834	-9,951
weeks in past year	(566)	(190)	(339)	(665)	(1,000)	(1,099)
Constant	9,811	-216	365	6,110	14,874	21,527
	(1,541)	(468)	(765)	(1,403)	(2,134)	(3,896)

B. 2SLS and QTE Estimates

	2SLS	Quantile				
		0.15	0.25	0.50	0.75	0.85
Training	1,593	121	702	1,544	3,131	3,378
	(895)	(475)	(670)	(1,073)	(1,376)	(1,811)
% Impact of Training	8.55	5.19	11.99	9.64	10.69	9.02
High school or GED	4,075	714	1,752	4,024	5,392	5,954
	(573)	(429)	(644)	(940)	(1,441)	(1,783)
Black	-2,349	-171	-377	-2,656	-4,182	-3,523
	(625)	(439)	(626)	(1,136)	(1,587)	(1,867)
Hispanic	335	328	1,476	1,499	379	1,023
	(888)	(757)	(1,128)	(1,390)	(2,294)	(2,427)
Married	6,647	1,564	3,190	7,683	9,509	10,185
	(627)	(596)	(865)	(1,202)	(1,430)	(1,525)
Worked less than 13	-6,575	-1,932	-4,195	-7,009	-9,289	-9,078
weeks in past year	(567)	(442)	(664)	(1,040)	(1,420)	(1,596)
Constant	10,641	-134	1,049	7,689	14,901	22,412
	(1,569)	(1,116)	(1,655)	(2,361)	(3,292)	(7,655)

Abadie & al.(2002): effects of training on earning quantiles

- **The OLS (panel A) provide a non-causal comparison of the earning distributions of trainees and non-trainees**.
- The OLS training coefficient at the mean suggests that trainees earn on average \$3,754 more than non-trainees (column 1).
- Quantile regression estimates show that all quantiles are significantly larger for trainees than non-trainees.
- For example, the 0.15 quantile is \$1,187 higher for trainees, while the median is \$ 4,420 larger and the 0.85 quantile is \$4,806 larger.
- The gap in quantiles by trainee status is much larger above than below the median.
- but is this because the training is effective or because participants in the training are positively selected?

Abadie & al.(2002): effects of training on earning quantiles

- To address the concern that trainees are self-selected and that the estimated effects may be contaminated by omitted variables bias, **IV estimates are presented in panel B**.
- At the mean, the effect of the program is reduced by more than half at $1,593 with a standard error of $895.
- Looking at QTE estimates (Panel B, columns 2-6), the program only increases quantiles above the median by $3,131 at 0.75.
- Up to the median, the effects of the training on the .15 or .25 quantile are positive but insignificant.
- This demonstrates the importance of looking at effects beyond the mean: if policy makers were trying to improve wages at the bottom they failed but wouldn't have realised if only looking at average effects!

Abadie & al.(2002): effects of training on earning quantiles

- This is also very different from the quantile regression estimates in panel A.
- For example, the QTE estimate (standard error) of the effect on the .15 quantile is \$121 (475), while the corresponding quantile regression estimate is \$1,187 (205).
- Similarly, the QTE estimate (standard error) of the effect on the .25 quantile is \$702 (670), while the corresponding QR estimate is \$2,510 (356).
- This suggests that the quantile regression estimates in the top half of the table (Panel A) are contaminated by positive selection bias.
- In this example the combination of IV + QR was critical to properly identify the impact of the program.

Summary

- QR provides a richer characterisation of the relationship between the treatment and the outcome, allowing us to consider the **impact on the entire distribution of y**, not just on its conditional mean.
- In addition, while OLS can be inefficient if the errors are highly non-normal, **QR is more robust to non-normal errors and to outliers**.

In Stata

- The Stata command **qreg** estimates a multivariate quantile regression with analytic standard errors. By default the quantile is 0.5, the median.
- A different quantile may be specified with the quantile() option.
  - For example a QR of y for quantile 0.1:

$$qreg\ y\ x1\ x2\ x3, quantile(.1)\ vce(robust)$$

  - Or a QR of y for quantile 0.8:

$$qreg\ y\ x1\ x2\ x3, quantile(.8)\ vce(robust)$$

- IV QTE are estimated using the **ivqreg** command.

In Stata

- Although the QR estimator is proven to be asymptotically normal with analytical standard errors, they are awkward to estimate.
- Bootstrap standard errors are often used in place of analytical standard errors.
- Bootstrapping is a way to calculate standard errors by performing inference from drawing N independent subsamples of identical size n (with replacement) from the data.
- The **bsqreg** command estimates the model with bootstrap standard errors, retaining the assumption of independent errors, i.e they are analogous to robust standard errors in linear regression.

Quantiles of the regressor

- Another use of QR is to look at different quantiles of the regressor (the main variable of interest, X) instead of the outcome.
- Here the question is whether the relationship between X and the outcome y is non-linear, i.e. if an increase of 1 unit in X has a different effect on y for low vs high values of X.
- Essentially, you want to identify the effect of X on y at different values for X.
- In practice, you can divide a continuous regressor into 4 quartiles, 5 quintiles or 10 deciles and create dummies for each quantile of the regressor: X_q1 equal to 1 if x is between 0 and $q1$ and 0 otherwise; X_q2 equal to 1 if x is between $q1$ and $q2$ and 0 otherwise...
- When running the regression, use $n-1$ of the quantile dummies instead of the continuous regressor.

Lavy and Schlosser (2011): gender peer effects at school

- This paper is interested in estimating gender peer effects on educational achievements of boys and girls in Israeli primary, middle, and high school.
- They rely on idiosyncratic variations in gender composition across adjacent cohorts within the same schools to identify causal effects: school fixed effects and school specific time trends.
- They find that having more girls in a cohort improves outcomes of both boys and girls.
- They then investigate mechanisms and find that more girls in a cohort has positive effects because it intrinsically changes the classroom environment by decreasing disruption, violence and teachers' fatigue and improving relationships rather than by changing the behaviour of the students themselves.

Lavy and Schlosser (2011): gender peer effects at school

- They use repeated cross-sectional data to estimate a reduced-form equation separately for boys and girls:

$$y_{igst} = \alpha_s + \theta_t + \beta_s year_{st} + \nu_g + X'_{igst}\gamma + Z'_{gst}\zeta + \delta P_{gst} + \epsilon_{igst}$$

- y_{igst}: achievement measures (e.g. average score, matriculation status, number of credit units etc) for student i in school s in grade g and year t;
- α_s, θ_t, ν_g: school, year and grade fixed effects;
- $\beta_s year_{st}$: school specific time trends;

Lavy and Schlosser (2011): gender peer effects at school

$$y_{igst} = \alpha_s + \theta_t + \beta_s year_{st} + \nu_g + X_{igst}\gamma + Z_{gst}\zeta + \delta P_{gst} + \epsilon_{igst}$$

- X_{igst}: vector of student's covariates (e.g. mother's and father's years of schooling, number of siblings, immigration status, ethnic origin, indicators for missing values in these covariates);
- Z_{gst}: vector of characteristics of grade g in school s and time t (e.g. quadratic function of enrollment, and average characteristics of the students in the grade;
- P_{gst}: proportion of girls in grade g, school s, time t (variable of interest);
- standard errors are clustered at the school level.

Lavy and Schlosser (2011): gender peer effects at school

Table 4. Estimates of the Effect of Proportion Female on Scholastic Outcomes in High School

	Main Results						Falsification Tests			
	Females			Males			Females		Males	
	Outcome means	Proportion female in the cohort		Outcome means	Proportion female in the cohort		Prop. female in t-1	Prop. female in t+1	Prop. female in t-1	Prop. female in t+1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Average Score	68.5	5.297	3.857	62.3	6.740	5.555	2.926	-1.995	0.145	-1.516
		(2.178)	(2.460)		(2.656)	(2.835)	(2.244)	(2.655)	(2.591)	(2.525)
Matriculation status	0.610	0.087	0.086	0.511	0.054	0.050	0.031	0.027	0.023	-0.030
		(0.040)	(0.042)		(0.043)	(0.045)	(0.042)	(0.046)	(0.042)	(0.045)
Number of credit units	20.4	1.413	1.135	18.9	1.332	1.154	0.378	-0.035	-0.565	-0.242
		(0.849)	(0.926)		(1.025)	(1.070)	(0.857)	(0.917)	(1.021)	(0.994)
Number of advanced level subjects in science	0.567	0.120	0.125	0.601	0.209	0.209	0.040	0.025	-0.059	-0.039
		(0.069)	(0.068)		(0.071)	(0.072)	(0.072)	(0.065)	(0.072)	(0.069)
Matriculation diploma that meets university requirements	0.549	0.070	0.072	0.460	0.081	0.076	0.015	0.014	0.019	-0.009
		(0.045)	(0.046)		(0.044)	(0.045)	(0.039)	(0.043)	(0.041)	(0.043)
Year effects		✓	✓		✓	✓	✓	✓	✓	✓
School Fixed Effects		✓	✓		✓	✓	✓	✓	✓	✓
School Time Trend		✓	✓		✓	✓	✓	✓	✓	✓
Enrollment (2nd Poly.)		✓			✓		✓	✓	✓	✓
Individual Pupil Controls		✓			✓		✓	✓	✓	✓
Cohort Mean Controls		✓			✓		✓	✓	✓	✓
Number of students	215,442	215,442	215,442	209,696	209,696	209,696	210,925	214,884	205,349	208,864
Number of schools	280	280	280	280	280	280	270	278	270	278

Notes: The table reports means of the dependent variables (columns 1 and 4), estimates for the effects of the proportion of female students in a grade on students achievement in high school (columns 2,3,5, and 6) and falsification tests using the proportion of female students of cohort in t-1 (columns 7 and 9) or in t+1 (columns 8 and 10). Proportion female is measured in 10^{th} grade. Individual controls include: both parents' years of schooling, number of siblings, immigration status, ethnic origin and indicators for missing values in these covariates. Cohort mean controls include students individuals controls averaged by school and year. The regressions include school fixed effects and school specific linear time trends. Robust standard errors clustered at the school level are reported in parenthesis.

Lavy and Schlosser (2011): gender peer effects at school

- The main linear results (left panel) indicate that the proportion of girls tends to improve test scores for boys and girls and has some borderline significant effect on other outcomes.
- These effects are moderate in size: a 100 percentage point increase in the proportion of female peers (i.e. from having none to having only girls) increases average scores of girls by 5.3 points and average scores of boys by 6.7 points.
- They also run some falsification tests (right panel) by estimating the effect of the proportion of girls in subsequent years ($t-1$ and $t+1$) on outcomes in t and find no significant effects on any outcomes.
- This supports the idea that their main estimates are not driven by omitted variable bias or other biases that would generate correlations between cohorts within schools.

Lavy and Schlosser (2011): gender peer effects at school

- To investigate non-linear effects, they compute quintiles of the proportion of girls and replace the single treatment P_{gst} with four quintile indicators.

Table A2. Quintiles of the Proportion Female in High Schools

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
			A. Summary Statistics		
Range	0.000-0.439	0.439-0.499	0.499-0.539	0.539-0.584	0.584-1.000
Mean	0.309	0.473	0.519	0.559	0.648
Median	0.346	0.477	0.519	0.558	0.628
Number of grades	405	400	411	406	406
Number of students	66,413	87,359	100,371	93,586	77,411

- The first quintile includes schools with a proportion of girls in the range [0-0.4390]. The second, third, fourth, and fifth quintiles are defined respectively as: [0.4391-0.4990], [0.4991-0.5389], [0.5390-0.5842], and [0.5843-1].

Lavy and Schlosser (2011): gender peer effects at school

Table 6. Nonlinear Estimates of the Effect of Proportion Female on Matriculation Outcomes and Enrollment in Advanced Math and Science Classes in High School

	Females				Males			
Quintile	II	III	IV	V	II	III	IV	V
Range	0.439-0.499	0.499-0.539	0.539-0.584	0.584-1.000	0.439-0.499	0.499-0.539	0.539-0.584	0.584-1.000
Mean	0.473	0.519	0.559	0.648	0.473	0.519	0.559	0.648
Median	0.477	0.519	0.558	0.628	0.477	0.519	0.558	0.628
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Main matriculation outcomes								
Average Score	-0.065	-0.046	0.055	0.831	0.246	-0.214	0.545	1.379
	(0.549)	(0.593)	(0.598)	(0.651)	(0.467)	(0.495)	(0.529)	(0.652)
Matriculation status	0.000	-0.004	0.002	0.017	0.005	-0.002	0.002	0.015
	(0.009)	(0.010)	(0.010)	(0.011)	(0.008)	(0.008)	(0.009)	(0.011)
Number of credit units	0.128	0.191	0.165	0.452	0.091	-0.024	0.150	0.528
	(0.204)	(0.211)	(0.216)	(0.241)	(0.200)	(0.198)	(0.211)	(0.264)
Number of advanced level subjects in science	0.009	0.023	0.028	0.039	0.016	0.026	0.030	0.051
	(0.015)	(0.017)	(0.017)	(0.019)	(0.014)	(0.015)	(0.015)	(0.020)
Matriculation diploma that meets university requirements	-0.003	-0.005	0.000	0.014	0.001	-0.002	0.004	0.019
	(0.009)	(0.009)	(0.010)	(0.011)	(0.008)	(0.009)	(0.009)	(0.011)
Enrollment in advanced classes								
Math	0.006	0.010	0.014	0.010	0.003	0.007	0.007	0.018
	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)
Physics	0.005	0.004	0.008	0.007	0.008	0.013	0.017	0.015
	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.006)	(0.008)
Computers	0.008	0.009	0.011	0.011	0.011	0.008	0.017	0.012
	(0.005)	(0.007)	(0.006)	(0.006)	(0.008)	(0.009)	(0.009)	(0.011)
Biology	-0.006	-0.005	-0.003	-0.005	-0.001	0.000	0.003	0.007
	(0.006)	(0.007)	(0.007)	(0.008)	(0.003)	(0.004)	(0.004)	(0.006)
Chemistry	0.005	0.007	0.008	0.016	0.005	0.008	0.005	0.016
	(0.005)	(0.005)	(0.006)	(0.007)	(0.004)	(0.005)	(0.005)	(0.007)

Notes: The table reports non-linear effects of the proportion of female students on main matriculation outcomes and students enrollment in advanced classes in Math Science, and English. The model replaces the single treatment variable with a set of quintile indicators for the proportion of female students. The

Lavy and Schlosser (2011): gender peer effects at school

- They re-run their regressions looking at the effects of being in the second, third, fourth, or fifth quintile in the proportion of female students (relative to the first quintile) on the outcomes.
- Most of the effects appear to increase with the quintiles, for both boys and girls.
- The significant effects are mainly concentrated in the fourth and especially in the fifth quintile, where the proportion of female students exceeds 58 percent.
- For example, having more than 58 percent girls (i.e. being in the fifth quintile) compared to being in the first quintile increases average scores of girls by 0.8 points and average scores of boys by 1.4 points.